COURS D'ANALYSE

Laurent Schwartz

Professeur à l'Ecole Polytechnique
et à la Faculté des Sciences de Paris

Analyse

MATHÉMATIQUE

Cours

professé à l'Ecole Polytechnique, Paris

I

Hermann

115 boulevard Saint-Germain Paris VI

TABLE

Chapitre I

THEORIE DES ENSEMBLES

TABLE

Chapitre II

TOPOLOGIE

/

TABLE

Chapitre III

CALCUL DIFFERENTIEL

NOTATIONS

	Paragraphe	Page
$(\vec{X} \mid \vec{Y})$	1	176
$\overleftarrow{E}$	1	178
$\langle \overleftarrow{\alpha}\ \vec{X} \rangle$	1	178
$\overleftarrow{\gamma}_{\vec{Y}}$	1	179
C^m, C^∞	2	187
$Df, D^m f$	2	187
$D_{\vec{X}} f$	3	194
$\overrightarrow{D_{\vec{X}} f(a)} = f'(a) \cdot \vec{X}$	3	198
$\frac{\partial f}{\partial x_1}(a_1, a_2)$	3	207
$\Vert f \Vert_1$	3	213
$\overrightarrow{D_{\vec{X}} D_{\vec{Y}}} f(a) = f''(a)(\vec{X}, \vec{Y})$	6	242
$\binom{p}{q} = C_p^q$	6	249
$p! = p_1!\, p_2! \ldots p_n!$	6	249
$(F^\Omega)_{b;m}$; $(F^\Omega)_{cb;m}$	6	251

INDEX

DEFINITIONS

DERIVEE TOTALE OU APPLICATION DERIVEE.

Soit f une application d'un ouvert Ω d'un espace affine normé E dans un espace affine normé F ; on dit que f admet au point a de Ω, une application dérivée ou dérivée totale ou différentielle totale L, si L est une application linéaire continue de $\vec{E}$ dans $\vec{F}$ et si l'on a pour $a + \vec{h} \in \Omega$:

$$f(a+\vec{h}) = f(a) + L\vec{h} + \varphi(\vec{h})\|\vec{h}\|$$

où $\varphi(\vec{h})$ tend vers 0 lorsque $\vec{h} \neq \vec{0}$ tend vers $\vec{0}$.

DERIVEES SUCCESSIVES.

On définit ainsi par récurrence $f^{(m)}$ en $a \in \Omega$.

Soit $f^{(m-1)}$, la dérivée d'ordre $m-1$, identifiée à une fonction sur Ω à valeurs dans $\mathcal{L}_{m-1}(\overrightarrow{E^{m-1}}, \vec{F})$. Alors $(f^{(m-1)})'(a)$ est un élément de $\mathcal{L}(\vec{E}; \mathcal{L}(\vec{E}^{m-1}; \vec{F}))$ identifié à un élément de $\mathcal{L}_m(\vec{E}^m; \vec{F})$ noté encore $f^{(m)}(a)$ par :

$$f^{(m)}(a).(\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_m) = ((f^{(m-1)})'(a).\vec{X}_1).(\vec{X}_2, \ldots, \vec{X}_m)$$

TABLE

Chapitre IV

CALCUL INTEGRAL

TABLE

Chapitre V

EQUATIONS DIFFERENTIELLES

INDEX

NOTATIONS

PRÉFACE

Comme on l'a souvent dit, il n'est pas de "Mathématiques sans larmes" à l'usage des physiciens et des ingénieurs. Le physicien et l'ingénieur moderne ont besoin d'un énorme volume de connaissances mathématiques, dans les domaines les plus divers. Il n'est absolument plus possible, à ces "utilisateurs", de connaître tous les résultats dont ils ont besoin, avec toutes les démonstrations, conduites avec la rigueur qui est de règle en Mathématiques. On se trouve donc dans la situation suivante. Ou bien on fait un exposé court, parce que contenant peu de résultats solidement démontrés; le Mathématicien y trouvera satisfaction, pas le Physicien. Ou bien encore on fait un exposé court, riche en résultats, mais avec des démonstrations seulement esquissées, sinon absentes; l'esprit cartésien du lecteur en est incommodé. Nous avons adopté une troisième solution. Nous avons fait un cours long, très long même, comportant beaucoup de théorèmes, et des démonstrations généralement complètes. C'est donc plutôt un livre, un document, qu'un cours proprement dit. Les conférences orales n'en donneront qu'un résumé. Les élèves n'auront à apprendre, à titre obligatoire, qu'une partie des feuilles, qui sera chaque fois très précisément spécifiée, et qui comportera beaucoup d'énoncés et peu de démonstrations. Ils devront s'exercer à comprendre les idées et les structures nouvelles qu'ils rencontreront, à connaître les théorèmes et leur esprit, et à savoir les appliquer avec exactitude, feuilles en main. Ce n'est pas aussi facile qu'il peut le paraître; quelqu'un qui n'a jamais réfléchi à un énoncé de théorème est, à coup sûr, incapable de l'appliquer à brûle-pourpoint, même avec l'aide d'un livre ! Seules les démonstrations les plus instructives et les plus caractéristiques seront obligatoires. Mais les élèves pourront, et cela leur est vivement conseillé, étudier une partie des autres à titre facultatif, en choisissant les passages les plus conformes à leur goût, et en suivant les conseils que les Maîtres de Conférences et moi-même ne demanderons qu'à leur donner. Des goûts et des niveaux divers seront ainsi satisfaits; et si tous les Polytechniciens d'une même promotion n'ont pas approfondi exactement les mêmes choses, ce sera tout bénéfice.

Laurent SCHWARTZ

I

THÉORIE DES ENSEMBLES

§ 1 ENSEMBLES. OPÉRATIONS ÉLÉMENTAIRES

On appelle Ensemble une collection d'objets.

Exemples : L'ensemble des élèves d'une promotion;
l'ensemble des points d'un ~~plan~~;
l'ensemble des quadriques non dégénérées d'un espace à 3 dimensions;
l'ensemble $\mathbb{N}$ des entiers ≥ 0, → - de Naturels
l'ensemble $\mathbb{Z}$ des entiers quelconques, → de ZAHLEN
l'ensemble $\mathbb{Q}$ des nombres rationnels, → de Quotient
l'ensemble $\mathbb{R}$ des nombres réels,
l'ensemble $\mathbb{C}$ des nombres complexes.

Parties d'un ensemble

La notation $x \in E$ signifie : " x est (ou étant) élément de l'ensemble E ".* Si un ensemble A est formé d'éléments appartenant à un autre ensemble E, il est appelé partie ou sous-ensemble de cet autre. L'ensemble des ellipsoïdes est une partie de l'ensemble des quadriques.

- Parmi les parties d'un ensemble E, figurent E lui-même, la partie vide notée $\emptyset$, la partie réduite à un élément a, notée $\{a\}$, la partie formée des 3 éléments a, b, c qu'on note $\{a, b, c\}$. Ne pas confondre a, élément de E, et $\{a\}$, partie de E réduite à a. Pour désigner l'ensemble des éléments qui vérifient une certaine propriété P, on emploie souvent la notation $\{x ;\ x \text{ vérifie } P\}$. Par exemple : $\{x ;\ x \in \mathbb{R},\ x \geq 0\}$ veut dire : l'ensemble des nombres réels ≥ 0; $\{\operatorname{Arc tg} x ;\ x \in \mathbb{R},\ e \leq x \leq 2e\}$ veut dire : l'ensemble des nombres $\operatorname{Arc tg} x$, lorsque x prend toutes les valeurs réelles entre e et $2e$ inclusivement.

* Un ensemble s'appelle aussi parfois un espace, ses éléments s'appellent souvent des points.

On note par $\mathfrak{P}(E)$ l'ensemble des parties de E .
C'est un nouvel ensemble qu'on peut former à partir de E .
On peut ensuite considérer $\mathfrak{P}(\mathfrak{P}(E))$, etc.
- Si E a n éléments, $\mathfrak{P}(E)$ en a 2^n *.

Relations d'inclusion, complémentaire

- Si X et Y sont deux parties de E , on dit que X est contenue dans Y ou que Y contient X , si tout élément de X est élément de Y , et on écrit $X \subset Y$ ou $Y \supset X$. Ainsi $\emptyset \subset X \subset E$, $\emptyset \subset \emptyset$, $E \subset E$. Si $X \subset Y$ et $Y \subset Z$, alors $X \subset Z$.

L'ensemble des éléments de E qui n'appartiennent pas à une partie A se nomme complémentaire de A par rapport à E , et s'écrit $\complement_E A$ ou simplement $\complement A$ si aucune confusion n'est à craindre. On a $\complement \complement A = A$. Si $A \subset B$, on a $\complement A \supset \complement B$.
Si A et B sont deux parties de E , et si $A \supset B$, on écrit parfois aussi $A - B$ pour désigner l'ensemble des éléments de A qui n'appartiennent pas à B .

Réunion, Intersection

On appelle réunion d'un ensemble de parties de E la partie de E formée des éléments qui appartiennent au moins à l'une d'entre elles. Pour plusieurs parties nommément désignées A, B, C , on représentera leur réunion par $A \cup B \cup C$.

* On appelle C_n^p , ou mieux $\binom{n}{p}$ le nombre des parties à p éléments d'un ensemble à n éléments.
Alors, en comptant toutes les parties de E (sans oublier $\emptyset$ ni E), on voit que leur nombre est

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n-1} + \binom{n}{n} = 2^n .$$

On peut dire aussi, si $E_n = \{1, 2, \ldots n\}$ que $\mathfrak{P}(E_n)$ comprend, pour $n \geqslant 1$, deux ensembles disjoints, l'ensemble des parties qui contiennent n , et l'ensemble de celles qui ne contiennent pas n ; chacun d'eux a un nombre d'éléments qui est celui de $\mathfrak{P}(E_{n-1})$. Si donc $\mathfrak{P}(E_n)$ à A_n éléments, on a $A_n = 2A_{n-1}$, et comme $A_0 = 1$ (E_0 est l'ensemble vide, il a une partie $\emptyset$), $A_n = 2^n$.

- S'il s'agit d'une famille de parties A_i numérotées à l'aide d'un ensemble d'indices I, on représentera par $\bigcup\limits_{i\in I} A_i$ leur réunion.

- On appelle <u>intersection d'un ensemble de parties de E</u> la partie de E formée des éléments qui appartiennent à la fois à toutes ces parties. On emploiera comme précédemment les notations $A \cap B \cap C$, $\bigcap\limits_{i\in I} A_i$, pour les intersections.

Deux parties sont disjointes si leur intersection est la partie vide.

- Le complémentaire de l'intersection d'une famille de parties est la réunion de leurs complémentaires, et le complémentaire de la réunion d'une famille de parties est l'intersection de leurs complémentaires :

(I, 1; 1) $\complement(A \cap B) = \complement A \cup \complement B$, et $\complement(A \cup B) = \complement A \cap \complement B$.

- La transformation qui, à chaque partie de E , fait correspondre son complémentaire, est donc une correspondance qui transforme $\subset$ en $\supset$, $\supset$ en $\subset$, $\cup$ en $\cap$, $\cap$ en $\cup$.

Ensemble produit

- <u>On appelle produit $E \times F$ de deux ensembles E , F ,</u> l'ensemble de tous les couples ordonnés (x, y) formés d'un élément x de E et d'un élément y de F . On considère ici que le couple (x, y) est différent du couple
2
(y, x) , si $x \neq y$; c'est particulièrement important si l'ensemble E et l'ensemble F sont identiques.

- On peut de même définir le produit de plusieurs ensembles ou d'une famille d'ensembles. On peut d'ailleurs identifier les produits $(E \times F) \times G$, $E \times (F \times G)$, $E \times F \times G$. C'est l'associativité de la multiplication.

- Les produits $E \times E$, $E \times E \times E$ etc, se notent aussi E^2, E^3, etc...

- C'est ainsi que l'on note $\mathbb{R}^n$ le produit de n ensembles identiques à l'ensemble $\mathbb{R}$ des nombres réels; un point de $\mathbb{R}^n$ est donc un système ordonné $(x_1, x_2, \cdots x_n)$ de n nombres réels arbitraires.

§ 2 APPLICATIONS, FONCTIONS

Soit E et F deux ensembles, on appelle application de E dans F ou fonction définie sur E à valeurs dans F toute correspondance f qui, à chaque élément x de E, fait correspondre un élément, noté $f(x)$, de F. La notation $E \xrightarrow{f} F$ signifie que f est une application de E dans F. E s'appelle ensemble initial, F ensemble final de l'application.

- Il y a lieu de distinguer soigneusement f, qui est l'application, et $f(x)$, qui est l'élément correspondant à x par cette application. Cependant cette distinction, pour des raisons pratiques, n'est pas toujours facile à faire dans l'usage courant ! Ainsi il est incorrect (mais commode !) de dire "la fonction $\sin x$", on devrait dire "la fonction $\sin$", alors que $\sin x$ est la valeur de cette fonction au point x. On remédie à cet inconvénient en disant la fonction "$x \to \sin x$," ou "la fonction f définie par $f(x) = \sin x$".
Une application d'un ensemble dans lui-même s'appelle aussi un opérateur.

Exemples d'applications

1 La fonction f, définie par $f(x) = \frac{1}{x}$, est une application de l'ensemble des éléments $\neq 0$ de la droite réelle $\mathbb{R}$, dans $\mathbb{R}$. Il est donc inexact de dire que c'est là une fonction réelle d'une variable réelle, puisque la variable ne peut pas prendre toutes les valeurs réelles. Mais on peut très bien considérer l'application g de $\mathbb{R}$ dans $\mathbb{R}$, définie par : $g(x) = \frac{1}{x}$ pour $x \neq 0$, $g(0) = 2$; cette fonction g est différente de la fonction f, elle n'a pas même ensemble de définition.
On peut aussi considérer l'application h, définie par : $h(x) = \frac{1}{x}$ pour $x \neq 0$, $h(0) = +\infty$, de l'ensemble $E = \mathbb{R}$, dans l'ensemble F formé de $\mathbb{R}$ et d'un élément supplémentaire noté $+\infty$.

2 Si E est l'ensemble des fonctions réelles définies sur un intervalle réel $[a,b]$, et intégrables, l'intégrale : $\varphi \longrightarrow \int_a^b \varphi(x)\,dx$, est une application de E dans la droite réelle $\mathbb{R}$.

3 Si E est l'ensemble des courbes de longueur finie du plan euclidien, on peut définir une application de E dans la demi-droite $\mathbb{R}_+$ (ensemble des éléments $\geqslant 0$ de $\mathbb{R}$), qui, à chaque courbe, fait correspondre sa longueur.

- On appelle <u>application identique</u> d'un ensemble E, l'application f de E dans E, définie par $f(x) = x$.

Si E est une partie d'un ensemble F, on appelle <u>injection canonique de E dans F</u> l'application f de E dans F, définie par $f(x) = x$.

- Si $E \times F$ est un produit de 2 ensembles, on appelle <u>projection sur E</u> l'application de $E \times F$ dans E qui, à chaque couple $(x, y) \in (E \times F)$, fait correspondre l'élément x. On définit de même la <u>projection sur F</u>.

Soient E, F, G 3 ensembles. Une application f de $E \times F$ dans G fait correspondre à tout couple $(x,y) = z \in E \times F$ un élément de G, qui est $f(z) = f((x,y))$. On remplace généralement la double parenthèse par une simple parenthèse : $f(x, y)$, et on dit que f est une fonction de 2 variables. Par exemple, si E est l'ensemble des fonctions réelles d'une variable réelle, intégrables sur tout intervalle fini, l'intégrale $\int_a^b \varphi(x)\, dx$ définit une application de $E \times \mathbb{R} \times \mathbb{R}$ dans $\mathbb{R}$, car elle est fonction de 3 variables, $\varphi \in E$, $a \in \mathbb{R}$, $b \in \mathbb{R}$.

Si maintenant on considère une application de E dans $F \times G$, elle est de la forme $x \longrightarrow (f(x), g(x))$, où f (resp. g) est une application de E dans F (resp. G) : la donnée de cette fonction est équivalente à celle d'un système de 2 fonctions.

Plus généralement, la donnée d'une application d'un produit $E_1 \times E_2 \times \cdots \times E_n$ dans un produit $F_1 \times F_2 \times \cdots \times F_m$ est équivalente a celle d'un système de m fonctions de n variables.

- Une application f de E dans F est dite constante si, pour tous les $x \in E$, $f(x)$ est le même élément de F.

Injections, surjections, bijections

On dit qu'une application f de E dans F est <u>injective</u> ou encore que c'est une <u>injection</u> si 2 éléments distincts de E ont pour images par f deux éléments distincts de F. L'injection canonique d'une partie d'un ensemble dans cet ensemble est bien une injection.

On dit que f est <u>surjective</u>, ou que c'est une <u>surjection</u>, si tout élément de F est l'image par f d'au moins un élément de E.

On dit que f est <u>bijective</u> ou que c'est une <u>bijection</u>, si tout élément de F est l'image par f d'un élément et d'un seul de F? . Une application est bijective si et seulement si elle est à la fois injective et surjective. Une bijection d'un

? ⇒ E !

ensemble sur lui-même s'appelle aussi une permutation ou transformation.

Soit f une bijection, et soit $y \in F$. Appelons $f^{-1}(y)$ l'unique élément x de E tel que $f(x) = y$. Nous venons de définir une application f^{-1} de F dans E. C'est encore une bijection, on dit que c'est l'application réciproque ou la bijection réciproque de f ; on l'appelle souvent improprement fonction inverse.

Image directe et image réciproque d'une partie

- Soit f une application de E dans F, et soit A une partie de E. On appelle $f(A)$ la partie de F formée de tous les éléments $f(x)$, $x \in A$. Evidemment $f(\emptyset) = \emptyset$. On voit que nous venons de définir une application $A \longrightarrow f(A)$ de $\mathcal{P}(E)$ dans $\mathcal{P}(F)$, à partir de l'application f. Cette application conserve les symboles $\subset$, $\supset$, $\cup$, en ce sens que :

$$(I,2;1) \qquad \begin{cases} \text{Si } A \subset B, \text{ on a } f(A) \subset f(B). \\ f(A \cup B) = f(A) \cup f(B). \end{cases}$$

Par contre elle ne conserve pas les symboles $\cap$, $\complement$, et on a seulement

$$(I,2;2) \qquad f(A \cap B) \subset f(A) \cap f(B) \text{ *}.$$

La partie $f(A)$ s'appelle image directe ou image de la partie A par l'application f.

Soit maintenant B une partie de F, on appelle $f^{-1}(B)$ la partie de E formée de tous les x tels que $f(x) \in B$. ** Evidemment $f^{-1}(\emptyset) = \emptyset$ ***. Nous venons ici encore d'associer à f une application $B \longrightarrow f^{-1}(B)$ de $\mathcal{P}(F)$ dans $\mathcal{P}(E)$.

* Soit f une application constante, $f(x) = b$ quel que soit $x \in E$. Soient A, B, deux parties disjointes de E. On a $A \cap B = \emptyset$, donc $f(A \cap B) = \emptyset$, qui est distinct de $f(A) \cap f(B) = \{b\}$.

** Avec la notation abrégée de la page 1, on a $f^{-1}(B) = \{x ; f(x) \in B\}$. $f(A) = \{f(x) ; x \in A\}$.

*** On peut aussi avoir $f^{-1}(B) = \emptyset$ pour $B \neq \emptyset$. Par exemple, si f est l'application $x \longrightarrow x^2$ de $\mathbb{R}$ dans $\mathbb{R}$, $f^{-1}(\{-1\}) = \emptyset$.

Cette application conserve les 5 symboles $\subset, \supset, \cup, \cap, \complement$ en ce sens que l'on a :

$$(I,2;3) \quad \begin{cases} \text{Si } A \subset B, \quad f^{-1}(A) \subset f^{-1}(B). \\ f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B), \quad f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B), \quad f^{-1}(\complement A) = \complement f^{-1}(A). \end{cases}$$

- On remarque donc que l'application f^{-1} ainsi définie est plus simple que l'application f définie plus haut. La partie $f^{-1}(B)$ s'appelle image réciproque de B par l'application f. Il y a lieu de remarquer que cette définition ne suppose nullement que f soit bijective. De toute façon, si $y \in F$, on a le droit de parler de $f^{-1}(\{y\})$, mais c'est une partie de E et non un élément de E ; elle peut comprendre plus d'un élément, si f n'est pas injective, et elle peut être la partie vide, si f n'est pas surjective (*). Si f est bijective, on a exactement $f^{-1}(\{y\}) = \{f^{-1}(y)\}$.

En outre les deux significations possibles du symbole $f^{-1}(B)$ sont les mêmes : c'est l'image réciproque de B par f, ou l'image directe de B par la bijection réciproque f^{-1}. Naturellement, si f est bijective, alors l'image directe conserve aussi les 5 symboles $\subset, \supset, \cup, \cap, \complement$. On a

$$(I,2;3^{bis}) \quad f^{-1}(f(A)) \supset A \text{ pour } A \in \mathfrak{P}(E), \quad f(f^{-1}(B)) \subset B \text{ pour } B \in \mathfrak{P}(F).$$

Ensembles d'applications. Familles, suites

Si E et F sont deux ensembles, on peut donc parler d'un nouvel ensemble, qui est l'ensemble des applications de E dans F. Si E n'a que 2 éléments, cet ensemble admet une bijection sur le carré F^2 ; car une application de E dans F est entièrement déterminée par un couple $(x,y) \in F^2$, le couple des images des deux éléments de E. Si E a n éléments, $a_1, a_2, \cdots a_n$, cet ensemble admet une bijection sur F^n, car il est équivalent de se donner une application de E dans F, ou de se donner le système $(x_1, x_2, \cdots x_n) \in F^n$, ou $x_1, x_2, \ldots x_n$, sont les images par cette application de $a_1, a_2, \cdots a_n$. C'est pourquoi on a l'habitude de noter F^E l'ensemble des applications de E dans F.

Naturellement, on aura à considérer des sous-ensembles de F^E : l'ensemble des applications continues de E dans F, si E et F sont des espaces métriques, l'ensemble des applications linéaires de E dans F, si E et F sont des espaces vectoriels, etc....

On appelle encore famille $(x_i)_{i \in I}$ d'éléments de E indexée par un ensemble d'indices I, une application de I dans E. L'ensemble des familles d'éléments de E, indexée par I, n'est autre que l'ensemble E^I des applications de I dans E.

* Voir (***) page 8.

En particulier, ce qu'on appelle une suite d'éléments de E n'est pas autre chose qu'une famille d'éléments de E, indexée par l'ensemble $\mathbb{N}$ des entiers ≥ 0, ou encore une application de $\mathbb{N}$ dans E. L'ensemble des suites d'éléments de E n est donc autre que $E^{\mathbb{N}}$. En réalité une suite n'est pas toujours indexée par $\mathbb{N}$, et on peut aussi parler d'une suite indexée par l'ensemble $\mathbb{N}_1$ des entiers ≥ 1 ou d'une suite finie indexèe par l'ensemble fini des entiers $1, 2, \dots n$. Il y aura toujours intérêt a préciser le sens du mot "suite", si l'ensemble d'indices n'est pas $\mathbb{N}$

Application composée

- Soient E, F, G 3 ensembles, et soit f une application de E dans F et g une application de F dans G. On appelle application composée $g \circ f$ l'application de E dans G donnée par $g \circ f(x) = g(f(x))$. On remarquera qu'on écrit $g \circ f$ dans l'ordre inverse de celui dans lequel les opérations sont effectuées :

(I, 2 ; 4) $$E \xrightarrow{f} F \xrightarrow{g} G .$$

Z C'est une règle absolue en mathématiques, que, dans une composition d'opérations $g \circ f$, on doit commencer par celle, f, qui est indiquée a droite.

Si A est une partie de E, on a $g \circ f(A) = g(f(A))$. Si B est une partie de G, on a $(g \circ f)^{-1}(B) = f^{-1}(g^{-1}(B))$. La composition des applications est associative: si f, g, h sont des applications de E dans F, F dans G, G dans H respectivement, on a $(h \circ g) \circ f = h \circ (g \circ f)$, qui s'écrit simplement $h \circ g \circ f$.

Si f^{-1} est la bijection réciproque d'une bijection f de E dans F, on a $f^{-1} \circ f = I_E$, application identique de E, et $f \circ f^{-1} = I_F$, application identique de F. Inversement, si f est une application de E dans F, et si g est une application de F dans E, telle que $g \circ f = I_E$, $f \circ g = I_F$, alors f est une bijection, et g est sa bijection réciproque.

Si f est une bijection de E sur F, g une bijection de F sur G, $g \circ f$ est une bijection de E sur G, et sa bijection réciproque est $f^{-1} \circ g^{-1}$.

Soit f une application de E dans F et soit A une partie de E ; on appelle restriction de f a A l'application f_A, souvent notée $f \mid A$, de A dans F, donnée par $f_A(x) = f(x)$ pour $x \in A$

On dit aussi que f est un prolongement à E de l'application f_A de A dans F. Naturellement, le plus souvent, on continuera à écrire f au lieu de f_A

De même si f est une application de E dans F, et si $f(E) \subset B$, f définit une application f_B de E dans B

donnée par $f_B(x) = f(x)$. Pratiquement, on continuera toujours à l'écrire f au lieu de f_B

Changements de variables et changements de fonctions

Soit f une fonction définie sur E , à valeurs dans **F.** Si u est une application d'un ensemble E_1 dans E , on peut définir une nouvelle fonction $f_1 = f \circ u$, définie sur E_1 à valeurs dans F . On dit qu'on a effectué le <u>changement de variable</u> u , ou le changement d'ensemble initial $E_1 \xrightarrow{u} E$, et que f_1 <u>est l'image réciproque de f par ce changement de variable.</u> On obtient l'expression de $f_1(x_1)$, $x_1 \in E_1$, en faisant, dans l'expression de $f(x)$, la substitution $x = u(x_1)$. On écrit aussi $f_1 = u^* f$, ou même f^* s'il n'y a pas lieu d'indiquer le changement de variable u ; ainsi :

$$(I,2;5) \qquad u^* f\,(x_1) = f(u(x_1)), \quad \text{ou} \quad f^*(x_1) = f(u(x_1)).$$

Couramment aussi, par un abus de langage parfois utile mais parfois imprudent et pouvant mener à de graves contradictions, on identifie f_1 et f et on dit que c'est la même fonction (sic !), représentée à l'aide de la variable x_1 au lieu de la variable x .

Si maintenant v est une application de F dans un ensemble F_2 on peut définir une nouvelle fonction, $f_2 = v \circ f$, définie sur E à valeurs dans F_2 . On dit qu'on a effectué le <u>changement de fonction v</u> ou le changement d'ensemble final $F \xrightarrow{v} F_2$,et que <u>f_2 est l'image directe de f</u> par ce changement de fonction.

On peut effectuer à la fois un changement de variable et un changement de fonction et considérer $f_3 = v \circ f \circ u$, f_3 est l'image de f par le changement de variable u et le changement de fonction v mais ce n'est plus ni une image réciproque ni une image directe.

§ 3 RELATIONS D'ÉQUIVALENCE, ENSEMBLE QUOTIENT

On se donne une <u>relation binaire</u> sur un ensemble E , si l'on se donne l'ensemble des couples (x, y) qui vérifient cette relation; ainsi les relations :

$$x = y, \quad x \leq y, \quad x > y, \quad x^2 = y^3, \quad (x^2 + y^2 \quad 81, \quad x \neq 0)$$

sont des relations binaires entre nombres réels. On voit qu'avec précision, une <u>relation binaire sur **E** n'est pas autre chose qu'une partie R du produit $E \times E$.</u>

- Nous étudierons dans ce paragraphe et dans le suivant, des relations binaires d'une importance particulière.

- Une relation binaire R est appelée <u>relation d'équivalence</u> si elle vérifie les trois propriétés suivantes :

(I,3;1) $\begin{cases} \text{a) } \underline{\text{réflexivité}} : (x,x) \in R, \text{ quel que soit } x \in E ; \\ \text{b) } \underline{\text{symétrie}} : (x,y) \in R \text{ implique } (y,x) \in R ; \\ \text{c) } \underline{\text{transitivité}} : \text{si } (x,y) \in R \text{ et } (y,z) \in R, \text{ alors } (x,z) \in R. \end{cases}$

Au lieu d'écrire $(x,y) \in R$,

on écrira souvent $x \sim_R y$ ou $x \equiv y \pmod{R}$, que l'on énonce:

x est congru à y, modulo R ; ou plus simplement $x \sim y$ ou $x \equiv y$, si, après une définition de R, on n'éprouve pas le besoin d'indiquer qu'il s'agit toujours de la même relation.

<u>Exemples</u> : Les relations données au début ne sont pas des relations d'équivalence, sauf la première.
On démontrera aisément que toutes les relations binaires suivantes sont des relations d'équivalence :

1°/ $\mathbb{Z}$ étant l'ensemble des entiers de signe quelconque, appelons E le sous-ensemble de $\mathbb{Z} \times \mathbb{Z}$ formé des couples (p,q) pour lesquels $q \neq 0$, nous prendrons comme relation d'équivalence :

(I,3;2) $$(p,q) \equiv (p',q'), \quad \text{si } pq' - p'q = 0 .$$

2°/ Prenons pour E l'ensemble des fonctions réelles ≥ 0 d'une variable réelle, c'est-à-dire l'ensemble des applications de $\mathbb{R}$ dans $\mathbb{R}_+ = \{x ; x \in R, x \geq 0\}$. Nous prendrons comme relation d'équivalence celle dans laquelle $f \equiv g$ si et seulement si f et g sont "équivalentes" pour $x \to +\infty$, c'est-à-dire si, quel que soit $\varepsilon > 0$, il existe x_0 tel que $x \geq x_0$ entraîne $f(x)(1-\varepsilon) \leq g(x) \leq f(x)(1+\varepsilon)$.

3°/ Dans l'ensemble $\mathbb{Z}$ des entiers, prenons la relation d'équivalence dans laquelle $p \equiv q$ si, et seulement si $p - q$ est divisible par un entier donné une fois pour toutes : m.
On écrit cela couramment, en arithmétique, $p \equiv q \pmod{m}$.

4°/ soit E l'ensemble des droites d'un plan. Définissons une relation d'équivalence par $D \equiv D'$, si D et D' sont parallèles ou confondues.

5°/ soit E le produit du plan par lui-même. Un élément de E est aussi ce qu on appelle un <u>vecteur</u> (ou <u>vecteur fixe</u>) d'origine A et d'extrémité B. Définissons dans E la relation d'équivalence $(A,B) \sim (A',B')$ si les vecteurs (A,B) et (A',B') sont équipollents.

6°/ Dans le même ensemble E, établissons la relation d'équivalence dans laquelle $(A,B) \sim (A',B')$, si (A,B) et (A',B') sont portés par une même droite et équipollents.

Classes d'équivalence. Partitions

- SI E est un ensemble muni d'une relation d'équivalence R, on appelle <u>classe d'équivalence</u> une partie de E formée exactement de tous les éléments equivalents à l'un d'entre eux.

<u>Théorème 1 - 2 classes d'équivalence sont toujours ou confondues ou disjointes.</u>

<u>Démonstration</u>. Soient A et B deux classes d'équivalence . si elles ne sont pas disjointes, soit x un point d'intersection. A est exactement formé de tous les éléments iquivalents à l'un de ses éléments a, mais, comme x et a sont équivalents, la propriété de transitivité c) montre que A est aussi formé de tous les éléments équivalents à x : mals alors, pour la même raison, B a la même propriété, et A et B sont confondues. Cette démonstration vient en outre de nous montrer qu'une classe d'équivalence est l'ensemble de tous les éléments équivalents à l'<u>un quelconque</u> des éléments de cette classe. On voit ainsi que les classes d'équivalence définissent une <u>partition de E</u>, c'est-à-dire <u>un ensemble de partiesde E non vides, 2 à 2 disjointes, de réunion E</u> SI l'on connait ces parties, on connait la relation d'équivalence, car x et y sont équivalents si et seulement si ils appartiennent à la même classe d'équivalence. Inversement, d'ailleurs, toute partition de E : $E = \bigcup_{i \in I} A_i$, ou les A_i sont non vides et 2 à 2 disjointes, définit une relation d'équivalence, à savoir la relation $x \equiv y$ s'il existe un. $i \in I$ tel que $x \in A_i$, $y \in A_i$; et alors les A_i sont les classes d'équivalence de cette relation. En particulier, si f est une <u>surjection</u> de E sur un ensemble F, la relation $f(x) = f(y)$ est une relation d'équivalence sur E, et les $f^{-1}(\{z\})$, $z \in F$, sont les classes d'équivalence.

Ensemble quotient

On appelle ensemble quotient de E par la relation d'équivalence R , et on note E/R l'ensemble dont chaque élément est une des classes d'équivalence. Si l'on convient de noter $\dot{x}$ la classe d'équivalence de x , alors $x \in \dot{x}$, et $\dot{x}$ est un élément de l'ensemble quotient.

- L'application, qui, à chaque élément x de E fait correspondre l'élément $\dot{x}$ qui est sa classe d'équivalence, est appelée <u>surjection canonique</u> de E sur E/R - C'est bien une surjection. Inversement, nous venons de voir plus haut que toute surjection f de E sur un ensemble F définit une relation d'équivalence R. , l'application qui, à tout $z \in F$, associe la classe d'équivalence $f^{-1}(\{z\})$, est une bijection de F sur E/R , ce qui permet de considérer F comme un "modèle" du quotient E/R .

On peut donner une Interprétation simple des ensembles quotient pour les exemples de relations d'équivalence donnes antérieurement page 12 :

Dans le premier, l'ensemble quotient n'est autre que l'ensemble $\mathbb{Q}$ des nombres rationnels, puisque justement on définit habituellement un nombre rationnel par une famille de couples (p,q), $q \neq 0$, 2 couples (p,q), (p',q') définissant le même nombre rationnel si $pq' - p'q = 0$.

-Dans le troisième, l'ensemble quotient est appelé l'ensemble des entiers module m, et un élément de cet ensemble est appelé un entier modulo m .

- Dans le quatrième, l'ensemble quotient est l'ensemble des directions de droites du plan.

- Dans le cinquième, c'est l'ensemble des <u>vecteurs libres</u> du plan. On l'appelle espace vectoriel attaché au plan; la classe d'équivalence d'un vecteur (A,B) se note $\overline{AB}$ et s'appelle le vecteur libre A B .

- Dans le sixième, c'est l'ensemble des "vecteurs glissants" du plan, utilisé en mécanique.

Quotient d'un groupe par un sous-groupe invariant

Soit G un groupe, où la loi de composition est notée multiplicativement : $(x,y) \longrightarrow xy$, application de G x G dans G . Soit Γ un sous-groupe invariant * . On considérera

* On dit aussi distingué au lieu d'invariant. Rappelons qu'on entend par là un sous-groupe Γ tel que $a \in G$, $x \in \Gamma$ implique $axa^{-1} \in \Gamma$.

la relation binaire : " $x \equiv y$ s'il existe $a \in \Gamma$ tel que $y = ax$, ou encore si $yx^{-1} \in \Gamma$ " . On montre que c'est une relation d'équivalence; les classes d'équivalence sont les "classes à droite" de G par rapport à Γ . Comme Γ est invariant, la relation d'équivalence :" $x \equiv y$ s'il existe $a \in \Gamma$ tel que $y = xa$, ou encore si $x^{-1}y \in \Gamma$ ", est identique à la précédente, et les classes à droite sont aussi des classes à gauche. Soient α, β, deux de ces classes; quels que soient $x \in \alpha$, $y \in \beta$, le produit xy appartient toujours à la même classe; on note cette classe $\alpha\beta$ La loi de composition $(\alpha,\beta) \longrightarrow \alpha\beta$ fait de l'ensemble quotient un nouveau groupe, que l'on note G/Γ, et que l'on appelle <u>groupe quotient de G par le sous-groupe invariant Γ</u>

Quotient d'un espace vectoriel par un sous-espace vectoriel *

Soit E un espace vectoriel sur un corps K , et soit F un sous-espace vectoriel. Considérons la relation binaire sur **E** : "la différence $x - y$ est dans F ". C'est une relation d'équivalence. D'ailleurs, pour la loi d'addition, **E** est un groupe abélien, **F** est un sous-groupe; tout sous-groupe est invariant dans un groupe abélien. La relation ci-dessus est alors un cas particulier de la précédente. Soit E/F le groupe quotient. Il est encore abélien, et nous y noterons encore additivement la loi de composition. Soit $\lambda \in$ **K** un scalaire, et α une classe d'équivalence; si x est un élément quelconque de $\alpha (\dot{x} = \alpha)$, λx appartient toujours à une même classe d'équivalence; on la note $\lambda\alpha$. La loi d'addition sur E/F, et la loi de multiplication par les scalaires: $(\lambda,\alpha) \longrightarrow \lambda\alpha$, font de **I/F** un espace vectoriel, qu'on appelle espace vectoriel quotient de E par le sous-espace vectoriel F . D'après la définition même des opérations vectorielles de E/F, on a , pour $x \in E, y \in E, \lambda \in K : (x+y)^{\cdot} = \dot{x} + \dot{y}$, $(\lambda x)^{\cdot} = \lambda \dot{x}$. Autrement dit, la surjection canonique $x \longrightarrow \dot{x}$ de **E** sur E/F est une application linéaire.

bien entendu, la surjection canonique n'est pas bijective (sauf si **F** $= \{0\}$) . Mais soit G un sous- espace vectoriel de E supplémentaire ** de F . La restriction à G de l'application canonique est une bijection de G sur E/F . En effet : a/ elle est injective, car, si x et y sont deux éléments de G tels que $\dot{x} = \dot{y}$, on a $x - y \in F$; mais on a aussi $x - y \in G$ puisque G est un sous-espace vectoriel; comme F et G sont supplémentaires, leur intersection est réduite à l'élément 0 , donc $x = y$;

* Voir Cours d'Algèbre, Chapitre 1, n°8, page 31.

** Tout élément de E s'exprime, d'une manière unique, comme somme d'un élément de **F** et d'un élément de G

b/ elle est surjective, car si α est une classe d'équivalence, et si x est un élément de cette classe, x s'écrit sous la forme $x = x' + x''$, $x' \in F$, $x'' \in G$; alors $x - x'' \in F$ donc $\dot{x}'' = \dot{x} = \alpha$, avec $x'' \in G$, donc l'application est bien surjective.

Ainsi l'application $x \to \dot{x}$ de G sur E/F est une bijection linéaire, c'est-à-dire une bijection conservant la structure vectorielle; elle permet de donner G comme modèle de l'espace vectoriel quotient E/F . En particulier G a même dimension que E/F *.

§ 4 RELATIONS D'ORDRE

Une relation binaire **R** sur un ensemble **E** est appelée relation d'ordre si elle vérifie les propriétés suivantes :

(I,4;1)

a) Reflexivité : $(x,x) \in R$;

b) Transitivité : si $(x,y) \in R$ et si $(y,z) \in R$, alors $(x,z) \in R$;

c) Antisymétrie : si $(x,y) \in R$ et $(y,x) \in R$, alors $x = y$.

Au lieu d'écrire $(x,y) \in R$,

on écrit aussi $x \preccurlyeq_R y$, ou $x \preccurlyeq y$ si, la relation d'ordre ayant Bté indiquée au début, il n'apparait pas nécessaire de la répéter.

-Dans ce cas, $y \succcurlyeq x$ signifie $x \preccurlyeq y$.

Les inégalités (I,4;1) peuvent donc aussi s'écrire :

a) $x \preccurlyeq x$;

b) si $x \preccurlyeq y$ et $y \preccurlyeq z$, alors $x \preccurlyeq z$;

c) si $x \preccurlyeq y$ et $y \preccurlyeq x$, alors $x = y$.

* - L'espace vectoriel quotient E/F existe toujours, puisque nous l'avons défini explicitement; mals existe-t-il toujours des supplémentaires G de F dans E ? C'est évident si **E** est de dimension finie, mais c'est vrai dans tout les cas; nous l'admettrons.

Exemples de relation d'ordre:

1°/ Dans l'ensemble $\mathbb{N}$ des entiers $\geqslant 0$, dans l'ensemble $\mathbb{Z}$ de tous les entiers, dans l'ensemble $\mathbb{Q}$ des nombres rationnels, dans l'ensemble $\mathbb{R}$ des nombres réels, la relation : $(x, y) \in R$ si et seulement si $x \leqslant y$, est une relation d'ordre.

Z

- Il y a lieu de remarquer que la relation : $(x, y) \in R$ si et seulement si $x \geqslant y$, est aussi une relation d'ordre, dite opposée à la précédente.

Le symbole $\preccurlyeq$ a été utilisé par analogie avec le symbole $\leqslant$; en fonction de cette analogie, les relations $x \preccurlyeq y$, $x \prec y$, $y \succcurlyeq x$, $y \succ x$, s'énoncent respectivement; x Inférieur à y, x strictement inférieur à y, y supérieur à x, y strictement supérieur à x ; $x \prec y$ signifie "$x \preccurlyeq y$ et $x \neq y$".

Z

- <u>Notons que nous rompons ici avec l'usage antérieurement acquis</u> en appelant Inférieur ce qu'on appelait inférieur ou égal, et strictement inférieur ce qu on appelait inférieur.-La-raison d'être de ces changements, pleinement justifiés par la suite, est que la notion la plus généralement utilisée est $\preccurlyeq$ et qu'il est bon qu'elle ait l'appellation la plus courte. On devra toujours utiliser $\preccurlyeq$ plutôt que $\prec$ toutes les fois que cela sera possible; quand on écrira'une inégalité stricte avec $\prec$, ce sera pour avertir le lecteur qu'il y a un point délicat, et que l'inégalité large $\preccurlyeq$ ne conviendrait pas. Par exemple, la continuité d'une fonction réelle f d'une variable réelle en un point a s'écrira ainsi :

"quel que soit $\varepsilon > 0$, il existe $\eta > 0$ tel que

$|x - a| \leqslant \eta$ entraîne $|f(x) - f(a)| \leqslant \varepsilon$:

Nous avons mis le symbole d'inégalité large toutes les fols que c'était possible, et n'avons employé l'inégalité stricte > 0 que là où c'était absolument nécessaire à l'énoncé.

- Ceci noté, l'analogie que nous avons utilisée entre une relation d'ordre quelconque $\preccurlyeq$ et la

Z relation particulière $\leqslant$ dans l'ensemble des nombres réels, peut conduire à certaines difficultés, puisqu'on peut être amené à noter $x \preccurlyeq y$ et à dire " x inférieur à y ", même si l'on a choisi la relation d'ordre $x \geqslant y$.

2°/ Dans l'ensemble des mots de la langue française, il existe une relation d'ordre, dite ordre alphabétique (si l'on convient d'identifier des homographes).

3°/ Dans l'ensemble $E = \mathcal{P}(F)$ des parties d'un ensemble F, il existe une relation d'ordre naturelle: $X \preccurlyeq Y$ si $X \subset Y$.

4°/ Dans l'ensemble $E = \mathbb{R}^F$ des fonctions définies sur un ensemble quelconque F et à valeurs réelles, il existe aussi une relation d'ordre naturelle : $f \preccurlyeq g$, si, quelque soit x de F, $f(x) \leqslant g(x)$.

On remarquera que, dans cette relation,

Z $f \prec g$ signifie que, quelque soit x, $f(x) \leqslant g(x)$, et que, pour au moins un x, $f(x) < g(x)$; elle ne signifie nullement que l'on a, quel que soit x, $f(x) < g(x)$.

5°/ Dans l'ensemble $\mathbb{N}_1$ des entiers $\geqslant 1$, existe une relation d'ordre essentielle en arithmétique : $a \preccurlyeq b$ si a divise b.

6°/ Dans un ensemble E quelconque, la relation : $x \preccurlyeq y$ si $x = y$, est une relation d'ordre. On dit que c'est l'ordre chaotique sur E.

- On dit qu'une relation d'ordre est <u>totale</u> ou que l'ensemble E muni de cette relation est <u>totalement ordonné</u> si, étant donné deux éléments quelconques x, y, de E, on a nécessairement $x \prec y$, $x = y$, ou $x \succ y$.

- Il en est ainsi des relations d'ordre naturel dans

Z $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$, dans l'ordre alphabétique des mots de la langue française; mais il n'en est pas ainsi pour les relations d'ordre considérées dans les exemples 3°, 4°, 5°.

Dans chacun de ces cas, si deux éléments x et y ne vérifient aucune des 3 relations indiquées, on dit qu'ils sont <u>non comparables</u>. Dans l'exemple 3°, deux parties non vides disjointes de F sont non comparables; dans 4°/, les fonctions 0 et x sont non comparables; dans 5°/ les entiers 2 et 3 sont non comparables. Dans la relation d'ordre chaotique 6°, 2 éléments distincts quelconques sont non comparables.

Parties majorées, majorants, maximum, borne supérieure

- On dit qu'une partie d'un ensemble ordonné E est majorée, s'il existe au moins un élément de E supérieur à tous les éléments de cette partie; un tel élément s'appelle un majorant de cette partie.

- Définition analogue pour *partie* minorée et élément minorant.

- Une partie à la fois majorée et minorée s'appelle "bornée".

- On dit qu'une partie de E admet un maximum, s'il existe un majorant de cette partie appartenant à cette partie

Une partie n'a pas nécessairement un maximum, mais si elle en a un, il est unique: car, si a et b sont deux maxima d'une partie, on a à la fois $a \preccurlyeq b$ et $b \preccurlyeq a$, donc $a = b$. Le maximum d'une partie **A**, **s'il** existe, se note $\operatorname{Max}_{x \in A} x$ ou $\operatorname{Max}(A)$

- Définition analogue pour un minimum; notation Min.

- On dit qu'une partie A de E admet une borne supérieure si l'ensemble de ses majorants admet un minimum, et ce minimum est appelé borne supérieure de la partie considérée.

La borne supérieure est donc le plus petit majorant; tout élément qui majore **A** majore aussi sa borne supérieure.

- Une partie n'admet pas nécessairement une borne supérieure, mais si elle en a une, elle est unique. La borne supérieure d'une partie A, si elle existe, se note $\operatorname{Sup}_{x \in A} x$ ou $\operatorname{Sup}(A)$.

Si la borne supérieure appartient à **A**, alors c'est un maximum, et réciproquement. Définition analogue pour la borne inférieure; notation Inf.

Théorème 2 - Sur la droite réelle $\mathbb{R}$, munie de sa relation d'ordre naturelle, toute partie majorée non vide a une borne supérieure et toute partie minorée non vide * a une borne inferieure. En outre la borne supérieure d'une partie majorée A est caractérisée par les relations:

a/ Pour tout $x \in A$, $x \leqslant b$;

b/ Quel que soit $b_1 < b$ il existe au moins un $x \in A$ tel que $b_1 \leqslant x \leqslant b$.

* Le "non vide" est essentiel. La partie $\emptyset$ est majorée, tout point de $\mathbb{R}$ en est un majorant, donc l'ensemble des majorants est $\mathbb{R}$ elle-même, qui n'a pas de minimum ! $\emptyset$, quoique bornée, n'a ni borne supérieure, ni borne inférieure !

Ce théorème est supposé avoir été démontré antérieurement. Naturellement une partie non majorée n'a pas de borne supérieure. La droite $\mathbb{R}$ elle-méme n'est ni majorée ni minorée.

Dans l'ensemble E des parties d'un ensemble **F** , pour la relation d'ordre X c Y (exemple **3")** toute partie admet une borne supérieure et une borne inférieure. Une partie de E est en effet un ensemble de parties de F, la borne supérieure est leur réunion et la borne inférieure leur intersection.

- Dans l'ensemble E des fonctions définies sur un ensemble **F** à valeurs réelles (exemple 4°), toute partie majorée'à une borne supérieure et toute partie minorée a une borne inférieure. SI en effet A est cette partie, la borne supérieure de A est la fonction f_0 donnée par

(I,4;2) $f_0(x) = \operatorname{Sup}_{f\in A} f(x)$, pour tout x de F .

On notera qu'il y a deux notions entièrement différentes, toutes deux utiles, et qu'il faudra bien se garder de confondre : d'une part, la borne supérieure de A dans l'ensemble ordonné $\mathbb{R}^F$, c'est une fonction réelle qu'on appelle enveloppe supérieure des fonctions $f \in$ A ; d'autre part la borne supérieure dans $\mathbb{R}$ de l'ensemble des valeurs de ces fonctions, $\operatorname{Sup}_{\substack{x\in F\\ f\in A}} f(x)$, qui est, Si elle

Z

existe, un nombre réel, et qu'on appelle la borne supérieure des fonctions $f \in A$ Si **A** est réduit à un seul elément f , on parlera de la borne supérieure de f , $\operatorname{Sup}_{x\in F} f(x)$, qui est un nombre réel, la borne supérieure de l'ensemble de ses valeurs, et on dira que c'est un maximum si elle est atteinte pour une valeur de x .

Remarque analogue pour enveloppe inférieure et borne inférieure. Ces notations, couramment adoptées, sont un peu dangereuses, parce que pas très cohérentes.

- Dans l'ensemble $\mathbb{N}_1$ des entiers ≥ 1 , pour la relation de divisibilité 5°, toute partie finie a une borne supérieure, qui est le p.p.c.m, et une borne Inférieure, qui est le p.g.c.d .

Fonctions croissantes

Soient E, F deux ensembles ordonnés.

Une application f de **E** dans **F** est dite croissante, si

(I,4;3) $x \preccurlyeq y$ entraîne $f(x) \preccurlyeq f(y)$.

Si, en outre,

(I,4;4) $x \prec y$ entraîne $f(x) \prec f(y)$, elle est dite strictement croissante.

- Définition modifiée convenablement pour décroissante et strictement décroissante.

- Si E est totalement ordonné, une fonction à la fois croissante et décroissante est constante. En effet, soient x et y deux éléments quelconques de E. On a ou bien $x \preccurlyeq y$ ou bien $x \succcurlyeq y$. Dans chacun des deux cas, comme f est à la fois croissante et décroissante, on a à la fois $f(x) \preccurlyeq f(y)$ et $f(x) \succcurlyeq f(y)$, donc $f(x) = f(y)$, et f est bien constante* ,

- Si E est ordonné, on appelle intervalle fermé $[a, b]$ $a \preccurlyeq b$, l'ensemble des éléments x de E qui vérifient $a \preccurlyeq x \preccurlyeq b$. -On appelle intervalle ouvert $]a, b[$ l'ensemble de ceux qui vérifient $a \prec x \prec b$ - On note par $[a, b[$ l'ensemble de ceux qui vérifient $a \preccurlyeq x \prec b$, et par $]a, b]$ l'ensemble de ceux qui vérifient $a \prec x \preccurlyeq b$. - On les appelle intervalles semi-ouverts. Dans tous ces cas, a est l'origine et b l'extrémité de l'intervalle; on dit aussi que a et b sont les extrémités de l'intervalle. Parfois nous désirerons considérer un intervalle d'origine a et d'extrémité b, sans désirer spécifier s'il est ouvert, semi-ouvert ou fermé; nous l'appellerons $|a, b|$.

- On appelle section droite fermée $[a,\rightarrow)$ l'ensemble des x tels que $a \preccurlyeq x$, et section droite ouverte $]a,\rightarrow)$ l'ensemble des x tels que $a \prec x$.

- Définition analogue pour les sections gauches $(\leftarrow, a]$ et $(\leftarrow, a[$.

* - Sur un ensemble E , nous avons signalé que la relation $x = y$ est une relation d'ordre. deux éléments distincts quelconques sont non comparables: Si F est un ensemble ordonné quelconque, toute application de E dans F est à la fois croissante et décroissante; elle n'est pas ndcessalrement constante. Ainsi l'hypothèse que E est totalement ordonné n'est pas superflue.

- Dans le cas de la droite **réelle** R munie de sa structure d'ordre naturel, de telles sections sont appelées **demi-**droites.

On convient de considérer que les sections sont aussi des intervalles, ainsi que l'ensemble E entier.

Droite achevée $\overline{\mathbb{R}}$. On appelle <u>droite achevée $\overline{\mathbb{R}}$</u> l'ensemble formé, d'un@ part des **éléments** de la droite réelle $\mathbb{R}$, d'autre part de deux éléments, notés $-\infty$ et $+\infty$. On **met** sur $\overline{\mathbb{R}}$ la relation d'ordre, notée par $\leqslant$, définie comme suit : $a \leqslant b$ si a et b sont finis et s'ils vérifient $a \leqslant b$ sur R, ou si $a = -\infty$, ou si $b = +\infty$. $\overline{\mathbb{R}}$ est encore totalement ordonnée, mais possède en outre un minimum $-\infty$ et un **maximum** $+\infty$, toute partie de $\overline{\mathrm{R}}$ est bornée et le **théorème** 2 est **encore** valable sur $\overline{\mathbb{R}}$ sans avoir même besoin de **supposer** A majorée puisque c'est toujours **vrai.*** Un intervalle **de** $\overline{\mathbb{R}}$ tel que $[a, +\infty[$ est contenu dans $\mathbb{R}$, **si** $a \in \mathrm{R}$; c'est ce qui permet d'appeler souvent $[a, +\infty[$ la demi-droite $[a, \rightarrow)$ de $\mathbb{R}$. De même, si A est une partie non majorée **de** $\mathbb{R}$, on écrit souvent $\mathrm{Sup.}(A) = +\infty$, ce qui revient **à** considérer A comme une partie de $\overline{\mathbb{R}}$.

§ 5 PUISSANCES. ENSEMBLES **DÉNOMBRABLES**

Considérons l'application **de** $\mathbb{N}$ dans $\mathbb{N}$, **qui, à** chaque entier, fait correspondre son double. Cette application est bi-univoque, elle permet donc de dire, dans un certain sens, qu'il existe autant de nombres pairs que de nombres entiers. On volt ainsi que, dans le cas d'ensembles infinis, **il** peut exister une bijection d'un ensemble sur une partie de cet ensemble distincte de lui-m2me.

Il peut paraitre téméraire d'essayer malgré cela de comparer entre eux les ensembles Infinis, c est néanmoins ce qu'on peut faire **grâce** a la notion de bijection.

* - La borne supérieure de la partie vide $\emptyset$ est le minimum **de** $\overline{\mathbb{R}}$ (voir note (1) page 21) donc $-\infty$; la borne **in**férlcure est $+\infty$. Pour une partie A non vide, on a toujours $\mathrm{Sup.}(A) \geqslant \mathrm{Inf}\ (A)$; c'est faux pour la partie **vide** !

Théorème 3 - (BERNSTEIN)

Soient E et F , deux ensembles auelconaues :

1°/ ou bien il existe une injection de **E** dans **F** , **ou** bien il existe une injection de **F** dans E (Les deux **circons**-tances n'étant pas exclusives l'une de l'autre).

2°/ S'il existe à la fois une injection de **E dans** F, et une injection de **F** dans E , alors il existe aussi une **bijec**-tion de E **sur** F .

Nous admettrons ce théorème.

Corollaire - Etant donné deux ensembles E, F, il n'y a aue trois possibilités :

a) Il existe une injection de **E** dans F , et il n'existe pas d'injection **de** F dans E . On dit dans ce cas aue F est strictement plus puissant **que** E , ou E strictement moins puissant aue **F** .

b) Il existe une injection de F dans E , et il n'existe pas d'injection de E dans F . Alors E **est** strictement plus puissant **que** F , ou F strictement moins puissant aue E .

c) Il existe une bijection de E sur F , on dit **alors** que E **et** F sont également puissants ou équipotents *

Puissances. Cardinaux

La relation "E est **équipotent à** F " est une relation d'équivalence entre ensembles.** **Une classe d'équivalence,** c'est-à-dire la classe de tous les ensembles équipotents **à** un ensemble donné, est ce qu'on appelle une puissance ou un nombre cardinal. Les nombres cardinaux finis sont les classes d'équivalence des ensembles finis, ces nombres sont par définition les entiers naturels 0, 1, 2, etc . . . . . (Noter que nous avons admis comme une notion première les nombres entiers naturels, mais que leur définition mathématique en **toute**

* La **démonstration** du théorème de BERNSTEIN est délicate; nous l'admettrons. Le ?Corollaire se déduit **immédiatement** du théorème. Si **E** est **équipotent à** F ou strictement moins puissant que F , on dira **que** E est moins puissant que F .

** Volontairement nous n'insistons pas sur les difficultés logiques: nous n'avons parlé de relations d'équivalence que sur un ensemble, or il n'y a pas d'ensemble de tous les ensembles.

rigueur est pleine de complications; en particulier il n'est pas si facile de définir à priori les ensembles finis; on adopte fréquemment comme définition : un ensemble est fini s'il n'est équipotent à aucune de ses parties distincte de lui-même, et on démontre que les nombres cardinaux finls possèdent les propriétés que nous connaissons aux entiers naturels). Un nombre cardinal non fini, c'est-à-dire la puissance d'un ensemble Infini, s'appelle nombre cardinal transfini ou nombre transfinl.

Il existe une relation d'ordre dans la classe des cardinaux : $\alpha \leq \beta$, si α est le cardinal d'une partie d'un ensemble de puissance β . C'est le théorème 3 qui nous assure que cette relation est bien. antisymétrlque, c'est-à-dire est bien une relation d'ordre; il nous assure en outre que c'est une relation d'ordre totale : deux cardinaux sont nécessairement comparables.

- Notons que, si E et F sont deux ensembles, et s'il existe une surjection f de E sur F , F est moins puissant que E . En effet l'image réciproque de tout point de F est non vide, et si, dans chacune de ces images réciproques, nous choisissons exactement un élément, nous formons une partie de E qui est équlpotente à F * . Ainsi le quotient de E par une relation d'équivalence est toujours moins puissant que E .

On peut définir sur les nombres cardinaux des opérations d'addition de multiplication et d'exponentiation, comme sur les entiers naturels finis.

1°/ Soient α et β 2 cardinaux, et soient E et F des ensembles de puissances respectives α et β . On appelle $\alpha + \beta$ la puissance d'une "somme" de E et F , c'est-à-dire de tout ensemble, admettant une partition formée de deux ensembles respectivement équipotents à E et F .

2°/ - On appelle $\alpha\beta$ la puissance du produit E x F . C'est aussi le cardinal d'un ensemble réunion de a parties disjointes, toutes de cardinal β.

3°/ - On appelle α^{β} la puissance de l'ensemble E^F des applications de F dans E .

* Etant donné un nombre fini d'ensembles, il est aisé de choisir un élément dans chacun d'eux. Faire le même choix pour une infinité d'ensembles est plus embarrassant! Après de grandes discussions au début de ce siècle, il a étd reconnu que la possibilité de faire ce choix ne pourrait que reposer sur un axiome de la théorie des ensembles, l'axiome de choix ou axiome de Zermelo.

Théorème 4 - Les opérations précédentes définies sur les nombres cardinaux possèdent les propriétés suivantes :

- associativité et commutativité de l'addition;
- associativité et commutativité de la multiplication;
- distributivité de la multiplication par rapport à l'addition;

en outre :

(I,5;1) $(\alpha^{\beta})(\alpha^{\gamma}) = \alpha^{\beta+\gamma}$; $\alpha^{\gamma}\beta^{\gamma} = (\alpha\beta)^{\gamma}$; $(\alpha^{\beta})^{\gamma} = \alpha^{\beta\gamma}$

Démonstration. Seules les égalités (I,5;1) ne sont pas évidentes

Soient alors E, F, G, des ensembles, de puissances respectives α, β, γ . Alors E^F, E^G, ont pour puissances $\alpha^{\beta}, \alpha^{\gamma}$; pour se définir une application de l'ensemble somme F + G dans E , il suffit de se définir la restriction de cette application à F et à G , c'est-a-dire une application arbitraire de F dans E et une application arbitraire de G dans E; Autrement dit on obtient l'élément le plus général de E^{F+G} comme couple d'un élément de E^F et d'un élément de E^G : or un tel couple est un élément arbitraire de $E^F \times E^G$; donc E^{F+G} est équipotent à $E^F \times E^G$ et $\alpha^{\beta+\gamma} = \alpha^{\beta}\alpha^{\gamma}$

Nous avons vu page 7 qu'une application de G dans E x F, c'est-à-dire un élément de $(E \times F)^G$, est le couple d'une application de G dans E et d'une application de G dans F, c est-à-dire d'un élément de E^G et d un élément de F^G ; or un tel couple est un élément arbitraire de $E^G \times F^G$, donc $(E \times F)^G$ est équipotent à $E^G \times F^G$, et $(\alpha\beta)^{\gamma} = \alpha^{\gamma}\beta^{\gamma}$

Soit maintenant f une application de F x G dans E. Pour y fixé dans G , l'application $x \to f(x,y)$ de F dans E est ce qu'on appelle l'application partielle f_y ; elle est donc définie par $f_y(x) = f(x,y)$. Ainsi $f \in E^{F\times G}$ définit une application $y \to f_y$ de G dans l'ensemble E^F des applications de F dans E , c'est-a-dire un élément de $(E^F)^G$. Inversement si $y \to g_y$ est une application de G dans E^F , elle provient, par le procédé précédent de l'application g de F x G dans E définie par $g(x,y) = g_y(x)$. Ainsi nous avons établi une bijection de $E^{F\times G}$ sur $(E^F)^G$, donc $\alpha^{\beta\gamma} = (\alpha^{\beta})^{\gamma}$.

- Ces formules pourraient laisser croire que les nombres cardinaux même transfinis, possèdent toutes les propriétés simples des entiers naturels, il n'en est rien, car voici une propriété plus surprenante, que nous admettrons,

Théorème 5 - Si α et β sont deux nombres cardinaux $\neq 0$, et si l'un au moins d'entre eux est transfini, la somme $\alpha + \beta$ et le produit a β sont égaux au plus grand des deux *

Il en résulte en particulier qu'on ne peut pas définir de soustraction parmi les nombres cardinaux transfinis, car, parmi les nombres qui, ajoutés à α, font a figure 0, mais aussi n'importe quel nombre fini ou α lui-même.

Théorème 6 - Quel que soit l'ensemble E, l'ensemble des parties de E est strictement plus, puissant que E.

Ce théorème montre que la succession de; nombres cardinaux infinis est illimitée.

Supposons en effet qu'il existe une surjection f de E sur $\mathcal{P}(E)$. Alors, pour $x \in E$ $f(x)$ est un élément de $\mathcal{P}(E)$, c'est-à-dire une partie de E. Appelons A la partie de E formée des $x \in E$ tels que $x \notin f(x)$. Alors il n'existe pas d'élément y de E tel que $f(y)$ soit la partie A, puisque y appartient à l'une des deux parties $f(y)$, A et pas à l'autre (par définition, on a $y \in A$ si et seulement si $y \notin f(y)$). Nous aboutissons donc à une contradiction.

Comme par ailleurs il existe une injection de E dans $\mathcal{P}(E)$, à savoir $x \longrightarrow \{x\}$, E est moins puissant que $\mathcal{P}(E)$, donc strictement moins puissant.

Remarque 1 L'ensemble $\mathcal{P}_f(E)$ des parties finies de E est par contre équipotent à E si E est infini. On peut le démontrer facilement en utilisant le théorème 7 donné plus loin. L'application $(x_1, x_2, \dots x_n) \longrightarrow \{x_1, x_2, \dots x_n\}$, qui, à chaque élément $(x_1, x_2, \dots x_n)$ de E^n $(n \geq 1)$, fait correspondre la partie de E formée de ces éléments (non nécessairement tous distincts), est une surjection de E^n sur l'ensemble $\mathcal{P}_n(E)$ des parties non vides de E formées d'au plus n éléments. Donc $\text{card}.\mathcal{P}_n(E) \leq \text{card}.E^n = \text{card}.E$ (théorème 5), et comme. $\text{card}.\mathcal{P}_n(E) \geq \text{card}.E$, on a $\text{card}.\mathcal{P}_n(E) = \text{card}.E$. Soit $f_n : x \longrightarrow f_n(x)$, une bijection de E sur $\mathcal{P}_n(E)$. Posons $f_0(x) = \emptyset$ quel que soit $x \in E$. Alors $(n, x) \longrightarrow f_n(x)$ est une

* Démonstration délicate.

surjection de $\mathbb{N} \times E$ sur $\mathcal{P}_f(E)$, donc $card.\mathcal{P}_f(E) \leq card(\mathbb{N} \times E)$ $= \nu\ card.E = card.E$ (théorème 5, parce que $\nu \leq card.E$, voir plus loin théorème 7), donc $card.\mathcal{P}_f(E) = card.E$.

Remaraue 2 - On appelle fonction caractéristique d'une partie A de E la fonction φ_A définie sur E , à valeur dans l'ensemble à 2 éléments 0, 1, telle que

$$(I,5;2)\quad \varphi_A(x) = 1 \text{ si } x \in A , \qquad \varphi_A(x) = 0 \text{ si } x \notin A .$$

La connaissance de cette fonction détermine la partie A sans ambigufté; et d'ailleurs toute fonction sur E prenant les valeurs 0 et 1 détermine une partie unique. Il existe ainsi une bijection de l'ensemble $\mathcal{P}(E)$ des parties de E sur l'ensemble $\{0,1\}^E$ des applications de E dans l'ensemble $\{0,1\}$. Le cardinal de $\mathcal{P}(E)$ est donc $2^{card.E}$.

- Le théorème précédent peut donc s'énoncer :

Quel aue soit le nombre cardinal α , on a $2^\alpha > \alpha$. *

- Nous allons étudier maintenant les deux puissances transfinies les plus importantes : la puissance du dénombrable et la puissance du continu

Ensembles dénombrables -

- On appelle puissance du dénombrable et on notera Y la puissance de l'ensemble $\mathbb{N}$ des entiers naturels. Un ensemble équipotent à $\mathbb{N}$ est dit dénombrable **

Théorème 7 - 3 est le plus petit cardinal transflni.

Cela veut dire seulement que tout ensemble infini E contient au moins une partie dénombrable. Or d'après le théorème 3, si l'on n'a pas $card\ E > 3$, c'est qu il existe une bijection de E sur une partie (infinie) P de $\mathbb{N}$; mais il existe une bijection canonique de $\mathbb{N}$ sur P , a savoir $n \longrightarrow x_n$, où x_n est le $(n+1)$ième élément de P par ordre de grandeur; alors $card.E = \nu$.

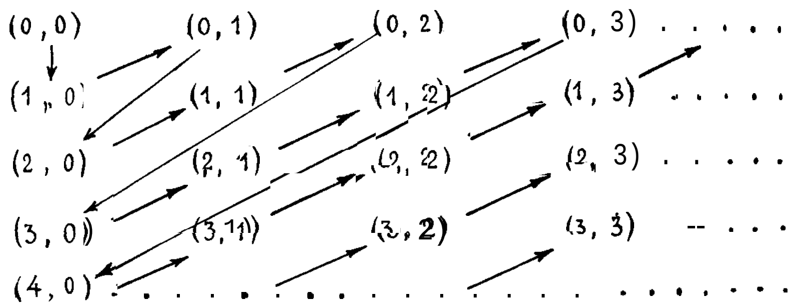
* Ceci montre, en particulier, que, pour tout entier $n \geq 0$, $2^n > n$.

** Dire d'un ensemble qu'il est dénombrable, c'est dire qu'il admet au moins une bijection sur $\mathbb{N}$, cela ne veut pas dire qu'une telle bijection soit donnée, et en général il n'en existe pas une qui s'impose plus qu'une autre. Cela veut dire encore qu'on peut ranger les éléments de cet ensemble en une suite $x_0, x_1, x_2, \ldots, x_n, \ldots$; mais cette suite n'est pas donnée. On appelle souvent $\aleph_0$ ($\aleph$ est la première lettre de l'alphabet hébreu) le cardinal de $\mathbb{N}$ Dans beaucoup d'ouvrages, on appelle dénombrable un ensembie de cardinal $\leq \nu$. E est fini ou dénombrable ($card.\ E \leq 3$), si et seulement s'il existe une injection de E dans $\mathbb{N}$, ou, en supposant $E \neq \emptyset$, si et seulement s'il existe une surjection de $\mathbb{N}$ sur E .

<u>Théorème 8</u> - <u>On a $m\,\nu = 3$, pour tout entier fini $m \geq 1$</u> .
<u>On a $\nu^m = \nu$, pour tout entier fini $m \geq 1$</u> .

Cela résulte immédiatement de théorème 5, qui montre que $m\nu = \nu$ puisque $m \leq 3$,et que $\nu^2 = \nu\nu = \nu$, donc que $\nu^m = \nu$ par récurrence sur m . Mais nous avons admis le théorème 5. Or on peut démontrer très élémentairement que $\nu\nu = Y$, d'où l'on déduit aussitôt tout le théorème 8 :

nous allons demontrer que <u>$\mathbb{N} \times \mathbb{N}$ est équipotent à $\mathbb{N}$</u>.
Si l'on énumère les éléments de $\mathbb{N} \times \mathbb{N}$ par parallèles successives à la bissectrice du tableau carré :



on obtient la suite $(0,0), (1,0), (0,1), (2,0), (1,1), (0,2), \ldots$. qui définit une bijection de $\mathbb{N}$ sur $\mathbb{N} \times \mathbb{N}$.

<u>Corollaires</u> 1°/ <u>La réunion d'un nombre fini ou d'une infinité dénombrable de parties finies ou dénombrables d'un ensemble E est finie ou dénombrable.</u>

soit en effet I une partie de $\mathbb{N}$, et A_i , $i \in I$,les parties de E considérées. On peut supposer qu'aucune n'est vide, car celles qui sont vides ne changent rien à la réunion. Soit f_i une surjection de $\mathbb{N}$ sur A_i . Alors $(i,n) \longrightarrow f_i(n)$ est une surjection de $I \times \mathbb{N}$ sur $\bigcup_{i \in I} A_i$; comme $I \times \mathbb{N}$ est dénombrable (ou vide si I est vide), $\bigcup_{i \in I} A_i$ est finie ou dénombrable.

2°/ <u>L'ensemble $\mathbb{Z}$ de tous les entiers est dénombrable</u>, comme réunion de 2 ensembles dénombrables. <u>L'ensemble $\mathbb{Q}$ des nombres rationnels est dénombrable</u>, car l'application qui, à chaque couple (p, q) , $q \neq 0$, de $\mathbb{Z} \times \mathbb{Z}$,fait correspondre le nombre rationnel

$\frac{r}{q}$, est une surjection sur $\mathbb{Q}$ d'un sous ensemble de $\mathbb{Z} \times \mathbb{Z}$.

Donc $\mathbb{Q}$ est au plus dénombrable, et, comme il contient $\mathbb{N}$, il est dénombrable.

3°/ <u>L'ensemble des nombres algébriques réels est dénombrable.</u>

On appelle nombre algébrique réel un nombre réel qui est racine d'un polynôme non identiquement nul à coefficients entiers.

Comme un polynôme de degré $\leq m$ non identiquement nul à coefficients entiers, a $m+1$ coefficients, qui sont des entiers arbitraires non tous nuls, la puissance de l'ensemble de ces polynômes est $\nu^{m+1} = \nu$. Or un tel polynôme a au plus m racines algébriques réelles. Donc l'ensemble des nombres algébriques de degré $\leq m$ apparait comme la réunion d'une infinité dénombrable d'ensemble finis. Il est donc au plus dénombrable, et comme il est infini, il est dénombrable. Lorsquen prend toutes les valeurs, on obtient ainsi l'ensemble de tous les nombres algébriques comme réunion d'une infinité dénombrable d'ensembles dénombrables, et cet ensemble est donc bien dénombrable.

4°/ L'ensemble de tous les points de $\mathbb{R}^n$ dont les coordorinées sont rationnelles ou algébriques est dénombrable; en effet son cardinal est $\nu^n = \nu$.

Puissance du continu

<u>Théorème 9</u> - L'<u>ensemble des nombres réels n'est pas dénombrable</u>

<u>Démonstration</u> - Nous allons même montrer que l'ensemble E des nombres réels vérifiant $0 \leq x < 1$ n'est pas dénombrable. Si en effet il l'était, il existerait une bijection de $\mathbb{N}$, sur E, c'est-à-dire une suite d'éléments deux à deux distincts de E : $x_1, x_2, \ldots x_n, \ldots$.

Or considérons le nombre réel ξ suivant : avant la virgule, nous mettrons 0 ; après, nous lui prendrons pour $j^{ième}$ décimale n'importe quel entier entre 1 et 8 distinct de la $j^{ième}$ décimale de x_j . Nous formons ainsi un développement décimal illimité qui représente bien un nombre ξ . On a nécessairement $\xi \neq x_n$, puisque la $n^{ième}$ décimale de ξ est distincte de la $n^{ième}$ décimale de x_n et que toutes les décimales de ξ sont distinctes de 0 et 9. (Il pourrait en effet y avoir des difficultés avec les décimales 0 et 9, puisqu'un nombre dont le développement décimal ne contient que des zéros, à partir d'un certain moment, admet une autre représentation à l'aide d'un développement décimal ne contenant que des 9 à partir d'un certain moment (Exemple : 0,10200000.... = 0,1019999....)

Ainsi nous avons abouti à une contradiction : l'ensemble $\{x_1, x_2, \ldots, x_n, \ldots\}$ ne contient pas ξ, alors qu'il devrait être $[0,1[$ tout entier.

Nombres transcendants. On appelle nombre **réel** transcendant un nombre réel qui n'est pas algébrique. Comme l'ensemble des nombres algébriques est dénombrable et que l'ensemble des nombres réels ne l'est pas, on voit qu'il existe des nombres transcendants, et même que "la plupart" des nombres réels sont transcendants. Il n'est pas pour cela plus facile de nommer explicitement un nombre transcendant. On peut montrer (mais ce n'est nullement évident !) que les nombres e et π sont transcendants.

Remarque La partie de $E = [0,1[$ formée des nombres dont le développement **décimal** ne contient que les chiffres 1, 2, 3, 4, 5, 6, 7, 8, est un ensemble équipotent à $\{1, 2, 3, 4, 5, 6, 7, 8\}^{\mathbb{N}}$, donc de puissances 8^{ν}.

Quant à l'ensemble E lui-même, sa puissance est $\leq 10^{\nu}$ (nous mettons $\leq$ et non $=$, à cause de la double représentation décimale signalée précédemment).

On a donc $8^{\nu} \leq \text{card}.E \leq 10^{\nu}$, et par suite à fortiori $2^{\nu} \leq \text{card}.E \leq 16^{\nu} = (2^4)^{\nu} = 2^{4\nu} = 2^{\nu}$. Donc $\text{card}.E = 2^{\nu}$.

Le théorème que nous venons de démontrer là n'est donc qu'un cas particulier du théorème 6. D'ailleurs la démonstration que nous venons d'en donner n'est qu'un cas particulier de celle du **théorème** 6 (avec une légère complication due aux doubles représentations décimales).

Nous désignerons par γ la puissance de l'ensemble E précédent, on l'appelle puissance du continu; c'est aussi la puissance de l'ensemble $\mathbb{R}$ des nombres reels (car $x \longrightarrow \log \frac{x}{1-x}$ est une bijection de $]0,1[$ sur $\mathbb{R}$).

Théorème 10 - On a les **égalités** : $m\gamma = \nu\gamma = \gamma\gamma = \gamma^m = \gamma^{\nu} = \gamma$, **où** m **est** un entier **fini** ≥ 1 arbitraire.

Démonstration - **Tous** ces **nombres cardinaux étant** $\leq \gamma^{\nu}$ et $\geq \gamma$, **il** suffit de montrer que $\gamma^{\nu} = \gamma$. Or $\gamma^{\nu} = (2^{\nu})^{\nu} = 2^{\nu^2} = 2^{\nu\cdot\nu} = \gamma$.

Corollaires * 1°/ L'ensemble des nombres complexes a la **puissance** du continu. Car il est **équipotent à** $\mathbb{R}^2$.

* Ces corollaires ainsi qu'un grand nombre d'autres propriétés **élémentaires** du dénombrable et du continu pourront être démontrées par les éléves'ou dans les petites classes, directement, **c'est-à-dire** sans utiliser les grands théorèmes (admis sans démonstration): 3,4,5.

2°/ <u>Tout espace vectoriel de dimension finie n sur le corps des réels ou des complexes a la puissance du continu</u>. Car, en choisissant une base on définit une bijection d'un tel espace sur $\mathbb{R}^n$, qui a la'puissance du continu d'après l'égalité $\gamma^n = \gamma$. Il en résulte en particulier cette conséquence assez paradoxale qu'il existe une bijection de $\mathbb{R}$ sur le plan $\mathbb{R}^2$, ces deux ensembles sont équipotents. On voit que lorsque, pour comparer des familles de points, de courbes, de surfaces, dépendants d'un certain nombre de paramétres, on dit, un peu rapidement, que l'une de ces familles est plus grande qu'une autre puisque ses éléments dépendent de 3 parcmètres réels alors que les éléments de l'autre dépendent de 2 paramètres réels, on raisonne un peu à la légère, puisque $\mathbb{R}^2$ et $\mathbb{R}^3$ sont équipotents. Nous verrons au chapitre II de la topologie, à quoi peuvent correspondre ces autres types de comparaisons, mais en aucun cas il ne peut s'agir de comparaison des infinis au sens actuel.

3°/ L'ensemble des suites de nombres réels ou des suites de nombres complexes a la puissance du continu, car son cardinal est $\gamma^{\nu} = \gamma$.

4°/ L'ensemble E des fonctions continues réelles d'une variable réelle a la puissance du continu. Car à une telle fonction on peut faire correspondre une suite de nombres réels, à savoir la suite de ses valeurs aux points d'abscisses rationnelles, et l'on peut supposer ces points en correspondance bi-univoque avec $\mathbb{N}$, puisque $\mathbb{Q}$ est dénombrable. Naturellement la suite de ces valeurs n'est pas arbitraire: si on choisit n'importe comment les valeurs d'une fonction aux points d'abcisses rationnelles, on ne peut pas la prolonger en une fonction continue sur la droite réelle. Mais cette suite de nombres réels, attribuée à une fonction continue, la détermine complètement. On peut donc trouver une bijection de l'ensemble des fonctions continues sur une partie de l'ensemble des suites de nombres réels. Donc cet ensemble E a au plus la puissance du continu. Et, comme l'application qui, à chaque fonction continue, fait correspondre sa valeur à l'origine, est une surjection de E sur $\mathbb{R}$, il a au moins la puissance du continu, donc il a exactement la puissance du continu.

5°/ L'ensemble de toutes les fonctions réelles d'une variable réelle, ou même l'ensemble de toutes celles qui ne prennent que les valeurs 0 et 1, a une puissance strictement supérieure à celle du continu, car ces puissances valent respectivement γ^{γ} et 2^{γ} ; or $\gamma^{\gamma} = (2^{\nu})^{\gamma} = 2^{\nu\gamma} = 2^{\gamma} > \gamma$. On déduit de 4° et 5° que "la majorité" des fonctions ont au moins un point de discontinuité.

Hypothèse du continu - Cette hypothèse consiste à supposer que 2^{ν} est le cardinal immédiatement au-dessus de 3. Plus généralement l'hypothèse du continu généralisée consiste à supposer que pour tout cardinal transfini α, le cardinal 2^{α} est immédiatement au-dessus de lui.

On a démontré récemment (Paul Cohen, 1963) que l'hypothèse du continu était indécldable : on ne peut pas la démontrer, ni démontrer le contraire; on peut la rajouter comme axiome, ou rajouter comme axiome le contraire.

§6 QUELQUES PRINCIPES DE LOGIQUE

Un théorème consiste généralement à énoncer qu'une certaine propriété A , appelée Hypothèse, entraîne une certaine propriété B, appelée Conclusion.

On abrège généralement l'expression "A implique B " ou " A entraîne B " par la formule $A \Longrightarrow B$.

Le théorème réciproque, qui n'est pas toujours vrai, est alors : $B \Longrightarrow A$.

Si le théorème et sa réciproque sont tous les deux vrais, les propriétés A et B sont équivalentes, et l'on peut énoncer le théorème sous la forme $A \Longleftrightarrow B$, qui veut dire aussi :

"Pour que A . . . . il faut et il suffit que B.... "

Le contraire ou la négation d'une propriété A s'écrit : "Non A " ou $\not{A}$. Par exemple, si x n'appartient pas à l'ensemble E on écrira : $x \notin E$ Si X et Y sont des parties de E , et si X n'est pas contenue dan8 Y , on écrira $X \not\subset Y$,

(ce qui ne signifie nullement que X contient Y).

On a non non A = A , et on a toujours " A ou non A "(principe du tiers exclu).

Théorème 11 - La proposition $A \Longrightarrow B$ est vraie si et seulement si la proposition $\not{B} \Longrightarrow \not{A}$ est vraie.

Démonstration :

1°/ Supposons que $A \Longrightarrow B$ soit vraie et supposons que l'on ait $\not{B}$; alors on ne peut pas avoir A , sans quoi l'on aurait B , ce qui est contraire à l'hypothèse $\not{B}$; donc on a $\not{A}$, et il est bien montré que $\not{B} \Longrightarrow \not{A}$ est vraie.

2°/ - Si $\not B \Longrightarrow \not A$ est vraie, alors, d'après 1°/, $\not\not A \Longrightarrow \not\not B$ est vraie, mais $\not\not A = A$, $\not\not B = B$, donc $A \Longrightarrow B$ est vraie.

Exemple - Il est équivalent de dire que toute fonction réelle, partout continue, sur un intervalle réel $[a, b]$ borné, est bornée, ou de dire que toute fonction réelle définie sur $[a, b]$ et non bornée est discontinue en au moins un point.

Si un objet vérifie la propriété A ou la propriété B, on écrit qu'il vérifie $A \vee B$, ou encore " A ou B ". En mathématique, la conjonction ou ne marque jamais des événements exclusifs les uns des autres. A ou B ne signifie nullement que l'on n'ait pas à la fois A et B. Par exemple : si A est la propriété $x \leq 0$, et si B est la propriété $x \geq 0$, on a toujours A ou B, ces deux propriétés n'étant pas exclusives puisque l'on peut avoir $x = 0$.

Si l'on a à la fois les propriétés A et B, on écrit que l'on a $A \wedge B$ ou encore " A et B ".

Dans un théorème de mathématique s'introduisent généralement plusieurs fois les expressions : "Quelque soit ......" et "Il existe . . ..tel que ......". On les note respectivement $\forall$ et $\exists$. On les appelle les quantificateurs.

Ces : "Quelque soit ......" et ces : "Il existe . . . ...' sont souvent accompagnés de certaines restrictions. On note entre parenthèses ( ) ces restrictions.

Supposons que nous voulions exprimer, pour une fonction réelle f d'une variable réelle, la propriété d'être continue en tout point. Dire qu'elle est continue en un point a, c'est dire que, quel que soit $\varepsilon > 0$, il existe $\eta > 0$ tel que $|x - a| \leq \eta$ entraîne $|f(x) - f(a)| \leq \varepsilon$; autrement dit : Quelque soit $\varepsilon > 0$ il existe. $\eta > 0$ tel que, quelque soit $x \in \mathbb{R}$ tel que $|x - a| \leq \eta$, on ait la propriété P : $|f(x) - f(a)| \leq \varepsilon$.

Pour exprimer maintenant que la fonction est continue en tout point, nous devrons écrire :

(Quelque soit $a \in \mathbb{R}$) (quel que soit $\varepsilon > 0$) (Il existe $\eta > 0$ tel que)

(quelque soit $x \in \mathbb{R}$ tel que $|x - a| \leq \eta$), on ait P : $|f(x) - f(a)| \leq \varepsilon$.

Cette phrase s'écrira en abrégé :

$$(I,6;1)\quad (\forall a \in \mathbb{R})(\forall \varepsilon > 0)(\exists \eta > 0)(\forall x \in \mathbb{R},\ |x-a| \leq \eta) : |f(x) - f(a)| \leq \varepsilon .$$

Théorème 12 - La négation d'une propriété contenant un certain nombre de fois les quantificateurs $\forall$, $\exists$, et ensuite l'énoncé d'une propriété P , s'obtient en remplaçant chaque quantificateur $\forall$ par le quantificateur $\exists$ et vice versa, et la propriété P par sa négation $\not{P}$. Ainsi la propriété pour la fonction f réelle d'une variable réelle de ne pas être partout continue, c'est-à-dire d'être discontinue en au moins un point, s'exprime par la ligne unique :

$$(I,6;2)\quad (\exists a \in \mathbb{R})(\exists \varepsilon > 0)(\forall \eta > 0)(\exists x \in \mathbb{R},\ |x-a| \leq \eta) : |f(x) - f(a)| > \varepsilon$$

Ce procédé pour nier une propriété doit devenir aussi automatique que la règle des signes dans la multiplication ou dans la suppression des parenthèses,

Démonstration :

Supposons qu'il n'y ait qu'un seul quantificateur, par exemple $\forall$. Notre propriété a donc la forme :

($\forall x$ vérifiant S) : P.

Sa négation est évidemment : il existe un x vérifiant S , qui cependant ne vérifie pas P , ou

($\exists x$ vérifiant S) : $\not{P}$.

Le théorème est ainsi démontré dans ce cas, et aussi, de manière analogue, s'il n'y a qu'un quantificateur $\exists$.

Il suffit alors de faire une récurrence sur le nombre de quantificateurs. Supposons le théorème démontré lorsqu'il y a $n-1$ quantificateurs, montrons-le lorsqu'il y en a n. Alors la propriété s'écrit, par exemple,

($\forall x$ vérifiant s) : Q ,

où Q est une propriété contenant $n-1$ quantificateurs.

Sa négation est donc

($\exists x$ vérifiant S) : $\not{Q}$,

en vertu de ce qui a été vu pour un seul quantificateur; mais $\not{Q}$ s'obtient en appliquant le théorème, puisque Q ne contient que n-1 quantificateurs, et alors le théorème est encore vrai dans ce cas; il en est de même si le premier quantificateur est $\exists$, et le théorème est vrai dans le cas général.

Remarque - Toutes les fois qu'un quantificateur $\exists$ est précédé d'un certain nombre d'autres quantificateurs, la lettre qui le suit est éventuellement fonction de toutes les lettres figurant dans les quantifications antérieures. Par exemple, dans la propriété pour une fonction d'être partout continue, η dépend de a et de ε. Il peut arriver en fait qu'on puisse choisir η dépendant de ε mais non de a ; on dit alors que la fonction est uniformément continue. La propriété pour une fonction d'être uniformément continue est plus forte que la propriété d'être continue, elle peut s'écrire de la manière suivante :

$$(\mathrm{I},6;3)\quad (\forall \varepsilon > 0)(\exists \eta > 0)(\forall a \in \mathbb{R})(\forall x \in \mathbb{R}, |x-a| \leqslant \eta) : |f(x) - f(a)| \leqslant \varepsilon .$$

Comme on le voit, une interversion de quantificateurs modifie considérablement la propriété énoncée; on voit immédiatement que la propriété la plus forte est celle où le symbole $\exists$ est placé le plus tôt. Avec une autre interversion, on aura :

$$(\mathrm{I},6;4)\quad (\exists \eta > 0)(\forall a \in \mathbb{R})(\forall \varepsilon > 0)(\forall x \in \mathbb{R}, |x-a| \leqslant \eta) : |f(x) - f(a)| \leqslant \varepsilon .$$

Choisissons ce η dont il est dit au début qu'il existe. Pour tout x de l'intervalle $[a-\eta, a+\eta]$, on aura $|f(x)-f(a)| \leqslant \varepsilon, \forall \varepsilon > 0$, donc $|f(x)-f(a)| = 0$. Donc on aura $f(x) = f(a)$, pour tout a et tout x de l'intervalle $[a-\eta, a+\eta]$. En prenant pour a tous les nombres $\frac{p}{10^{b}}$, p entier allant de $-\infty$ à $+\infty$, b fixe tel que $\frac{p}{10^{b}} \leqslant \eta, \ldots$, on voit que (I, 6 ; 4) signifie que f est constante, ce qui est une propriété de "continuité" encore plus forte !

N.B. Nous avons considéré toute la théorie des ensembles et la logique du point de vue "naif": les mots ensemble, égal, quelque soit, il existe, implique, etc .... sont pris avec le sens qu'ils ont dans la langue française et l'intuition courante. Il est bien évident qu'il n'y a pas là un fondement mathématique sérieux. Les. symboles $\forall$, $\exists$, $\in$, $=$, $\Longrightarrow$, doivent être seulement des signes, soumis à certaines "règles du jeu", comme le cheval et la tour aux échecs, sont des pièces aux mouvements réglementés, et non un vrai cheval et une vraie tour; comme aussi le Plan d'Euclide n'est pas une Surface d'eau et la boule n'est Pas une orange. La logique mathématique (comprenant la théorie des ensembles) est à elle Seule une branche des mathématiques modernes. Signalons qu'il n'est pas prouvé que les logiques actuelles ne soient pas "contradictoires". Une théorie logique est'dite contradictoire s'il existe une proposition P telle qu'on puisse à la fois démontrer P et $\not P$. Alors toute proposition Q (et aussi $\not Q$) est vraie. En effet $P \Longrightarrow (P$ ou $Q)$, et $((P$ ou $Q)$ et $\not P) \Longrightarrow Q$; donc $(P$ et $\not P) \Longrightarrow Q$. Si l'actuelle théorie des ensembles est contradictoire, cela ne signifie pas que nous soyons perdus sans remède, mais qu'il faut diminuer le nombre des axiomes de la théorie; ce qui, sans doute, ne changerait pas sensiblement l'ensemble des mathématiques!

II
TOPOLOGIE

§ 1 ESPACES MÉTRIQUES. EXEMPLES ÉLÉMENTAIRES

On appelle espace métrique un ensemble E muni d'une fonction distance, c'est-à-dire d'une application d de $E \times E$ dans la demi-droite $\mathbb{R}_+ = \{x ; x \in \mathbb{R}, x \geq 0\}$ de $\mathbb{R}$, qui, au couple (x, y) de $E \times E$ fait correspondre un nombre $d(x,y) \geqslant 0$, appelé distance de x et de y.

Cette distance doit posséder les 3 propriétés suivantes :

(II,1;1)
1°) symétrie : $d(x,y) = d(y,x)$;

2°) positivité : $d(x,y) > 0$ si $x \neq y$; et $d(x,x) = 0$.

3°) inégalité triangulaire : $d(x,z) \leq d(x,y) + d(y,z)$

(tout côté d'un triangle est au plus égal à la somme des 2 autres).

On sait que ceci entraîne, comme conséquence, que tout côté d'un triangle est au moins égal à la différence des deux autres :

$$d(x,z) \geqslant |d(x,y) - d(y,z)| . \tag{II,1;2}$$

Cela entraîne aussi que si $x_1, x_2, \ldots x_n$, sontn points arbitraires de E, on ait :

$$d(x_1, x_n) \leqslant d(x_1,x_2) + d(x_2,x_3) + \cdots + d(x_{n-1}, x_n) . \tag{II,1;3}$$

Donnons tout de suite des exemples importants d'espaces métriques :

1°/ la droite réelle $\mathbb{R}$, le plan complexe $\mathbb{C}$, munis de la distance $d(x,y) = |x-y|$; on appelle cette métrique la métrique naturelle de $\mathbb{R}, \mathbb{C}$; sauf mention expresse du contraire $\mathbb{R}$ ou $\mathbb{C}$ est toujours muni de sa métrique naturelle.

2°/ la droite réelle $\mathbb{R}$, munie de la distance $d(x,y) = |F(x) - F(y)|$, où F est n'importe quelle fonction réelle strictement monotone d'une variable réelle.

3°/ L'espace euclidien réel $\mathbb{R}^n$ ou hermitien complexe $\mathbb{C}^n$ à n dimensions, dans lesquels la distance du point $(x_1, x_2, \ldots x_n)$ et du point $(y_1, y_2, \ldots y_n)$ est $\left(\sum_{i=1}^{n} |x_i - y_i|^2\right)^{\frac{1}{2}}$.

Nous démontrerons plus loin rigoureusement qu'il s'agit bien d'un espace métrique (page 40).On appelle cette métrique la métrique naturelle de $\mathbb{R}^n$ ou $\mathbb{C}^n$; sauf mention expresse du contraire, $\mathbb{R}^n$ ou $\mathbb{C}^n$ sera toujours muni de sa métrique naturelle.

4°/ E étant un ensemble quelconque, on peut le munir de la métrique discrète, dans laquelle $d(x,y) = 1$ si $x \neq y$, et $d(x,x) = 0$.

Sphères, boules

On appelle sphère de centre a et de rayon R fini > 0 l'ensemble des x de E tels que $d(a,x) = R$. Rien ne dit qu'une telle sphère ne soit pas vide, ni que deux sphères de centres distincts ne puissent pas coïncider (exemple : dans la métrique discrète, toutes les sphères de rayon 2 coïncident et sont vides).

On appelle boule ouverte (resp. fermée) de centre a et de rayon R fini > 0, et on note $B_0(a,R)$ (resp. $B(a,R)$) l'ensemble des x de E tels que $d(a,x) < R$ (resp. $\leq R$) *
Quand on dira boule, sans autre spécification, il s'agira de boule fermée. Dans la métrique discrète, une boule fermée de rayon < 1 se réduit à son centre, une boule (fermée) de rayon ≥ 1 est l'espace entier. (Pour R = 0 , la boule ouverte serait vide, la sphère et la boule fermée se réduiraient à
Z leur centre; on supposera toujours, même si ce n'est pas dit explicitement, que le rayon des sphères ou des boules est fini et > 0).

* Contrairement à l'usage courant dans les lycées, on distinguera toujours boule et sphère. Dans le cas du plan $\mathbb{R}^2$, on dit aussi circonférence ou cercle au lieu de sphère, et disque au lieu de boule.

Une partie d'un espace métrique est dite <u>bornée</u> si elle est contenue dans au moins une boule (de rayon fini, comme toujours). Ainsi $\mathbb{R}$ n'est pas bornée, ni la partie $\mathbb{N}$ de $\mathbb{R}$; si E a la métrique discrète, il est borné.

Espaces vectoriels normés

Soit E un espace vectoriel sur le corps K des nombres réels ou des nombres complexes. On appelle alors <u>norme</u> sur l'espace vectoriel E toute fonction, notée $\vec{x} \mapsto \|\vec{x}\|$, * possédant les propriétés suivantes;

$$(\text{II},1;4)\quad \begin{cases} 1°)\ \text{positivité} : \|\vec{x}\| > 0 \quad \text{pour} \quad \vec{x} \neq \vec{0},\ \|\vec{0}\| = 0 ; \\ 2°)\ \text{transformation par les homothéties} : \|\lambda\vec{x}\| = |\lambda|\,\|\vec{x}\|,\ \lambda \in K ; \\ 3°)\ \text{inégalité de convexité} : \|\vec{x}+\vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| . \end{cases}$$

Naturellement, de 2°) et de 3°), on déduit aisément l'inégalité générale de convexité

$$(\text{II},1;5)\qquad \|\vec{x}_1 + \vec{x}_2 + \ldots + \vec{x}_n\| \leq \|\vec{x}_1\| + \|\vec{x}_2\| + \cdots + \|\vec{x}_n\| ,$$

et aussi

$$(\text{II},1;6)\qquad \|\vec{x} - \vec{y}\| \geq \big|\, \|\vec{x}\| - \|\vec{y}\| \,\big| .$$

Si E est muni d'une telle norme, on l'appelle <u>espace vectoriel normé</u>.

Naturellement on ne peut employer le symbole $\|\ \|$ que s'il s'agit d'une norme bien précisée une fois pour toutes. SI, dans un même problème, Interviennent plusieurs normes différentes sur un espace vectoriel, on devra bien les représenter par des symboles différents.

soit E un espace vectoriel normé; on peut le munir de la fonction distance définie par $d(\vec{x},\vec{y}) = \|\vec{x}-\vec{y}\|$, qui vérifie bien les axiomes requis. <u>Tout espace vectoriel normé est donc automatiquement un espace métrique</u>, sa distance posséde en outre des propriétés particulières

* Nous conviendrons de toujours surmonter d'une flèche les éléments d'un espace vectoriel. Bien distinguer le nombre 0 et l'origine $\vec{0}$ de l'espace vectoriel.

compatibles avec sa structure vectorielle :

d'une part la distance est invariante par translation, autrement dit $d(\vec{x}-\vec{a}, \vec{y}-\vec{a}) = d(\vec{x}, \vec{y})$, d'autre part une homothétie de rapport? multiplie la distance par $|\lambda|$: $d(\lambda\vec{x}, \lambda\vec{y}) = |\lambda| d(\vec{x}, \vec{y})$.

Réciproquement, on voit sans peine que toute distance sur un espace vectoriel, ayant les deux propriétés précédentes, est nécessairement définie à partir d'une norme; celle-ci n'est autre que $\|\vec{x}\| = d(\vec{0}, \vec{x})$.

Dans un espace vectoriel, on appelle boule ouverte de rayon R, sans préciser le centre, la boule ayant pour centre l'origine de l'espace vectoriel , et pour rayon R . De même pour la boule fermée . On les notera $B_0(R)$ et $B(R)$.

En particulier la boule unité ouverte (resp. fermée) est la boule ouverte (resp. fermée) de centre 0 et de rayon 1, sans autre spécification. Boule unité, sans spéclfication, veut dire : boule unité fermée.

Dans un espace vectoriel sur les corps des réels ou des complexes, on appelle segment d'extrémités a et b l'ensemble des points $t\vec{a} + (1-t)\vec{b}$, $0 \leq t \leq 1$. On le notera $[a, b]$, de la même manière qu'un intervalle fermé dans un ensemble ordonné. On dit qu'une partie de E est convexe si, toutes les fois qu elle contient 2 points distincts, elle contient tout le segment qui les a pour extrémités. L'inégalité de convexité nous' montre que dans un espace vectoriel normé, toute boule est un ensemble convexe. En effet, si $\|\vec{x}\| \leq R$ et $\|\vec{y}\| \leq R$, on a

$$\|t\vec{x} + (1-t)\vec{y}\| \leq \|t\vec{x}\| + \|(1-t)\vec{y}\| = |t|\,\|\vec{x}\| + |1-t|\,\|\vec{y}\| \leq [t+(1-t)]R = R$$

Donnons quelques exemples d'espaces vectoriels normés :

1°/ Sur le corps des scalaires $\mathbb{R}$ ou $\mathbb{C}$, la fonction $x \rightarrow |x|$ est une norme, elle définit la métrique vue plus haut (I°, page 37). On l'appellera la norme naturelle.

2°/ Sur l'espace vectoriel $\mathbb{R}^n$ ou $\mathbb{C}^n$, les 3 fonctions suivantes sont des normes :

$$(x_1, x_2, \dots x_n) \rightarrow \operatorname{Max}_{i=1}^{i=n} |x_i| \,, \quad \text{ou} \quad \sum_{i=1}^{i=n} |x_i| \,, \quad \text{ou} \quad \left(\sum_{i=1}^{i=n} |x_i^2|\right)^{\frac{1}{2}} .$$

C'est évident pour les deux premières, il suffit de le voir pour la 3ème. Sbient alors $(x_1, x_2, \dots x_n)$, $(y_1, y_2, \dots y_n)$, 2 systèmes de n nombres complexes.

Nous voulons démontrer l'inégalité de convexité

(II,1;8) $$\left(\sum |x_i+y_i|^2\right)^{\frac{1}{2}} \leqslant \left(\sum |x_i|^2\right)^{\frac{1}{2}} + \left(\sum |y_i|^2\right)^{\frac{1}{2}}.$$

Comme le premier membre est majore par $\left(\sum(|x_i|+|y_i|)^2\right)^{\frac{1}{2}}$, il suffit, en élevant au carré, de montrer

(II,1;9) $$\sum(|x_i|+|y_i|)^2 \leqslant \sum |x_i|^2 + \sum |y_i|^2 + 2\left(\sum |x_i|^2\right)^{\frac{1}{2}} \left(\sum |y_i|^2\right)^{\frac{1}{2}},$$

ou

(II,1;10) $$\sum |x_i||y_i| \; 6 \; \left(\sum |x_i|^2\right)^{\frac{1}{2}} \left(\sum |y_i|^2\right)^{\frac{1}{2}},$$

ce qui est l'inégalité de CAUCHY - SCHWARZ (Cours d'Algèbre, Chapitre III, n°7, théorème 7.1).

Cette norme $x = (x_1, x_2, \dots x_n) \longrightarrow \sqrt{\sum_{i=1}^{i=n} |x_i|^2}$ s'appellera la norme naturelle de $\mathbb{R}^n$ ou $\mathbb{C}^n$.

Nous verrons plus tard d'autres exemples essentiels d'espaces vectoriels normés, de dimension Infinie.

§ 2 OUVERTS. FERMES. VOISINAGES. INTERIEUR. FRONTIERE. ADHERENCE, SOUS ENSEMBLES DENSES

Parties ouvertes

Soit E un espace métrique. Une partie **A** de E est appelée ouverte si, toutes les fois qu'elle contient un point de E, elle contient au moins une boule ouverte (de rayon > 0) ayant pour centre ce point * . Les parties ouvertes de E possédent évidemment les propriétés suivantes :

* En abrégé : **A** est ouverte si : $(\forall x \in A)(\exists \rho > 0)(\forall y \in E, d(x,y) < \rho)$: $y \in A$. Ou plus rapidement : $(\forall x \in A)(\exists \rho > 0)$: $B_o(x,\rho) \subset A$.

Naturellement ce ρ dépend de x . On peut, dans la définition remplacer "boule ouverte" par "boule fermée" car la boule fermée de rayon ρ contient la boule ouverte de rayon ρ, et la boule ouverte de rayon ρ contient la boule fermée de rayon $\frac{\rho}{2}$.

(II,2;1)

a) **E** lui même , et la partie vide $\emptyset$ sont ouvertes *

b) Toute intersection d'un nombre fini d'ensembles ouverts est ouverte.

c) Toute réunion d'une famille finie ou infinie d'ensembles ouverts est ouverte.

Démontrons par exemple b).

Soient $\mathcal{O}_1, \mathcal{O}_2, \dots \mathcal{O}_n$,des ouverts de **E** . Si x appartient à leur intersection, il existe, pour chaque i , un nombre $R_i > 0$ tel que la boule $B(x, R_i)$ soit tout entière contenue dans $\mathcal{O}_i$. Alors, si on pose $R = \underset{i=1,2,\dots n}{\mathrm{Min}} (R_i)$, la boule $B(x, R)$ est contenue dans l'intersection; donc celle-ci est bien ouverte.

Il existe en outre une **4ème** propriété interessante appelée "Axiome de séparation de HAUSDORFF".

d) Quels que soient les points a et b distincts de **E**, il existe deux ouverts, contenant respectivement a et b, d'intersection vide.

Si en effet d est la distance $d(a, b)$, il suffit de prendre les boules ouvertes, de centres a et b , de rayon $\frac{d}{3}$. Elles ne peuvent avoir de point commun, car, s'il existait un tel point commun c , l'inégalité triangulaire $d(a, b) \leq d(a, c) + d(c, b)$ donnerait $d(a, b) \leq \frac{2d}{3}$, ce qui serait absurde.

* **A** est une partie ouverte si, pour tout $x \in$ **A** , il existe une boule de centre x contenue dans A . Si **A** $= \emptyset$, on ne peut pas trouver de point $x \in A$, alors la propriété est bien vérifiée, donc $\emptyset$ est ouverte . Voici un autre exemple du même raisonnement logique. Nous dirons qu'un homme a la propriété (P) s'il est plus grand que tous ses enfants: alors un homme qui n'a pas d'enfants a la propriété (P). Passons d'ailleurs à la propriété négative. Une partie A est non-ouverte s'il existe un $x \in$ **A** qui ne soit pas centre d'une boule contenue dans **A** ; alors $\emptyset$ n'est pas non-ouverte car il n'existe pas de $x \in \emptyset$; donc $\emptyset$ est ouverte.

Exemples - Les boules ouvertes sont bien des ensembles ouverts, ce qui justifie leur nom. En particulier, sur la droite $\mathbb{R}$, munie de sa métrique naturelle, les intervalles ouverts sont des ensembles ouverts. On vérifie aisément au contraire qu'un intervalle fermé ou semi-ouvert n'est pas un ensemble ouvert.

Pour qu'une partie de E soit ouverte, il est donc nécessaire, d'après la définition, qu'elle soit une réunion de boules ouvertes; c'est suffisant, puisqu'une boule ouverte est ouverte et qu'une réunion d'ouverts est ouverte (propriété c). C'est donc une condition nécessaire et suffisante pour qu'une partie soit ouverte.

Dans $\mathbb{R}^n$, munie de l'une quelconque des métriques définies par les normes données page 40 , un pavé ouvert, c'est-à-dire un ensemble défini par des inégalités strictes

$$\{x = (x_1, x_2, \dots x_n) \,;\, a_i < x_i < b_i \,,\, i = 1, 2, \dots n\} \text{ est ouvert.}$$

Dans tout espace métrique, l'ensemble des points x vérifiant

$$d(a,x) > R \text{ est ouvert.}$$

Dans la métrique discrète (exemple 4°, page 38), toutes les parties de E sont ouvertes.

Parties fermées

On appelle partie fermée de E toute partie de E dont le complémentaire est ouvert. En transformant par passage aux complémentaires les propriétés a, b, c des ouverts, on en déduit immédiatement des propriétés équivalentes pour les parties fermées :

(II, 2 ; 2)

a') E et $\emptyset$ sont fermées;

b') Toute réunion d'un nombre fini d'ensembles fermés est fermée.

c') Toute intersection d'une famille finie ou Infinie d'ensembles fermés est fermée. (*).

L'axiome de HAUSDORFF nè se traduit pas de manière intéressante pour les ensembles fermés.

* Le corollaire du théorème 15 donnera le meilleur critère pratique pour reconnaître qu'une partie est fermée.

Exemples - Toute partie réduite à un point ou à un nombre fini de points est fermée. Toute boule fermée est une partie fermée, ce qui justifie son nom. En particulier , sur la droite réelle $\mathbb{R}$, munie de sa métrique naturelle, tout intervalle fermé est fermé, alors qu'un intervalle semi-ouvert ou ouvert, n'est pas fermé. L'ensemble $\{x ; d(a,x) \geq R\}$ est fermé. Toute sphère est fermée.

Dans $\mathbb{R}^n$, muni de l'une quelconque des métriques définies par les normes de la page 40 , un pavé fermé, défini par des inégalités larges : $\{x = (x_1, x_2, \dots x_n) ; a_i \leq x_i \leq b_i , i = 1,2,\dots\}$ est un ensemble fermé.

Dans la métrique discrète, toutes les parties de E sont fermées.

Remarques

Z 1°/ Nous avons, dans les propriétés des ouverts et des fermés, soigneusement distingué s'il s'agissait d'intersections ou de réunions d'une famille quelconque, ou seulement d'un nombre fini de parties. Ces distinctions sont absolument essentielles.

Par exemple, sur la droite réelle $\mathbb{R}$, munie de sa métrique naturelle, une partie réduite à un point est un ensemble fermé mais n'est pas un ensemble ouvert, alors qu'elle est intersection d'une infinité dénombrable d'intervalles ouverts. De même un intervalle ouvert n'est pas fermé, alors qu'il est une réunion d'une infinité dénombrable d'intervalles fermés strictement plus petits.

Z 2°/ Il existe naturellement des parties de E qui ne sont ni ouvertes, ni fermées, par exemple les intervalles semi-ouverts sur $\mathbb{R}$ munie de sa métrique naturelle.

Z 3°/ Il peut exister, en dehors de $\emptyset$ et E , des parties de E qui sont à la fois ouvertes et fermées, comme le montre l'exemple de la métrique discrète * .

Voisinages

On appelle voisinage d'un point a de E toute partie de E contenant au moins un ouvert contenant lui-même a (ou encore contenant une boule ouverte ou une boule fermée de centre a).

Les voisinages d'un point a possèdent les propriétés suivantes, outre la propriété de contenir tous a :

(II,2;3)

a") Toute partie qui contient un voisinage de a est un voisinage de a ;

b") Toute intersection d'un nombre fini de voisinages de a est un voisinage de a .

c") Axiome de séparation de HAUSDORFF : Quels que soient les points distincts a et b de E , il existe un voisinage de a et un voisinage de b qui sont disjoints.

* Voir à ce sujet le § 9.

On doit noter que le mot français "voisinage" est ici quelque peu trompeur : on pourrait croire qu'un voisinage de a est un ensemble très voisin de a, et qu'il est d'autant plus facile pour une partie de E d'être un voisinage de a qu'elle est plus petite.

Or il n'en est rien. La propriété a) des voisinages montre au contraire qu'il est d'autant plus facile, pour une partie A, d'être un voisinage de a, qu'elle est plus grande. L'espace entier E est un voisinage de a, et, sauf des cas exceptionnels (comme la métrique discrète), la partie réduite à a n'est pas un voisinage de a.

On dit qu'un point de E est <u>isolé</u> si la partie réduite à ce point en est un voisinage, c'est-à-dire est ouverte. Dans la métrique discrète tout point est isole. Dans~ , aucun point n est isolé.

Si A est une partie de E, on appelle voisinage de A toute partie de E contenant un ouvert contenant A. Si ε est un nombre > 0, on appelle voisinage d'ordre ε de A la réunion des boules de rayon ε de centre dans A. Il contient la réunion des boules ouvertes de rayon ε et de centre dans A qui est un ouvert (réunion de boules ouvertes) contenant A, donc c'est bien un voisinage de A.

<u>Théorème 1</u> : <u>Pour qu'une partie A de E soit ouverte, il faut et il suffit qu'elle soit un voisinage de chacun de ses points.</u>

<u>Démonstration</u> - Si A est voisinage de chacun de ses points, alors elle ne peut contenir un point sans contenir un ouvert de E qui le contient, c'est-à-dire sans contenir une boule ayant pour centre ce point. Donc A est ouverte.

Réciproquement, si A est ouverte, et si $a \in A$, A contient bien un ouvert contenant a, à savoir A lui-même, donc A est un voisinage de a.

Soit $(\mathcal{V}_i)_{i \in I}$ une famille de voisinages de a dans E. On dit que c'est un système fondamental de voisinages de a, si tout voisinage de a contient l'un des $\mathcal{V}_i$. Par exemple, dans un espace métrique E, les boules ouvertes (ou les boules fermées) de centre a, ou les boule de centre a et de rayon rationnel, forment des systèmes fondamentaux de voisinages de a.

Tout point possède un système fondamental de voisinages ouverts, a savoir les boules ouvertes de centre a ; et un système fondamental de voisinages fermes, à savoir les boules fermées de centre a.

Intérieur

Soit A une partie d'un espace métrique E . On appelle intérieur de A et on note $\mathring{A}$ la réunion de tous les ouverts de E contenus dans A . C'est donc un ouvert de E contenu dans A (propriété c) page 42), et c'est alors bien évidemment le plus grand ouvert de E contenu dans A . Bien entendu, $\mathring{A}$ peut être vide (exemple page 48 : si $E=\mathbb{R}$, $A=\mathbb{Q}$).

Théorème 2 - L'intérieur $\mathring{A}$ de A est l'ensemble des points de E qui sont centre d'au moins une boule contenue dans A . ou encore c'est l'ensemble des points de E dont au moins un voisinage est tout entier dans A , ou l'ensemble des points de E dont A est un voisinage.

Démonstration - Si en effet a est centre d'une telle boule ouverte, celle-ci est un ouvert de E contenu dans A, donc elle est contenue dans $\mathring{A}$, et à fortiori $a \in \mathring{A}$.

Réciproquement, si $a \in \mathring{A}$ comme $\mathring{A}$ est ouverte, il existe une boule de centre a contenue dans $\mathring{A}$, et à fortiori dans A .

Extérieur

On appelle extérieur de A l'intérieur de son complémentaire. C'est donc le plus grand ouvert de E qui soit disjoint de A , c'est l'ensemble des points de E qui sont centre d'au moins une boule de rayon > 0 disjointe de A. L'intérieur et l'extérieur de A sont évidemment disjoints.

Frontière

- On appelle frontière de A , et on note $\dot{A}$, l'ensemble des points de E qui n'appartiennent ni à son intérieur, ni à son extérieur.

Son complémentaire, réunion de 2 ouverts, est ouvert, donc elle est fermée.

D'après les propriétés de ces deux parties, il en résulte immédiatement les propriétés suivantes.

Théorème 3 - Pour qu'un point a appartienne à la frontière $\dot{A}$ de A , il faut et il suffit que tout voisinage de a contienne à la fois des points de A et de son complémentaire.

Tout d'abord, si tout voisinage $\mathcal{V}$ de a rencontre à la fois A et $\complement A$, il n'est contenu dans aucun d'entre eux, donc a n'est ni intérieur ni extérieur, et il est bien sur la frontière. Inversement, si a est sur la frontière, et si $\mathcal{V}$ est un voisinage de a $\mathcal{V}$ rencontre A, sans quoi il serait dans CA et a serait extérieur, et il rencontre $\complement A$, sans quoi il serait dans A et a serait intérieur.

L'intérieur, l'extérieur et la frontière de A ont pour réunion E , et ils sont deux à deux disjoints.

Adhérence

- On appelle adhérence de A , et on note $\bar{A}$, l'intersection de toutes les parties fermées de E qui contiennent A. Elle est donc fermée (propriété c') page 43), et c'est alors évidemment la plus petite partie fermée de E qui contienne A.

Un point appartenant à cette adhérence est dit adhérent à A

Théorème 4 - L'adhérence de A est la réunion de son intérieur et de sa frontière, autrement dit le complémentaire de son extérieur . Ou encore c'est l'ensemble des points de E dont tout voisinage rencontre A * .

Démonstration - Dire que $\bar{A}$ est le plus petit fermé qui contient A est equivalent, d'après la propriété de passage au complémentaire, à dire que $\complement\bar{A}$ est le plus grand ouvert contenu dans $\complement A$, donc n'est autre que l'extérieur, ce que nous voulions démontrer.

Soit maintenant a un point adhérent à A . Alors tout voisinage $\mathcal{V}$ de a rencontre A ; sinon, en effet, $\mathcal{V}$ serait contenu dans $\complement A$, il existerait alors un ouvert $\mathcal{O}$ contenant a et contenu dans $\mathcal{V}$ donc dans $\complement A$; alors $\mathcal{O}$, et donc a , serait dans la réunion de tous les ouverts contenus dans $\complement A$,qui est l'extérieur de A , ceci est contraire à l'hypothèse $a \in \bar{A}$. Réciproquement, si a est tel que chacun de ses voisinages rencontre A , a ne peut pas être dans l'extérieur de A , sans quoi cet extérieur serait un voisinage de a ne rencontrant pas A ; donc a est dans l'adhérence $\bar{A}$.

Nous avons donc bien montré que a est adhérent à A si et seulement si tout voisinage de a rencontre A .

Donnons quelques exemples : Dans un espace vectoriel normé, si A est une boule ouverte ou fermée de centre a et de rayon R > 0, son intérieur est la boule ouverte correspondante,son adhérence est la boule fermée correspondante, son extérieur est l'ensemble $\{x \; ; \; d(a,x) > R\}$, et sa frontière est la sphère de centre a et de rayon R . *

[Notons que cette circonstance n'est pas absolument générale.

Si E est muni de la métrique discrète, et si A est une boule fermée de rayon 1 , elle est identique à E donc identique à son intérieur, tandis que sa frontière et son extérieur sont vides. Une boule ouverte de rayon 1 est réduite à son centre).

* Le théorème 15 donnera une autre caractérisation essentielle de l'adhérence.

Sur la droite réelle $\mathbb{R}$, si A est l'ensemble des nombres rationnels, $\mathring{A} = \emptyset$, $\bar{A} = \bar{A} = \mathbb{R}$.

Pour qu'une partie A de E soit ouverte, il faut et il suffit qu'elle soit identique a son Intérieur.

Pour qu'elle soit fermée, il faut et il suffit qu'elle soit identique à son adhérence.

Sous-ensembles denses

Une partie A d'un espace métrique E est dite "dense" dans E si tout point de E lui est adhérent, c'est-à-dire si son adhérence est E lui-même. Cela veut encore dire que tout ouvert rencontre A. Un espace métrique E est dit séparable s'il est fini ou s'il contient une partie dénombrable dense. (Cette dénomination, assez largement adoptée, est fâcheuse, car separable n'a ainsi aucun rapport avec séparé !)

Exemples - Sur $\mathbb{R}$ munie de sa métrique naturelle, l'ensemble A des nombres rationnels, et l'ensemble B des nombres irrationnels sont denses. Comme l'ensemble des rationnels est dénombrable, $\mathbb{R}$ est séparable.

Sous-espace. Métrique induite

Soit F une partie d'un espace métrique E La restriction à $F \times F$ de la fonction distance définie sur $E \times E$, fait de F lui-même un nouvel espace métrique. On l'appelle un sous-espace-métrique de E, et on dit qu'il est muni de la métrique "induite".

Si alors A est une partie de F, il y a lieu de préciser avec soin, quand on dira qu'elle est ouverte ou fermée, si elle l'est dans l'espace métrique E, ou dans l'espace métrique F. Par exemple F lui-même est à la fois ouvert et fermé dans l'espace métrique F, alors qu'il ne l'est pas en général dans E.

Théorème 5 - Pour qu'une partie A de F soit ouverte (resp. fermée) dans l'espace métrique F, il faut et il suffit qu'elle soit l'intersection de F et d'une partie ouverte (resp. fermée) de l'espace métrique E.

Pour qu'une partie de F soit, dans l'espace métrique F, un voisinage de $a \in F$, il faut et il suffit qu'elle soit l'intersection de F et d'un voisinage de a dans E.

Démonstration -

Désignons par B des boules dans E, par β les boules dans F.

Soit A une partie de F, intersection de F et d'un ouvert A_1 de E . Si $a \in F$ est dans A , donc dans A_1 , il existe une boule $B_o(a, R)$ contenue dans A_1 • alors $\beta_o(a,R) = B_o(a,R) \cap F$ est contenue dans A ; A est bien ouvert dans F .

Réciproquement soit A un ouvert de F . Alors il est (page 43) une réunion d'une famille $\beta(a_i, R_i)$, $i \in I$, de boules ouvertes, $a_i \in A, R_i > o$; alors la réunion des $B(a_i, R_i)$ est un ouvert A_1 de E , et $A = A_1 \cap F$.

Ainsi la propriété relative aux ouverts est démontrée.

Soit maintenant A_1 une partie fermée dans E . soit C_1 son complémentaire dans E . A_1 et C_1 coupent F suivant deux parties complémentaires A et C de F : C est ouverte donc $A = A_1 \cap F$ est fermée.
Réciproquement, soit A une partie fermée de F . Soit C son complémentaire relativement à F . C est ouverte, donc il existe un ouvert C_1 de E tel que $C_1 \cap F = C$. Soit A_1 le complémentaire de C_1 dans E , A_1 est fermée. A_1 et C_1 étant complémentaires dans E , leurs intersections avec F sont complémentaires dans F ; comme $C_1 \cap F = C$, on a $A = A_1 \cap F$, A_1 étant fermée dans E , ce qui démontre la propriété relative aux fermés.

Enfin soit $\mathcal{V}_1$ un voisinage de $a \in F$ dans E . $\mathcal{V}_1$ contient un ouvert A_1 contenant a . Alors $\mathcal{V} = \mathcal{V}_1 \cap F$ contient l'ouvert $A = A_1 \cap F$ de F , contenant a , donc c'est un voisinage de a dans F.
Réciproquement, soit $\mathcal{V}$ un voisinage de a dans F . Il contient un ouvert A de F contenant a . Il existe alors un ouvert A_1 de E tel que $A = A_1 \cap F$ Alors $\mathcal{V}_1 = A_1 \cup \mathcal{V}$ est un voisinage de a dans E puisqu'il contient l'ouvert A_1 contenant a ; et $\mathcal{V} = \mathcal{V}_1 \cap F$.

Ceci montre deux choses :

1°/ Les ouverts, les fermés, et par conséquent les voisinages d'un sous-espaces métrique F sont parfaitement connus dès que l'on connait ceux de E , sans qu'il soit nécessaire de connaître la fonction distance elle-même. Nous en verrons plus tard l'intérêt (voir § 4)

2°/ <u>Quelle que soit la partie F de E , si une partie A de F est ouverte (resp. fermée) dans l'espace métrique E , elle l'est à fortiori dans le sous-espace métrique F; si elle est un voisinage de a, $\in F$ dans E , elle l'est a fortiori dans F .</u>

La réciproque n'est pas necessairement vraie, comme le montre le cas de **A** = **F** lui-même vu plus haut.

Mais :

Théorème 6 - a) Si F est un ouvert de **E** , toute partie **A** de F, ouverte dans l espace métrique F , est encore ouverte dans l'espace métrique E .

b) Si F est une partie fermée de E, toute partie **A** de F , fermée dans l'espace métrique F , est encore fermée dans l'espace métrique E .

c) Si F est un voisinage de a dans E , toute partie A de F , voisinage de a dans F , est encore un voisinage de a dans E .

Démonstration

a/ Si **A** est une partie ouverte de **F** , supposé ouvert dans **E** , il existe, d après le théorème **5**, une partie A_1 , ouverte de **E** , telle que **A** = $A_1 \cap$ **F** ; comme alors A_1 et F sont ouvertes dans E, $A = A_1 \cap F$ l'est aussi. On en déduit aussitôt c) .

b/ Démonstration analogue, en remplaçant ouvert par fermé .

§ 3 FONCTIONS CONTINUES. HOMÉOMORPHISMES

Soit f une application d'un espace métrique E dans un espace métrique F . On dit que f est continue en un point a de E , si, quel que soit $\varepsilon > 0$, il existe $\eta > 0$ tel que $d(a,x) \leq \eta$ entraîne $d(f(a), f(x)) \leq \varepsilon$.

On peut aussi dire : si quelle que soit la boule ayant pour centre $f(a)$, il existe une boule de centre a dont l'image par f soit dans la précédente.

On peut encore dire : si, quel que soit $\mathcal{V}$, voisinage de de $f(a)$, il existe $\mathcal{U}$, voisinage de a , tel que $f(\mathcal{U}) \subset \mathcal{V}$.

On peut encore dire : si l'image réciproque par f de tout voisinage de $f(a)$ est un voisinage de a .

Une application de E dans **F** est dite continue, si elle continue en tout point a de **E** .

La première définition de la continuité fait essentiellement intervenir la métrique. Les deux dernières, au contraire, ne font intervenir que les ouverts et les voisinages, mais non la métrique elle même. On en verra l'intérêt plus tard (§ 4).

Exemple - La fonction $\frac{1}{x}$, c'est-à-dire l'application $x \longrightarrow \frac{1}{x}$, de l'espace métrique E , complémentaire de l'origine sur la droite réelle, dans l'espace métrique $F = \mathbb{R}$, droite réelle, est une application partout continue.

Théorème 7 - Pour qu'une application d'un espace métrique E dans un espace métrique F soit continue, il faut et il suffit que l'image réciproque par f de tout ouvert de F soit un ouvert de E .

Démonstration - Montrons d'abord que la condition est nécessaire : Supposons f continue, soit B un ouvert de F, et posons $A = f^{-1}(B)$ soit $a \in A$. f est continue en a , or B est un voisinage de $f(a)$, donc A doit être un voisinage de a . Ainsi A est un voisinage de chacun de ses points, donc c'est un ensemble ouvert (théorème 1).

Montrons maintenant que la condition est suffisante : Supposons que l'image réciproque par f de tout ouvert de F soit un ouvert de E . Alors, pour tout point a de E , soit $\mathcal{V}$ un voisinage de $f(a)$ dans F .

Alors $\mathcal{V}$ contient un ouvert B contenant b , donc l'image réciproque $f^{-1}(\mathcal{V})$ contient $f^{-1}(B)$ qui est un ouvert contenant a. Alors $f^{-1}(\mathcal{V})$ est un voisinage de a , ce qui prouve que l'application f est continue en a .

Théoreme 8 - Pour qu'une application f d'un espace métrique E dans un espace métrique F , soit continue, il faut et il suffit, que l'image réciproque par f de tout fermé de F soit un fermé de E .

Démonstration : On passe du théorème précédent à celui-ci en remplaçant les parties ouvertes considérées dans E et dans F par leurs parties complémentaires fermées, et en utilisant la formule (1, 2;3).

Remarque : Si, dans les deux théorèmes précédents, on remplaçait les images réciproques par des images directes on aboutirait à des résultats inexacts. Considérons par exemple, en premier lieu, une application constante de E dans F. Une telle application est manifestement continue. Cependant l'image par cette application de n'importe quel ouvert de E, en particulier de E lui-même, est réduite à un point de F, et en général une partie réduite à un point n'est pas ouverte. Si par ailleurs nous considérons la fonction $\frac{1}{x}$ définie dans l'exemple ci-dessus, l'image par cette fonction de l'ensemble E tout entier c'est-à-dire d'une partie fermée, est, dans $F = \mathbb{R}$, le complémentaire de l'origine, qui n 'est pas une partie fermée, Théorème 9 - Si E est un espace vectoriel normé, sa norme, application de E dans la droite $\mathbb{R}$ munie de sa métrique naturelle, est une fonction continue.

Démonstration :

On en deduit en effet de (II,1;6) que, si $\varepsilon > 0$ est donné, il suffit de choisir $\eta = \varepsilon$ pour que $\|x-a\| \leqslant \eta$ entraîne $\big| \|x\| - \|a\| \big| \leqslant \varepsilon$.

Théorème 10 - L'application composée de deux applications continues est continue.

Démonstration Soient E, F, **G**, trois espaces métriques, et $h = g \circ f$ l'application composée d'une application! de **E** dans F, et d'une application g de F dans **G** . On suppose en outre f continue en un point a de E, et g continue au point $b = f(a)$ de F .

soit $c = g(b) = h(a)$. Soit $\mathcal{W}$ un voisinage de c dans **G** . L'application g étant continue au point 8 , l'image réciproque $\mathcal{V} = g^{-1}(\mathcal{W})$ est un voisinage de b dans F .

L'application f étant continue au point a , l'image réciproque $\mathcal{U} = f^{-1}(\mathcal{V})$ est un voisinage de a dans E . Mais $\mathcal{U} = f^{-1}(g^{-1}(\mathcal{W}))$ n'est autre que $h^{-1}(\mathcal{W})$, et alors ceci prouve bien que h est continue au point a .* . On en déduit bien évidemment que, si f et g sont partout continues, alors h est aussi partout continue. On peut d'ailleurs le voir directement en utilisant le théorème 7 ou le théorème 8.

Homéomorphismes

On appelle homéomorphisme d'un espace métrique E sur un espace métrique F toute bijection de E sur F qui soit continue ainsi que sa bijection réciproque.

* A titre d'exercice, donner une autre démonstration en utilisant la première définition (métrique) de la continuité, avec Ve, $\exists$. η. . .

Théorème 11 - Pour qu'une application f, déjà bijective et continue, de E dans F, soit un homéomorphisme, il est nécessaire et suffisant que l'image directe par f de tout ouvert de E soit un ouvert de F. Il est aussi nécessaire et suffisant que l'image directe par f de tout fermé de E soit un fermé de F.

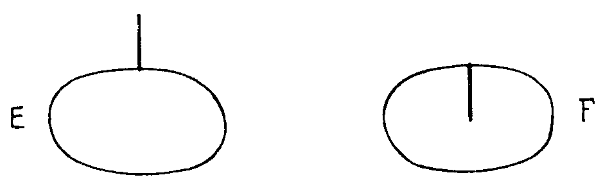
En effet ces images directes ne sont autres que des images réciproques, relatives à la bijection réciproque $g = f^{-1}$, et les conditions précédentes ne sont autres que celles qui sont données dans le théorème 7 et dans le théorème 8 pour la continuité de f^{-1}.

Remarque - Il ne faudrait pas croire que toute application bijective et continue soit nécessairement un homéomorphisme. Par exemple, si E est la droite $\mathbb{R}$ munie de sa métrique discrète, et si F est la droite $\mathbb{R}$ munie de sa métrique naturelle, l'application identique de E dans F est continue et bijective, mais n'est manifestement pas un homéomorphisme.

On dit que deux espaces métriques E et F sont homéomorphes, s'il existe au moins un homéomorphisme de l'un sur l'autre. Ces deux espaces ont alors les mêmes propriétés topologiques, c'est-à-dire les mêmes propriétés pour tout ce qui concerne les ensembles ouverts, les ensembles fermés et les voisinages.

Exemple - L'intérieur d'un disque et l'intérieur d'un triangle, dans un plan euclidien, sont des espaces métriques homéomorphes. Le demi-plan $y > 0$, la région $y > x^2$ située au-dessus de la parabole $y = x^2$, la région $y < x^2$ située au dessous de cette parabole, dans le plan $\mathbb{R}^2$, sont homéomorphes *.

Les deux espaces métriques définis par les lignes tracées sur la figure qui suit sont homéomorphes.



* A titre d'exercice, définir chaque fois un homéomorphisme entre les espaces métriques homéomorphes considérés.

Z

Mais attention ! Cela ne veut nullement dire qu'il existe un homéomorphisme du premier plan sur le deuxième, qui amène le premier sous-espace sur le deuxième !

La droite réelle, munie de sa métrique naturelle, et la droite réelle, munie de la métrique discrète, ne sont pas homéomorphes, puisque, sur cette dernière, toutes les parties sont ouvertes et qu'il n'en est pas de même sur la première *

§ 4 ESPACES MÉTRIQUES ET ESPACES TOPOLOGIQUES

A partir de la notion de distance, nous avons pu définir, sur un espace métrique E les notions d'ensemble ouvert, d'ensemble fermé, de voisinage, d'intérieur, d'extérieur, de frontière, d'adhérence, d'ensemble dense, d'application continue. Toutes ces notions de déduisent de celle d'ensemble ouvert.

Il peut arriver que deux métriques différentes, sur le même espace E, aient le même système d'ensembles ouverts. Elles ont alors les mêmes parties fermées, les mêmes voisinages de chaque point, etc....

Par exemple, si E est un espace métrique et si d est sa fonction distance, la fonction distance $2d$, c'est-a-dire telle que la distance de deux points x et y soit $2\,d(x,y)$, donne bien évidemment les mêmes ensembles ouverts.

On dit que deux métriques sur un même ensemble E sont équivalentes, si elles ont le même système d'ensembles ouverts. On dit encore qu'elles définissent la même topologie sur E. Cela revient à dire que l'application identique de E, muni de la 1ère métrique, sur E, munie de la 2ème metrique, est un homeomorphisme.

Sur un espace vectoriel, 2 normes sont dites équivalentes, si les métriques correspondantes-sont équivalentes.

Théorème 12 - Pour que 2 normes sur un espace vectoriel notées sous la forme $\| \; \|_1$ et $\| \; \|_2$ soient équivalentes, il faut et il suffit qu'il existe des Constantes $k', k'' > 0$ telles que l'on ait, pour tous les $\vec{x}$ de E :

$$\|\vec{x}\|_2 \leqslant k'\|\vec{x}\|_1 \; , \quad \|\vec{x}\|_1 \leqslant k''\|\vec{x}\|_2 \tag{II,4;1}$$

* A la page 53, nous avons indiqué que l'application identique n'était pas un homéomorphisme de 1 une sur l'autre. Nous disons ici qu'il n'existe aucun homéomorphisme de l'une sur l'autre, ce qui est un résultat plus fort.

<u>Démonstration</u> - Appelons $B_1(R)$ (resp. $B_2(R)$) la boule fermée de centre origine et de rayon R, pour la première norme (resp. pour la 2ème). Supposons alors que les deux normes soient équivalentes, c'est-à-dire donnent les mêmes ouverts.

Alors $B_1(1)$ contient un ensemble ouvert de E, dans la lère métrique, contenant l'origine: il contient donc aussi un ensemble ouvert, dans la 2ème, contenant l'origine; il existe par suite un nombre $\frac{1}{k''}$ tel que l'on ait $B_1(1) \supset B_2\left(\frac{1}{k''}\right)$. Par une homothétie de rapport $k''R$, on en déduit la relation d'inclusion $B_1(k''R) \supset B_2(R)$. Elle signifie que $\|\vec{x}\|_2 \leqslant R$ entraîne $\|\vec{x}\|_1 \leqslant k''R$; mais la première inégalité est vraie avec $R = \|\vec{x}\|_2$, la seconde donne alors $\|\vec{x}\|_1 \leqslant k''\|\vec{x}\|_2$. *

En opérant de même en sens inverse, la nécessité de la condition est bien démontrée. Exprimons maintenant que cette condition est suffisante. Si elle est réalisée, alors $\|\vec{x}\|_2 \leqslant \frac{R}{k''}$, entraîne $\|\vec{x}\|_1 \leqslant R$.

Donc : $$B_2\left(\frac{R}{k''}\right) \subset B_1(R).$$

Il en résulte que toute boule, pour la lère métrique, contient nécessairement une boule pour la 2ème, et vice versa; ce que nous venons de dire pour les boules ayant pour centre l'origine est vrai, par translation, pour les boules ayant un centre quelconque.

Comme un ouvert pour la topologie est un ensemble qui, toutes les fois qu'il contient un point, contient au moins une boule ayant pour centre ce point, les propriétés que nous venons de voir pour les boules entrainent l'identité des ouverts pour les deux métriques.

<u>Remarque</u> - Si d_1 et d_2 sont les distances définies par 2 normes équivalentes, il y a donc une constante k telle que, quels que soient x et y, $d_1(x,y) \leqslant k\, d_2(x,y)$, $d_2(x,y) \leqslant k\, d_1(x,y)$.

<u>Cette circonstance est très spéciale aux métriques équivalentes définies par des normes équivalentes sur un espace vectoriel.</u>

* Voici une autre démonstration. S'il n'existait pas de tel nombre k'', alors, pour tout entier $n \geqslant 0$ il existerait un point $\vec{x}_n \neq \vec{0}$ tel que $\|\vec{x}_n\|_1 \geqslant n\|\vec{x}_n\|_2$. En remplaçant au besoin $\vec{x}_n$ par un homothétique, on peut toujours supposer que $\|\vec{x}_n\|_1 = 1$. Alors $\|\vec{x}_n\|_2 \leqslant \frac{1}{n}$. Donc la suite des $\vec{x}_n$ convergerait vers 0 pour la 2ème norme et pas pour la première, et elles ne seraient pas équivalentes (cette démonstration utilise les suites convergentes, qui seront vues plus loin).

Mais, si d est la distance d'une métrique quelconque sur un ensemble E , on voit aussitôt que $d' = \mathrm{Inf}\,(d\ 1)$, définie par $d'(x\ y) = \mathrm{Min}(d(x,y),1)$, est aussi'une fonction distance' (vérifier (II,1;1)). Il est trivial que la métrique définie par d' est équivalente a la métrique initiale (les boules de rayon ≤ 1 sont les mêmes). Or, lorsque d varie de 0 à $+\infty$ d' varie de 0 à 1. Il n'y a donc pas entre elles d'inégalités du type (II,4;1). Si E est un espace vectoriel, et si d est la distance définie a partir d'une norme, il n'en est plus de même de d'. Cet exemple montre que, étant donné une métrique sur E , on peut toujours trouver une métrique eouivalente pour laquelle E soit borné.

Corollaire - Les 3 normes données au début pour l'espace vectoriel $\mathbb{R}^n$ sont équivalentes.

On a en effet, les inégalitis :

$$(\mathrm{II},4;2)\quad \begin{cases} \dfrac{1}{n}\displaystyle\sum_{i=1}^{i=n}|x_i| \leqslant \operatorname*{Max}_{i=1,2,\dots n}|x_i| \leqslant \sum_{i=1}^{i=n}|x_i| \\ \dfrac{1}{\sqrt{n}}\left(\displaystyle\sum_{i=1}^{i=n}|x_i|^2\right)^{\frac{1}{2}} \leqslant \operatorname*{Max}_{i=1,2,\dots n}|x_i| \leqslant \left(\displaystyle\sum_{i=1}^{i=n}|x_i|^2\right)^{\frac{1}{2}} . \end{cases}$$

Plus généralement nous admettrons le théorème suivant, dont la démonstration est délicate *

Théorème 13 - Sur un espace vectoriel de dimension finie sur le corps des réels ou des complexes, 2 normes quelconques sont équivalentes; il existe par suite un seul système d'ensembles ouverts, d'ensembles fermes, etc.... valable pour toutes les normes possibles.

Z Il est bon de remarquer, comme nous en verrons plus tard des exemples, que cette propriété ne subsiste absolument pas pour des espaces vectoriels de dimension infinie.

Nous sommes ainsi amenés, pour des espaces métriques, à introduire deux sortes de propriétés : les propriétés métri-

* On trouvera une démonstration (facultative) à la page 72 avec le théorème 23.

ques qui dépendent explicitement de la métrique elle-même, comme la distance de deux points, la propriété des côtés d'un triangle ou d'une figure formée par plusieurs points, les boules, etc...; d'autre part, les propriétés topologiques qui ne dépendent pas de la métrique elle-même mais seulement de l'ensemble des parties ouvertes, des parties fermées * etc.... P us généralement, on ooncoit qu'il soit même possible d'introduire une topologie, sans passer par l'intermédiaire d'une métrique :

On appelle espace topologique E , un ensemble E sur lequel on a aistinguéune famille de parties, appelées les parties ouvertes de la topologie.

C'est une famille absolument quelconque de parties astreinte seulement à satisfaire aux propriétés a, b, c, d, données page 42 . **

Nous voyons qu'un espace métrique est un espace topologique particulier, mais il existe des espaces topologiques qui ne peuvent pas être définis à partir d'une métrique. i.e. non-métrisables.

* Le fait, pour une partie A de E , d'être bornée,(voir définition page 39) est une propriété métrique et non topologique: l ensemble $\mathbb{N}_1$ des entiers $\geqslant 1$ n est pas borné si on le munit de la métrique $d(p,q) = |p-q|$, il devient borné si on le munit de la métrique $d(p,q) = \left|\frac{1}{p} - \frac{1}{q}\right|$, alors que ces deux métriques sont équivalentes. Mais, pour les espaces vectoriels normés, 2 normes équivalentes donnent les mêmes parties bornées, d après le theorème 12. Voir remarque de la page 55

** Généralement, on n'impose aux ouverts d'une topologie que les axiomes a, b, c . Si en outre l'axiome d est vérifié, on dit que l espace topologique est séparé. Les espaces non séparés n'ont qu'un usage très limite en analyse , et nous n'en aurons jamais besoin. C'est pourquoi, pour nous, un espace topologique sera toujours supposé sépare, et les ouverts devront vérifier l axiome d .

Dans un espace non sépare, une partie rdduite à un point n'est pas nécessairement fermée, et une suite convergente peut avoir plusieurs limites distinctes (voir théorème 14)!

On dit qu'un espace topologique est métrisable s'il existe une métrique qui donne naissance à sa topologie.

La plus grande partie des espaces topologiques que nous rencontrerons sont des espaces métrisables.

Il est bon de remarquer que presque toutes les définitions et les théorèmes que nous avons donnés jusqu'à présent pour les espaces métriques, faisaient simplement intervenir la topologie et non la métrique elle-même; on démontre facilement qu'ils sont vrais pour des espaces topologiques quelconques. Toutefois les théorèmes 2, 9, 12 et 13 ne sont vrais que pour des espaces métriques, car ils faisaient essentiellement intervenir la métrique dans leur énoncé même.

Tout point possède toujours un système fondamental de voisinages ouverts (puisque tout voisinage de a contient un ouvert contenant a, qui est un voisinage ouvert de a), mais pas nécessairement un système fondamental de voisinages fermés. Un espace topologique est dit régulier si tout point a un système fondamental de voisinages fermés; un espace topologique métrisable est régulier., Par contre, il est toujours vrai que l'intersection des voisinages fermés de a se réduit à a (donc a fortiori que l'intersection de tous les voisinages de a se réduit à a) : si en effet $b \neq a$ il existe, d'après l'axiome de Hausdorff, des ouverts $\mathcal{A}, \mathcal{B}$, contenant a, b, et d'intersection vide; $\mathcal{A}$ est un voisinage fermé de a, et $\mathcal{A} \subset \complement \mathcal{B}$ entraîne $\bar{\mathcal{A}} \subset \complement \mathcal{B}$, donc $\bar{\mathcal{A}}$ ne contient pas b, et b n'appartient pas à tous les voisinages fermés de a.

Le théorème 5 sert à définir la topologie induite : si E est un espace topologique, F une partie de E on peut faire de F un espace topologique, en prenant comme ouverts les intersections avec F des ouverts de la topologie de E. On dit alors que F est un sous-espace topologique de E, ou que sa topologie est la topologie induite par celle de E.

Dans la suite, nous énoncerons les théorèmes dans les espaces topologiques, toutes les fois que ce sera possible; mais nous ne nous gênerons pas pour ne donner la démonstration que dans les espaces métriques, si cela doit simplifier. Les élèves ne seront tenus de connaître que ce qui est relatif aux espaces métriques.

Topologie de la droite achevée $\bar{\mathbb{R}}$. (Voir page 22 du Chap. 1)

On munit $\bar{\mathbb{R}}$ d'une topologie, en définissant ses ensembles ouverts de la façon suivante : Une partie U de $\bar{\mathbb{R}}$ est ouverte lorsque :

a) Si elle contient un point x de $\mathbb{R}$, elle contient au moins un intervalle ouvert contenant x ;

b) si elle contient le point $-\infty$, elle contient au moins un intervalle $[-\infty, A[$;

c) si elle contient le point $+\infty$, elle contient au moins un intervalle $]A, +\infty]$.

On démontre que l'ensemble d'ouverts ainsi-défini satisfait bien à tous les axiomes voulus pour faire de $\overline{\mathbb{R}}$ un espace topologique. Il est facile de voir que $\overline{\mathbb{R}}$, pour cette topologie, est métrlsable, et nous donnerons plus loin une infinité de métriques equivalentes donnant naissance à cette topologie (théorème 38, page 93); mais aucune de ces métriques ne s'impose de façon naturelle plus que les autres. $\mathbb{R}$ est un sous-ensemble de $\overline{\mathbb{R}}$, et sur ce sous-ensemble la topologie induite par celle de $\overline{\mathbb{R}}$ est la topologie définie par la métrique naturelle.

§ 5 SUITES. LIMITES. CONVERGENCES

Une propriété essentielle des espaces topologiques est qu'il est possible, dans ces espaces, de parler de suites convergentes.

1°/ soit $x_0, x_1, x_2, \dots x_n, \dots$ Une suite de points d'un espace métrique E On dit que cette suite est convergente vers un point ℓ de E ou a pour limite ℓ * si la suite de nombres réels $d(\ell, x_0), d(\ell, x_1), \dots d(\ell, x_n), \dots$ converge vers 0 .

Cela revient à dire que, quel que soit $\varepsilon > 0$, il existe un entier n_0 tel que $n \geqslant n_0$ entraîne $d(\ell, x_n) \leqslant \varepsilon$. On Peut encore dire si, quel que soit le voisinage $\mathcal{V}$ de ℓ , il existe un entier n_0 tel que, pour $n \geqslant n_0$, tous les x_n appartiennent à $\mathcal{V}$.

On peut encore dire : si, quel que soit le voisinage $\mathcal{V}$ de ℓ , tous les x_n appartiennent à $\mathcal{V}$, sauf au plus pour un nombre fini de valeurs de l'entier n. **

Ces deux dernières définitions sont valables si E est un espace topologique; même si E est métrique, la convergence d'une suite est une propriété topologique, et non métrique.'

Il existe bien d'autres notions de limites qui ne sont pas relatives aux suites :

2°/ Considérons par exemple une suite double $x_{m,n}$, m entier $\geqslant 0$, n entier $\geqslant 0$, c'est-à-dire une application de $\mathbb{N} \times \mathbb{N}$ dans E.

On dit que cette suite double converge vers l'élément ℓ de E lorsque m et n tendent simultanément vers l'infini,

* On sous-entend l'expression : "Quand n tend vers + ∞ "
Si les x_n sont réels, on peut prendre $\ell = \pm \infty$; on considère alors que la suite est sur $\overline{\mathbb{R}}$.

** Cette définition montre en outre que la convergence d'une suite est Indépendante de l'ordre de ses termes. Changer l'ordre des termes d'une suite $x_0, x_1, x_2, \dots x_n, \dots$, c'est la remplacer par la suite $x_{p_0}, x_{p_1}, \dots x_{p_n}, \dots$, où $n \to p_n$ et une bijection de $\mathbb{N}$ sur lui-même; si la suite initiale converge vers ℓ, il en est de même de la suite modifiée.

si, quel que soit $\varepsilon > 0$, il existe des entiers m_0, n_0 tels que $m \geq m_0$, $n \geq n_0$, entraîne $d(x_{m,n}, \ell) \leq \varepsilon$ * .

3°/ On dit au contraire que $x_{m,n}$ converge vers ℓ lorsque m ou n tend vers l'infini, si quel que soit $\varepsilon > 0$, il existe des entiers m_0, n_0 tels que $m \geq m_0$ ou $n \geq n_0$ entraîne $d(x_{m,n}, \ell) \leq \varepsilon$ ** .

4°/ Si f est une application de la droite réelle $\mathbb{R}$ dans E l'expression " $f(x)$ tend vers ℓ lorsque x tend vers a par valeurs strictement supérieures" signifie què, quel que soit $\varepsilon > 0$, il existe $\eta > 0$ tel que $(|x-a| \leq \eta, x > a)$ entraîne $d(f(x), \ell) \leq \varepsilon$.

5°/ Dans les mêmes conditions, l'expression " $f(x)$ tend vers ℓ lorsque x tend vers $+\infty$ " signifie que, quel que soit $\varepsilon > 0$, il existe A réel tel que $x \geq A$ entraîne $d(f(x), \ell) \leq \varepsilon$.

Toutes ces limites se définissent aisément si E est un espace topologique, non nécessairement métrique. Elles rentrent dans un cadre bien plus général, que voici. Soit X un espace topologique, A une partie de X , a un point de X adhérent à A . On dira, si f est une application de A dans E, que " $f(x)$ tend vers ℓ lorsque x tend vers a par valeurs dans A *** ", si, quel que soit le voisinage $\mathcal{V}$ de ℓ dans E , il existe un voisinage $\mathcal{U}$ de a dans X tel que $f(\mathcal{U} \cap A) \subset \mathcal{V}$.
Ainsi, dans le cas des suites (cas 1°/), $X = \bar{\mathbb{R}}$, $A = \mathbb{N}$, $a = +\infty$
Dans 4°/, $X = \mathbb{R}$, $A = \{x \in \mathbb{R} ; x > a\}$.

Dans 5°/, $X = \bar{\mathbb{R}}$, $A = \mathbb{R}$, $a = +\infty$.

Dire qu'une application f d'un espace topologique E dans un espace topologique F est continue en un point a de E ,

* En abrégé : $(\forall \varepsilon > 0) (\exists m_0 \in \mathbb{N}) (\exists n_0 \in \mathbb{N}) (\forall m, m \geq m_0) (\forall n, n \geq n_0) : d(x_{m,n}, \ell) \leq \varepsilon$.

** En abrégé : $(\forall \varepsilon > 0) (\exists m_0 \in N) (\exists n_0 \in N) (\forall (m,n) \in N \times N$, $m \geq m_0$ ou $n \geq n_0) : d(x_{m,n}, \ell) \leq \varepsilon$.

*** Très souvent, $a \in \bar{A}$ mais $a \notin A$. C'est ce qu'on suppose toujours en taupe : quand on dit que x tend vers a , on suppose $x \neq a$. Ici nous ne nous placerons pas forcément dans ce cas : on peut avoir $a \in A$ ou $a \notin A$.

Z

c'est alors exactement dire que $f(x)$ tend vers $f(a)$ quand x tend vers a.

Théorème 14 - Si une suite admet une limite, cette limite est nécessairement unique.

Démonstration - Supposons qu'une suite $x_0, x_1, x_2, \dots x_n, \dots$ d'éléments de E puisse admettre deux limites distinctes a, b de E. On sait qu'il existe, d'après l'axiome de séparation de HAUSDORFF, un voisinage $\mathcal{U}$ de a, et un voisinage $\mathcal{V}$ de b, qui sont sans point commun. Alors il doit exister d'une part un entier m_0 tel que $n \geqslant m_0$ entraîne $x_n \in \mathcal{U}$, et d'autre part un entier n_0 tel que $n \geqslant n_0$ entraîne $x_n \in \mathcal{V}$.

Alors $n \geqslant \mathrm{Max}(m_0, n_0)$ entraîne $x_n \in \mathcal{U} \cap \mathcal{V}$, ce qui est absurde puisque cette intersection est vide.

Théorème 15 - Pour qu'un point a d'un espace métrisable E soit adhérent à une partie A de E, il faut et il suffit qu'il existe une suite d'éléments de A qui converge vers a *.

Démonstration - Il est évident que la condition est suffisante, car, s'il existe une telle suite, alors tout voisinage $\mathcal{V}$ de a contient au moins un point de la suite et par conséquent un point de A, et a est bien adhérent à A (théorème 4)

Réciproquement, si a est adhérent à A, et si nous choisissons une métrique définissant la topologie de E alors la boule de centre a et de rayon $\frac{1}{n}$ contient au moins un point x_n appartenant à A. La suite $x_0, x_1, x_2, \dots x_n, \dots$ ainsi formée appartient bien à A et converge bien vers a.

Ainsi l'adhérence $\bar{A}$ de A est l'ensemble des limites des suites de A qui sont convergentes dans E.

Corollaire - Pour qu'une partie d'un espace topologique métrisable E soit fermée, il faut et il suffit qu'elle contienne toutes les limites de ses suites convergentes dans E.

* Si E est topologique non métrisable, la condition reste suffisante mais non nécessaire.

Théorème 16 - Pour qu'une application f d'un espace **métrisable** E dans un espace métrisable F soit continue en un point a de E, il faut et il suffit que l'image par f de toute suite de points de E convergeant vers a soit une suite de points de F convergeant vers $f(a)$ * .

1°/ La condition est nécessaire. Supposons en effet f continue en a , et soit $x_0, x_1, x_2, \dots x_n, \dots$ une suite convergeant vers a dans E . Alors, quel que soit $\mathcal{V}$ voisinage de $f(a)$ dans F , son image réciproque par f est un voisinage $\mathcal{U}$ de a dans E. Alors, pour tous les entiers n sauf au plus un nombre fini, x_n est dans $\mathcal{U}$, et par suite $f(x_n)$ est dans $\mathcal{V}$, ce qui prouve bien la convergence de la suite des $f(x_n)$ vers $f(b)$.

2°/ La condition est suffisante. Supposons donc cette condition réalisée, et supposons choisie une métrique définissant la topologie de E . Si l'application f n'était pas continue en a , alors on pourrait trouver un nombre $\varepsilon > 0$ tel que, quel que soit $\eta > 0$ il existe *un* point x tel que $d(x,a) \leq \eta$ et pourtant'tel que $d(f(x), f(a)) > \varepsilon$.

En particulier, ε étant ainsi choisi, pour tout entier n il existerait au moins un point x_n tel que $d(x_n, a) \leq \frac{1}{n}$ et $d(f(x_n), f(a)) > \varepsilon$.

On voit que la suite des x_n serait convergente vers a dans E , et que pourtant la suite des $f(x_n)$ ne serait pas convergente vers $f(a)$ dans F , ce qui contredirait les hypothèses.

§ 6 TOPOLOGIE PRODUIT

Soit E_1 et E_2 deux espaces métriques, appelons d_1 et d_2 leurs fonctions-distance respectives. Il est possible sur l'ensemble produit $E_1 \times E_2$ d'introduire la métrique δ définie comme suit :

(II,6;1) $$\delta((x_1, x_2), (y_1, y_2)) = \mathrm{Max}(d_1(x_1, y_1), d_2(x_2, y_2)).$$

* Si E et F sont topologiques non métrisables, la condition reste 'nécessaire mais non suffisante.

métrique que nous pourrons abréger symboliquement par $\delta = \mathrm{Max}(d_1, d_2)$. De la même manière on pourrait aussi introduire la métrique abrégée par $d_1 + d_2$, avec $(d_1+d_2)\big((x_1,x_2),(y_1,y_2)\big) = d_1(x_1, y_1) + d_2(x_2, y_2)$, ou la métrique abrégée par $\sqrt{d_1^2 + d_2^2}$, avec $\sqrt{d_1^2+d_2^2}\big((x_1,x_2),(y_1,y_2)\big) = \sqrt{(d_1(x_1,y_1))^2 + (d_2(x_2,y_2))^2}$.

Il est bien évident qu'aucune de ces métriques ne s'impose spécialement plus que les autres. Par ailleurs on voit que ces métriques sont équivalentes c'est à dire qu'elles définissent sur $E_1 \times E_2$, la même topologie. Cet exemple montre que précisément il n'est pas spécialement intéressant de mettre une métrique plutôt qu'une autre sur un espace produit. En revanche il est très facile de définir une topologie naturelle sur le produit de 2 espaces topologiques. Cette topologie se définit comme suit :

Considérons un ouvert A_1 de E_1, et un ouvert A_2 de E_2, on peut définir leur produit $A_1 \times A_2$, c'est-à-dire l'ensemble des couples (x,y) de $E_1 \times E_2$, tels que $x_1 \in A_1$, $x_2 \in A_2$ *.

On dira alors qu'une partie de $E_1 \times E_2$ est ouverte pour la topologie produit si, toutes les fois qu'elle contient un point, elle contient au moins un produit d'ouverts $A_1 \times A_2$ contenant ce point.

On vérifie facilement que les ensembles ainsi définis sur $E_1 \times E_2$ possèdent tous les axiomes que doivent vérifier les ensembles ouverts d'une topologie : naturellement les produits d'ouverts sont des ouverts, mais il y en a bien d'autres. La topologie ainsi définie sur $E_1 \times E_2$ s'appelle topologie produit des deux topologies données sur E_1 et E_2. On définira de même la topologie produit de plusieurs espaces, (Par exemple si $E_1 = E_2 = \dots = E_n = \mathbb{R}$, leur produit est l'espace topologique connu $\mathbb{R}^n$, défini par exemple par sa métrique

* $A_1 \times A_2$ s'appelle parfois un rectangle ouvert, par analogie avec les rectangles de $\mathbb{R}^2$, à côtés parallèles aux axes, et qui sont des produits d'intervalles de $\mathbb{R}$.

naturelle). De la définition même de cette topologie, il résulte que <u>les projections canoniques de $E_1 \times E_2$ sur E_1 et E_2 sont continues</u>. Si en effet f est la projection de $E_1 \times E_2$ sur E_1, et si A_1 est un ouvert de E_1, son image réciproque est $f^{-1}(A_1) = A_1 \times E_2$, qui est bien un ouvert et même un produit d'ouverts de $E_1 \times E_2$.

Suites convergentes dans un produit

<u>Théoréme 17</u> - <u>Pour qu'une suite $(x_0, y_0), (x_1, y_1), \dots (x_n, y_n), \dots$ * d'éléments de $E \times F$, converge vers un point (a, b) pour la topologie produit, il faut et il suffit que la suite des x_n converge vers a dans E, et que la suite des y_n converge vers b dans F.</u>

Supposons d'abord la suite des (x_n, y_n) convergente vers (a, b). Comme les projections $(x,y) \to x$ et $(x,y) \to y$ sont continues, on déduit du théorème 16 que la suite des x_n converge vers a, et que la suite des y_n converge vers b.

Réciproquement supposons que la suite des x_n converge vers a dans E, la suite des y_n vers b dans F. Soit $\mathcal{V}$ un voisinage de (a, b) dans $E \times F$; il contient un ouvert contenant (a, b) donc, d'après la définition des ouverts de $E \times F$, un produit d'ouverts $A \times B$, $a \in A$, $b \in B$. Alors il existe, d'après l'hypothèse de convergence des suites x_n et y_n, des entiers n_1 et n_2 tels que $n \geq n_1$ entraîne $x_n \in A$, et que $n \geq n_2$ entraîne $y_n \in B$; alors $n \geq \mathrm{Max}(n_1, n_2)$ entraîne $(x_n, y_n) \in A \times B \in \mathcal{V}$, et la suite des (x_n, y_n) converge bien vers (a, b).

Plus généralement soit G un espace topologique. Une application h, de E dans $F \times G$ est donnée par un couple de deux applications f et g de E dans F et dans G respectivement, telles que l'on ait $h(x) = (f(x), g(x))$ (chap. 1, page 7). Pour que cette application soit continue, il faut et il suffit que les applications f et g, chacune séparément, soient continues. Démonstration évidente.

Fonctions continues de plusieurs variables

On a étudié en Mathématiques Spéciales la notion de fonction continue de deux variables.

* Nous changeons ici de notation. Nous appelons E et F les espaces topologiques, pour laisser la numérotation 1,2... aux éléments de la suite.

Soit f une fonction de deux varfables : $(x,y) \longrightarrow f(x,y)$. C'est simplement une application d un espace produit $E \times F$ dans G . Supposons que ces trois espaces soient topologiques. Quand dit-on que la fonction est continue en un point (a,b) ?

Dans le cas d'espaces métriques,on le dit si, quel que soit $\varepsilon > 0$, il existe $\eta > 0$ tel que $d(x,a) \leq \eta$, $d(y,b) \leq \eta$, entraîne $d(f(x,y), f(a,b)) \leq \varepsilon$ *

Mais on voit que cela ne signifie pas autre chose que la continuité de l'application f de E x F dans G , au point (a,b), lorsque l'on munit E x F de la topologie produit. C'est donc cela que nous appellerons, en général, continuité d'une fonction de deux variables. Ainsi la donnée d'une application continue d'un produit $E_1 \times E_2 \cdots \times E_n$ dans un produit $F_1 \times F_2 \cdots \times F_m$ est équivalente à la donnée d'un système de m fonctions continues de n variables.

Théorème 17 bis - Si E est un espace métrique, la fonction distance d, application de E x E dans la droite réelle $\mathbb{R}$, est continue.

En effet, quelque soit $\varepsilon > 0$, $d(x,a) \leq \frac{\varepsilon}{2}$ et $d(y,b) \leq \frac{\varepsilon}{2}$ entraîne $|d(x,y) - d(a,b)| \leq d(x,a) + d(y,b) \leq \varepsilon$,ce qui démontre la continuité.

Groupes topologiques, espaces vectoriels topologiques

On appelle groupe topologique un ensemble G qui, d'une part, est muni d'une structure de groupe, et, d'autre part, d'une topologie telle que les applications fondamentales définies par la structure de groupe, c'est-à-dire l'application $(x,y) \longrightarrow xy$ de Gx G dans G , et l'application $x \longrightarrow x^{-1}$ de G dans G , soient continues.

On appelle espace vectoriel topologique un ensemble, E , muni d'une part d'une structure d'espace vectoriel sur le corps des réels - ou des complexes,+ d'autre part d'une topologie telle que l'addition $(\vec{x},\vec{y}) \longrightarrow \vec{x}+\vec{y}$ soit continue de E x E dans E, ** et que la multiplication par les

* Nous avons représenté, pour simplifier, par le même symbole d , les distances dans E, F, G .

** La continuité de l'addition exprime que la limite d'une somme de 2 vecteurs est la somme des limites de ces 2 vecteurs.

scalaires, $(\lambda, \vec{x}) \longrightarrow \lambda\vec{x}$, soit continue de $K \times E$ dans E, où K est le corps des scalaires muni de sa topologie naturelle. On voit alors qu'<u>un espace vectoriel normé est bien un espace vectoriel topologique</u>. On a en effet :

$$(\mathrm{II}, 6; 2) \begin{cases} \|(\vec{x}+\vec{y}) - (\vec{a}+\vec{b})\| \leq \|\vec{x}-\vec{a}\| + \|\vec{y}-\vec{b}\| \\ \|\lambda\vec{x} - \alpha\vec{a}\| \leq |\lambda - \alpha| \|\vec{x}\| + |\alpha| \|\vec{x}-\vec{a}\| . \end{cases}$$

Alors, pour $\|\vec{x}-\vec{a}\| \leq \frac{\varepsilon}{2}$, $\|\vec{y}-\vec{b}\| \leq \frac{\varepsilon}{2}$, on aura bien $\|(\vec{x}+\vec{y}) - (\vec{a}+\vec{b})\| \leq \varepsilon$, ce qui exprime la continuité de l'addition. D'autre part, si une suite $(\lambda_n, \vec{x}_n)$ tend vers $(\alpha, \vec{a})$ dans $K \times E$, c'est-à-dire si la suite $\vec{x}_n$ tend vers $\vec{a}$, et si la suite λ_n tend vers α (théorème 17), alors $\|\vec{x}_n - \vec{a}\|$ tend vers 0 donc $|\alpha| \|\vec{x}_n - \vec{a}\|$ tend vers 0, d'autre part $|\lambda_n - \alpha|$ tend vers 0 et $\|\vec{x}_n\|$ reste borné, donc $|\lambda_n - \alpha| \|x_n\|$ tend vers 0 ; alors la 2ème Inégalité montre que $\lambda_n \vec{x}_n$ tend vers $\alpha\vec{a}$; d'après le théorème 16, cela prouve la continuité de l'application $(\lambda, \vec{x}) \longrightarrow \lambda\vec{x}$ de $K \times E$ dans E.

Le théorème 13 se généralise alors comme suit : <u>sur un espace vectoriel de dimension finie, il n'existe qu'une seule topologie d'espace vectoriel topologique</u>.

Continuité partielle d'une fonction de deux variables

Soit f une application de $E \times F$ dans G. Si (a, b) est un point de $E \times F$, et si nous fixons $x = a$, nous voyons qu'il existe une application, définie par f et par a, de F dans G, à savoir l'application $y \longrightarrow f(a, y)$. On note souvent f_a cette application, de sorte que $f_a(y) = f(a, y)$. On la note aussi $f(a, \ )$, en omettant la variable y.

Il se peut que cette application soit continue au point $y = b$. On dit alors que l'application f est <u>partiellement ou séparément</u> continue au point b par rapport à y, pour x fixé en a. Même définition pour l'application partielle $f_b : x \longrightarrow f(x, b)$, notée aussi $f(\cdot, b)$, et pour la notion d'application partiellement ou séparément continue au point a, par rapport à x, pour y fixé en b.

Cela revient tout simplement à dire que la restriction de l'application f au sous-espace $\{a\} \times F$ ou au sous-espace $E \times \{b\}$ est continue au point (a, b). On dit que f est séparément continue au point (a,b) si elle possède les 2 continuités séparées précédentes; on dit qu elle est séparément continue sur $E \times F$ si elle l'est en tout point (a,b) de $E \times F$.

Une fonction continue est bien évidemment partiellement continue; mais la réciproque est fausse. Par exemple, considérons la fonction f réelle de 2 variables réelles, définie par la formule suivante :

$$(\mathrm{II},6;3) \quad f(x,y) = \frac{xy}{x^2+y^2} \quad \text{pour} \quad (x,y) \neq (0,0) \quad ; \quad f(0,0) = 0 .$$

Nous voyons que cette fonction est partout continue sur le complémentaire de l'origine dans $\mathbb{R}^2$; elle est continue sur l'axe $x'x$ et sur l'axe $y'y$ sur lesquels elle est $\equiv 0$, donc elle est séparément continue à l'origine et par suite dans tout le plan. Mais elle n'est manifestement pas continue par rapport à l'ensemble des deux variables à l'origine, car, sur la droite $y = mx$, elle prend la valeur $\frac{m}{1+m^2}$ en dehors de l'origine; or $\frac{m}{1+m^2} \neq 0$, et par conséquent elle ne tend pas vers 0 quand (x,y) tend vers l'origine.

§ 7 ESPACES COMPACTS. PROPRIÉTÉS ÉLÉMENTAIRES

Soit E un espace topologique, on appelle recouvrement de E un ensemble de parties de E, tel que tout point de E appartienne au moins à l'une d'entre elles. Un sous-recouvrement d'un recouvrement est un recouvrement formé de parties appartenant au premier recouvrement. Un recouvrement est fini s'il est formé seulement d'un nombre fini de parties de E. Un recouvrement de E est dit ouvert si toutes les parties appartenant à ce recouvrement sont des ouverts de E. Ainsi l'ensemble des intervalles $]n-1, n+1[$, lorsque n parcourt $\mathbb{Z}$ est un recouvrement ouvert de $\mathbb{R}$; il n'a pas d'autre sous-recouvrement que lui-même, car, si on supprime l'intervalle $]n-1, n+1[$, le point n n'est plus recouvert * .

Définition - Propriété de HEINE-BOREL-LEBESGUE

Un espace topologique E est dit compact si tout recouvrement ouvert de E admet au moins un sous-reaouvrement fini. Cela signifie que si l'on considère n'importe quel recouvrement de E par un ensemble de parties ouvertes, il existe déjà un ensemble fini de ces parties qui suffit à recouvrir E .

Z

* L'ensemble des intervalles $]n-\frac{3}{4}, n+\frac{3}{4}[$ est aussi un recouvrement ouvert de $\mathbb{R}$; il est formé d'ouverts plus petits que le précédent, mais n'en est pas un sous-recouvrement, car il est un ensemble de parties entièrement différentes des premières

On notera que jusqu'à présent nous n'avions étudie les Propriétés topologiques que de certaines parties de E relativement à E lui-même, par exemple la propriété pour une partie A de E d'être ouverte, fermée, etc.... Ici au contraire, le fait d'être compact est une propriété de l'espace topologique lui-même. Si cependant E est un espace topologique et A une partie de E , on dira que A est compacte, si, en tant qu'espace topologique muni de la topologie induite, A est un espace compact. Cela n'impose nullement a E d'être compact ! On appelle partie relativement compacte de E toute partie dont l'adhérence est compacte.

On voit alors immédiatement que la réunion d'un nombre fini de parties compactes de E est encore compacte.

Soient en effet $A_1, A_2, \ldots A_n$ des parties compactes de E . Soit $\mathcal{R}$ un recouvrement ouvert de $A = A_1 \cup A_2 \ldots \cup A_n$. Alors les ouverts de $\mathcal{R}$ coupent A_1 suivant des ouverts de A_1 , et $\mathcal{R}$ définit donc un recouvrement ouvert de A_1 ; comme A_1 est compact, un nombre fini des ouverts de $\mathcal{R}$ suffit à recouvrir A_1. Comme il en est de même pour $A_2, \ldots A_n$, on peut bien, en prenant à la fois ces n systèmes finis d'ouverts de $\mathcal{R}$, recouvrir A avec un nombre fini d'ouverts de $\mathcal{R}$, et A est bien compacte. De même la réunion d'un nombre fini de parties relativement compactes est relativement compacte.

Exemples - Un espace ne contenant qu'un nombre fini de points est compact. La droite réelle $\mathbb{R}$, l'espace $\mathbb{R}^n$, un espace vectoriel normé de dimension finie ou infinie ne sont jamais compacts. En effet si nous considérons l'ensemble des boules ouvertes, de centre origine et de rayon > 0 , elles forment évidemment un recouvrement de l'espace. Or un nombre fini quelconque de ces boules, est contenu dans une même boule, de rayon fini, et par suite ne recouvre pas l'espace. Plus généralement, une partie non bornée d'un espace métrique, c'est-à-dire non contenue dans au moins une boule (de rayon fini), n'est sûrement jamais compacte.

Théorème 18 - Un intervalle fermé, borné $[a, b]$ de la droite réelle $\mathbb{R}$ est un espace compact.

Démonstration - Soit $\mathcal{R}$ un recouvrement ouvert de $[a, b]$. Soit c le milieu de $[a, b]$.

S'il n'était pas possible de trouver un nombre fini de parties, appartenant à $\mathcal{R}$ et recouvrant l'intervalle entier $[a, b]$, alors cela ne serait pas possible non plus pour au moins l'un des deux sous-intervalles $[a, c], [c, b]$, par exemple $[a, c]$.

Appelons $[a_1, b_1]$ ce sous-intervalle. Nous le partagerons aussi en deux et trouverons un sous-intervalle $[a_2, b_2]$ deux fois plus petit et possédant la même propriété. Nous formerions ainsi une suite Infinie $[a_0, b_0] = [a, b]$, $[a_1, b_1], \dots$ $[a_n, b_n], \dots$ de sous-intervalles de $[a, b]$ avec la même propriété : aucun d'eux ne pourrait être recouvert par un nombre fini de parties appartenant à $\mathcal{R}$.

La suite croissante majorée des a_n admettrait une limite α, la suite décroissante minorée des b_n admettrait une limite β, et comme la longueur de $[a_n, b_n]$ est $\frac{b-a}{2^n}$, on aurait nécessairement $\alpha = \beta$. D'après la définition de la limite d'une suite, tout intervalle ouvert contenant le point $\alpha = \beta$ contient, pour n assez grand, a_n et b_n donc tout l'intervalle $[a_n, b_n]$.

Or il existe nécessairement un des ouverts du recouvrement $\mathcal{R}$, soit $\mathcal{O}$, qui contient le point $\alpha = \beta$; comme $\mathcal{O}$ est ouvert, il existe un intervalle ouvert $]a', b'[$ contenu dans $\mathcal{O}$ et contenant ce point; or, pour n suffisamment grand, l'intervalle $[a_n, b_n]$ serait contenu dans $]a', b'[$ donc dans $\mathcal{O}$, et nous aboutirions ainsi à une contradiction puisque $[a_n, b_n]$ ne devrait pas pouvoir être recouvert par un nombre fini de parties appartenant à $\mathcal{R}$, alors qu'il est recouvert par une seule d'entre elles, à savoir $\mathcal{O}$. Cette contradiction prouve bien que $[a, b]$ est compact.

Plus généralement, <u>dans l'espace $\mathbb{R}^m$,</u> un <u>pavé fermé borné,</u> c'est-à-dire l'ensemble des points $(x_1, x_2, \dots x_m)$ définis par un système d'inégalités larges

$$a_1 \leqslant x_1 \leqslant b_1, \quad a_2 \leqslant x_2 \leqslant b_2, \dots \quad a_m \leqslant x_m \leqslant b_m,$$

$a_1, a_2, \dots a_m$, $b_1, b_2, \dots b_m$ finis,

<u>est un espace compact.</u> On le montrera par la même **méthode** de subdivision, avec cette différence qu'au **lieu de** diviser en deux, il faut à chaque opération diviser en 2^m pavés en divisant en 2 pour chacune des m coordonnées.

* <u>Théorème, 19</u> : <u>La droite achevée $\overline{\mathbb{R}}$, munie de sa topologie naturelle, est un espace compact.</u>

<u>Démonstration</u> - Soit $\mathcal{R}$ un recouvrement ouvert, de $\overline{\mathbb{R}}$. Le point $-\infty$ appartient à l'un au moins des ouverts de $\mathcal{R}$ soit $\mathcal{O}_-$, et de même $+\infty$ appartient à l'un de ces ouverts, $\mathcal{O}_+$. Il en résulte en particulier que le complémentaire de

la réunion de ces deux ouverts est contenu dans un intervalle fermé borné de $\mathbb{R}$, suffisamment grand, $[A, B]$. Comme il existe nécessairement alors un nombre fini des ouverts du recouvrement $\mathcal{R}$ qui suffit à recouvrir $[A, B]$ (théorème 19), on en déduit bien que ce nombre fini d'ouverts de $\mathcal{R}$, augmenté de $\mathcal{O}_-$ et $\mathcal{O}_+$, recouvre $\overline{\mathbb{R}}$, et que par conséquent $\overline{\mathbb{R}}$ est compact.

Remaroues 1°) Comme conséquence particulière de la définition d'un compact, on a la propriété suivante :

Si E est un espace compact et si $\mathcal{O}_1, \mathcal{O}_2, \ldots \mathcal{O}_n, \ldots$ est une suite croissante d'ensembles ouverts dont la réunion est E alors, déjà pour un entiern convenable, l'ouvert $\mathcal{O}_n$ est identique à E .

2°) Si deux espaces sont homéomorphes, et si l'un est compact, il en est de même de l'autre. Les exemples précédents nous prouvent donc que $\mathbb{R}$ et $\overline{\mathbb{R}}$ ne sont pas homéomorphes (il n'existe pas d'homéomorphisme de l'une sur l'autre). Un intervalle $[a, b]$ fermé borné de $\mathbb{R}$ n'est pas homéomorphe à $\mathbb{R}$ (alors qu'un intervalle ouvert est homéomorphe à $\mathbb{R}$: $x \longrightarrow \operatorname{tg} x$ est un homéomorphisme de $\left]-\frac{\pi}{2}, +\frac{\pi}{2}\right[$ sur $\mathbb{R}$). Les critères, permettant de prouver que deux espaces ne sont pas homéomorphes, sont précieux; car il est généralement facile de prouver que deux espaces sont homéomorphe-s, en exhibant l'homéomorphisme; il est toujours plus délicat de prouver que deux espaces ne sont pas homéomorphes.

Théorème 20 - Pour qu'un espace topologique E soit compact, il faut et il suffit que, pour tout ensemble de parties fermées de E dont l'intersection est vide, il existe déjà un nombre fini de ces parties dont l'intersection soit vide.

Démonstration. On passe de la définition à ce théorème en remplaçant les parties par leur complémentaire, ce qui passe de réunion à Intersection et de partie ouverte à partie fermée.

Corollaire 1 - Si E est un espace compact, et si $F_0, F_1, F_2, \ldots F_n, \ldots$ est une suite décroissante d'ensembles fermés dont l'intersection est vide, alors il existe déjà un entier n convenable, tel que F_n soit vide.

On remarquera bien que cette propriété n'est sûrement pas vraie pour la droite $\mathbb{R}$, ce qui confirme le fait déjà vu que $\mathbb{R}$ n'est pas compacte; si en effet nous considérons la suite décroissante d'intervalles fermés $[n, +\infty[$, leur intersection est vide alors qu'aucun d'eux n'est vide.

Remarque - Si l'on applique le théorème 11 du chapitre 1, on voit que le précédent corollaire est équivalent au suivant.

Corollaire 2 - Si E est compact, si $F_0, F_1, \dots F_n, \dots$ est une suite décroissante d'ensembles fermés, si aucun d'eux n'est vide, alors leur intersection n'est pas vide.

Théorème 21 - Soit E un espace topologique, F une partie compacte de E, alors nécessairement F est une partie fermee de E * .

Démonstration - Nous nous bornerons, pour simplifier, à la donner lorsque E est un espace métrique. Soit a un point adhérent à F, nous devons montrer que $a \in F$. Si nous appelons F_n l'intersection de F et de la boule fermée de centre a et de rayon $\frac{1}{n}$, n entier ≥ 1 nous voyons que les F_n forment une suite décroissante de parties fermées de F (théorème 5). Aucun des F_n n'est vide, puisque a-est adhérent à F (théorème 4) donc, F étant suppose compact, leur intersection n'est pas vide. Mais l'intersection des boules considérées se réduit à a. Alors a est le seul point qui puisse être l'intersection de tous les F_n, cela prouve que a appartient à tous les F_n, c'est-a-dire à F. Par conséquent F est fermée dans E.

Il est bien évident que la réciproque de ce théorème n'est pas exacte. Une partie fermée quelconque d'un espace topologique n'est pas nécessairement compacte, sans quoi l'espace lui-même, toujours fermé, serait toujnurs compact. Mais on a la réciproque suivante :

Théorème 22 - Toute partie fermée d'un espace compact est un espace compact.

Démonstration - Soit E un espace compact, F une partie fermée de E. Soit $(F_i)_{i \in I}$ une famille quelconque de parties de F, fermées dans F, et dont l'intersection soit vide. Comme les F_i sont fermées dans F et que F 'est supposée fermée dans E, les F_i sont fermées dans E (théorème 6,b)); comme alors E est supposé compact, il existe un nombre fini des F_i dont l'intersection est vide, et cela prouve, d'après le théorème 20, que F est compact.

Les théorèmes 21 et 22 prouvent que, dans un espace compact, les parties compactes sont identiques aux parties fermées. Ainsi pour qu'une partie de la droite achevée $\overline{\mathbb{R}}$ soit compacte, il faut et il suffit qu'elle soit fermée dans $\overline{\mathbb{R}}$.

Z

* - Attention ! Ce théorème compare une propriété intrinsèque de F, le fait d'être compacte (pour la topologie Induite), à une propriété de F relativement à E, le fait d'être fermée dans E.

Théorème 22 bis. Tout espace compact est régulier : tout point a un système fondamental de voisinages compacts.

Démonstration. soit $\mathcal{V}$ un voisinage de a, $\overset{\circ}{\mathcal{V}}$ son intérieur. Alors $K = \complement \overset{\circ}{\mathcal{V}}$ est fermé, donc compact d'après le théorème 22. Soit $\mathcal{W}$ un voisinage fermé, donc compact, de a. Alors $\mathcal{W} \cap K$ est un ensemble fermé de K. Lorsque $\mathcal{W}$ varie, ces ensembles fermés de K ont une intersection vide, puisque l'intersection des $\mathcal{W}$ est réduite à a (page 58) qui n'est pas dans K. K étant compact, il existe alors un nombre fini d'ensembles $\mathcal{W}_1, \mathcal{W}_2, \dots, \mathcal{W}_n$ tels que l'intersection des $\mathcal{W}_i \cap K$ soit vide. $\mathcal{W}_0 = \mathcal{W}_1 \cap \dots \cap \mathcal{W}_n$ est un voisinage compact de a et $\mathcal{W}_0 \cap K$ est vide, donc $\mathcal{W}_0 \subset \mathcal{V}$; ainsi tout voisinage $\mathcal{V}$ de a contient un voisinage compact $\mathcal{W}_0$, c.q.f.d

à revoir

Plus généralement, on démontre de la même manière que, dans un espace compact, tout compact a un système fondamental de voisinages compacts.

Théorèmes 2-j - Pour qu'une partie d'un espace vectoriel E normé de dimension finie soit compacte, il faut et il suffit qu'elle soit fermée et bornee.

Démonstration 1°/ La condition est nécessaire, Une partie compacte est nécessairement bornée comme nous l'avons vu au début, elle est nécessairement fermée d'après le théorème 21.

2°/ La condition est suffisante. Même si le corps des scalaires est $\mathbb{C}$, nous puuvons considérer E, espace vectoriel de dimension n sur $\mathbb{C}$, comme un espace vectoriel de dimension $2n$ sur $\mathbb{R}$. Nous supposerons donc que le corps des scalaires est $\mathbb{R}$.

a / Supposons d'abord que l'espace E soit $\mathbb{R}^n$, et que la norme soit la fonction $(x_1, x_2, \dots x_n) \longrightarrow \underset{i=1,2,\dots n}{\mathrm{Max}} |x_i|$.

Une partie bornée est alors contenue dans une boule, c'est-à-dire dans un pavé fermé borné, donc dans un compact; comme la partie considérée est supposée fermée, elle est une partie fermée d'un compact;il suffit alors d'appliquer le théorème 22. (Plus généralement dans tout espace métrique où toute boule fermée est compacte, toute partie fermée bornée est compacte).

b / Soit maintenant E un espace vectoriel normé quelconque de dimension finie n sur $\mathbb{R}$.

Nous allons d'abord montrer ce qui avait été admis au théorème 13. Choisissons une base $\vec{e}_1, \vec{e}_2, \dots \vec{e}_n$ de E, de sorte que chaque point $\vec{x}$ peut-être représenté par ses coordonnées $x_1, x_2, \dots, x_n$. Appelons $\|\vec{x}\|$ la norme donnée sur E, et posons $|\vec{x}| = \underset{i=1,2,\dots n}{\mathrm{Max}} |x_i|$. On a d'abord $\|\vec{x}\| = \|\sum x_i \vec{e}_i\| \leq \sum |x_i| \, \|\vec{e}_i\|$

$\leq \sum_i \|e_i\| \underset{i=1,2,\dots n}{\mathrm{Max}} |x_i|$; donc, si l'on pose $k = \sum_i \|\vec{e}_i\|$, on a la majoration

$$\|\vec{x}\| \leq k\,|\vec{x}| \qquad (\mathrm{II},7;1)$$

Il nous faut trouver une inégalité en sens Inverse. Mais (II,7;1) nous montre que l'application identique de E, muni de la norme $|\vec{x}|$, dans E, muni de la norme $\|x\|$, est continue (car $\|\vec{x}-\vec{a}\| \leqslant k|\vec{x}-\vec{a}|$, donc, $\varepsilon > 0$ étant donné, si l'on choisit $\eta = \frac{\varepsilon}{k}$., on voit que $|\vec{x}-\vec{a}| \leqslant \eta$ entraîne $\|\vec{x}-\vec{a}\| \leqslant \varepsilon$, ce qui est la définition de la continuité au point $\vec{a}$); alors l'image réciproque d'un fermé est fermée (théorème 8), autrement dit toute partie F de E , fermée dans la métrique $\| \; \|$, est fermée dans la métrique $| \; |$. En particulier la boule $\|\vec{x}\| \leqslant R$, fermée dans la métrique $\| \; \|$, l'est aussi dans la métrique $| \; |$. Appelons Q la sphère (cube creux) $|x| = 1$. C'est une partie fermée bornée dans **E** pour la métrique $| \; |$ donc un compact d'après ce que nous avons vu au début de'la démonstration. Soit F_R 1 intersection de Q avec la boule $\|\vec{x}\| \leqslant R$, fermée pour la métrique $| \; |$: c'est un fermé du compact Q . L'intersection de tous ces fermes est vide, car l'intersection des boules $\|\vec{x}\| \leqslant R$ est l'origine, qui n'appartient pas à Q Donc il y en a un nombre fini dont l'intersection est vide (théorème 20), autrement dit il existe un nombre $\rho > 0$ tel que la boule $\|\vec{x}\| \leqslant \rho$ ne rencontre pas Q . Alors $\|\vec{x}\| \leqslant \rho$ entraîne $|\vec{x}| < 1$; [sans quoi, il existerait un point $\vec{x}_0$ vérifiant $\|\vec{x}_0\| \leqslant \rho$, $|\vec{x}_0| \geqslant 1$; alors, pour $\lambda = \frac{1}{|\vec{x}_0|} \leqslant 1$, on aurait $\|\lambda\vec{x}_0\| \leqslant \lambda\rho \leqslant \rho$ et $|\lambda\vec{x}_0| = 1$, ce qui est contraire à ce que nous venons de voir] . Par homothétie de rapport μ , on en déduit que $\|\vec{x}\| \leqslant \mu\rho$ entraîne $|\vec{x}| < \mu$ donc $\leqslant \mu$; la première inégalité est vérifiée pour $\mu = \frac{\|\vec{x}\|}{\rho}$, donc la deuxième donne

$$|\vec{x}| \leqslant \frac{1}{\rho}\|\vec{x}\| , \tag{II,7;2}$$

ce qui, avec (II,7;1) prouve l'équivalence des normes $| \; |$ et $\| \; \|$, c'est-à-dire le théorème 13.

Alors le théorème 23 est démontré dans tous les cas, car les notions "fermée, bornée, compacte" sont les mêmes pour deux normes équivalentes (voir note (*) page **57**), et ce que nous avons montré dans a/, pour une norme particulière de E définie par une base, l'est aussi pour la norme donnée, qui lui est équivalente d'après le théorème **13.**

Espaces localement compacts

On dit qu'un espace topologique E est <u>localement compact</u> si tout point possède au moins un voisinage compact.

Ainsi <u>tout espace vectoriel normé de dimension finie est localement compact</u>, puisque toute boule fermée est compacte.

On démontre qu'un espace vectoriel normé de dimension infinie n'est jamais localement compact (théorème 45 bis); une boule fermée n'est pas compacte ; pour qu'une partie de l'espace soit compacte, il est alors nécessaire, mais non suffisant, qu'elle soit fermée et bornée.

Si deux espaces topologiques sont homéomorphes, et si l'un est localement compact, il en est de même de l'autre. Ceci nous donne, comme la compacité (page 70), un nouveau critère permettant de dire que deux espaces ne sont pas homéomorphes. Par exemple un espace vectoriel normé de dimension finie et un autre de dimension infinie ne sont jamais homéomorphes.

Tout espace localement compact est régulier, et tout point a un système fondamental de voisinages compacts. Si en effet $\mathcal{V}$ est un voisinage compact de a, a possède dans $\mathcal{V}$ un système fondamental de voisinages compacts d'après le théorème 22 bis; mais tout voisinage de a dans $\mathcal{V}$ est aussi un voisinage de a dans l'espace entier (théorème 6, c). Plus généralement, tout compact d'un espace localement compact a un système fondamental de voisinages compacts.

Point d'accumulation d'une suite*

soit $x_0, x_1, x_2, \ldots x_n, \ldots$ une suite d'éléments d'un espace topologique E, on dit que a est point d'accumulation de la suite si, pour tout voisinage $\mathcal{V}$ de a, il existe une infinité de valeurs de l'entier n telles que $x_n \in \mathcal{V}$. Si une suite converge vers a, elle admet a comme point d'accumulation.

Théorème 24 - Si E est un espace métrisable, a est point d'accumulation d'une suite $x_0, x_1, x_2, \ldots x_n, \ldots$ d'éléments de E, si et seulement si on peut extraire de cette suite une suite Partielle convergeant vers a.

Démonstration - On appelle suite partielle de la suite $x_0, x_1, x_2, \ldots x_n, \ldots$ une suite du type $x_{p_0}, x_{p_1}, \ldots x_{p_n}, \ldots$ où $n \to p_n$ est une application strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ **. Il est alors évident, en appliquant simplement la définition, que s'il existe une suite partielle de la suite initiale qui converge vers a, cette suite initiale admet a comme point d'accumulation (et ceci même si E est topologique non métrisable) Montrons la réciproque, en supposant choisie une métrique définissant la topologie. Supposons a point d'accumulation de la suite initiale. Puisque a est point d'accumulation, pour chaque boule $B(a, \frac{1}{n})$, il existe une infinité de valeurs de p telles que $x_p \in B(a, \frac{1}{n})$. Prenons d'abord un entier p_1 tel que $x_{p_1} \in B(a,1)$, prenons ensuite un entier $p_2 > p_1$ tel que $x_{p_2} \in B(a, \frac{1}{2})$,

* On dit souvent "point adhérent à une suite", mais cela introduit des confusions possibles avec la notion de point adhérent à un ensemble, c est pourquoi nous employons un mot différent.

** En prenant $p_n = n$, on voit que la suite elle-même est une suite partielle.

puis un entier $p_3 > p_2$ tel que $x_{p_3} \in B\left(a, \frac{1}{3}\right)$, et ainsi de suite; de proche en proche, nous formons ainsi une suite partielle $n \longrightarrow x_{p_n}$ de la suite initiale, et qui, bien évidemment, converge vers a .

; on dit aussi; théorème de..

Théorème 25 - (Propriété de Weierstrass Bolzano). Si E est un espace métrisable, pour qu'il soit compact, il faut et il suffit que toute suite d éléments de E admette au moins un point d'accumulation *

Démonstration 1°/ -Supposons **E** compact et soit $x_0, x_1, \dots x_n, \dots$ une suite d'éléments de **E** . Appelons **A,** l'ensemble $\{x_n, x_{n+1}, \dots\}$, et $\bar{A}_n$ son adhérence. Alors les A, forment une suite décroissante d'ensembles fermés et aucun d'eux n'est vide, donc leur intersection n'est pas vide. Soit a un point de cette intersection. Dire que a appartient à $\bar{A}_n$ ou est adhérent à A, , c'est dire que tout voisinage de a contient au moins un point de **A,** , et comme c'est vrai pour tout n , cela prouve bien que a est un point d'accumulation de la suite.

2°/ La réciproque est délicate. Nous passerons par l'intermédiaire de deux lemmes.

Lemme 1 - Soit E un espace métrique, dans lequel toute suite admet au moins un point d'accumulation. Soit& un recouvrement ouvert de E. Alors il existe un nombre $\varepsilon > 0$ tel que toute boule, de centre quelconque et de rayon $\leqslant \varepsilon$, soit contenue toute entière dans au moins l'un des ouverts du recouvrement.

Supposons en effet qu'il n'en soit pas ainsi. Alors, pour tout entier n , il serait possible de trouver un point a, de E tel que la boule de centre a, et de rayon $\frac{1}{n}$ ne soit pas contenue toute entière dans l'un au moins des ouverts du recouvrement. Nous formons ainsi une suite infinie $a_1, a_2, \dots$ a,, ... d'éléments de E . Cette suite admet au moins un point d'accumulation a . Comme $\mathcal{R}$ est un recouvrement, il existe un des ouverts de $\mathcal{R}$, soit $\mathcal{O}$, qui contient a , et cet ouvert lui-même contient une boule de centre a et de rayon α. Mais il existe une infinité de valeurs de n , donc au moins une , telles que l'on ait à la fois $\frac{1}{n} \leqslant \frac{\alpha}{2}$, et $d(a_n, a) \leqslant \frac{\alpha}{2}$.

On voit alors que la boule de centre a_n et de rayon $\frac{1}{n} \leqslant \frac{\alpha}{2}$

* Si E est topologique non métrisable, la condition reste nécessaire, mais non suffisante.

est toute entière contenue dans la boule $B(a,\alpha)$, et par conséquent dans l'ouvert $\mathcal{O}$ du recouvrement, ce qui est contraire à l'hypothèse faite sur la suite des a_n. Nous aboutissons ainsi à une contradiction.

Lemme 2 - Soit E un espace métrique, dans lequel toute suite admet au moins un point d'accumulation. Alors, quel que soit $\varepsilon > 0$, on peut recouvrir E tout entier à l'aide d'un nombre fini de boules de rayon ε.

En effet soit un point a_0 de E. Si $B_0(a_0,\varepsilon) = E$, le lemme est démontré. S'il n'en est pas ainsi, il existe au moins un point a_1 qui n'appartienne pas à $B_0(a_0,\varepsilon)$.

Si alors $B_0(a_0,\varepsilon) \cup B_0(a_1,\varepsilon) = E$, alors le lemme est démontré, et ainsi de suite. Nous pouvons former de cette manière une suite $B_0(a_0,\varepsilon), B_0(a_1,\varepsilon), \dots B_0(a_n,\varepsilon), \dots$ de boules de rayon ε.

Si nous ne sommes jamais arrêtés, cela prouve que nous pouvons former une suite infinie de points $a_0, a_1, a_2, \dots a_n, \dots$ dont les distances mutuelles sont $\geqslant \varepsilon$. Or il est facile de voir que cette circonstance est impossible, car cette suite infinie posséderait au moins un point d'accumulation a, et par suite il existerait une infinité de valeurs de n, dont au moins deux valeurs distinctes p et q, telles que $d(a,a_p) \leqslant \frac{\varepsilon}{3}$, $d(a,a_q) \leqslant \frac{\varepsilon}{3}$. On en déduirait $d(a_p,a_q) \leqslant \frac{2\varepsilon}{3}$, ce qui serait contradictoire avec l'hypothèse $d(a_p,a_q) \geqslant \varepsilon$. Il en résulte bien que nous sommes arrêtés, dans notre construction, à un certain entier n, et qu'alors on peut recouvrir E avec $n+1$ boules ouvertes de rayon ε.

Moyennant ces deux lemmes la démonstration du théorème est évidente. Choisissons une métrique définissant la topologie de E. Pour prouver que E est compact, nous devons considérer un recouvrement ouvert quelconque $\mathcal{R}$. D'après le lemme 1, il existe un nombre $\varepsilon > 0$ tel que toute boule de rayon $\leqslant \varepsilon$ soit contenue toute entière dans au moins l'un des ouverts du recouvrement $\mathcal{R}$. D'après le lemme 2, on peut recouvrir E tout entier par un nombre fini de boules $B_0, B_1, B_2, \dots, B_n$, de rayon ε. Comme chacune d'elle B_i est contenue toute entière dans un ouvert $\mathcal{O}_i$ du recouvrement $\mathcal{R}$ on obtient un nombre fini d'ouverts $\mathcal{O}_0, \mathcal{O}_1, \mathcal{O}_2, \dots \mathcal{O}_n$, de $\mathcal{R}$, qui suffit à recouvrir E.

Remarque 1. Il en résulte que les lemmes 1 et 2 sont des propriétés des espaces métriques compacts *.

* Le mot "compact" signifie (Dictionnaire Larousse) : serré, pressé. C'est bien de cela qu'il s'agit. Le lemme 2 indique que, même si des boules ont un petit rayon, un nombre fini d'entre elles suffit à recouvrir E, qui est donc très serré.

Remaraue 2 - Le **théorème** de WBIERSTRASS-BOLZANO n'est **manifestement** pas vrai pour la **droite** $\mathbb{R}$: la suite des entiers $\geqslant 0$ n'a aucun point d'accumulation. Par contre sur $\overline{\mathbb{R}}$, qui est compact, elle converge vers $+\infty$.

Théorème 26 - Le produit de deux espaces compacts, pour la topologie produit, est compact.

Démonstration - Nous nous bornerons **à** donner la démonstration dans le cas de deux espaces métriques compacts E et F. Soit alors $(x_0, y_0), (x_1, y_1), \dots (x_n, y_n), \dots$ **une** suite d'éléments de $E \times F$.

D'après le théorème 25 la suite $x_0, x_1, x_2, \dots x_n, \dots$ admet au moins un point d'accumulation a dans E D'après le théorème 24, on peut donc en extraire une suite partielle $x_{p_0}, x_{p_1}, x_{p_2}, \dots x_{p_n}, \dots$ convergeant vers a . Alors la **suite** $y_{p_0}, y_{p_1}, y_{p_2}, \dots y_{p_n}, \dots$ **admet** au moins un point d'accumulation b **dans** F **,** et **d'après** le théorème 24 on peut en extraire une **suite partielle** $y_{p_{q_0}}, y_{p_{q_1}}, \dots y_{p_{q_n}}, \dots$ convergeant vers b . Alors **finalement** $(x_{p_{q_0}}, y_{p_{q_0}}), \dots (x_{p_{q_n}}, y_{p_{q_n}}), \dots$ est une suite partielle **de la suite initiale et elle converge** vers (a, b) On en déduit que la suite Initiale admet au moins un point d'accumulation dans $E \times F$, et cela prouve bien que ce produit est compact.

Théorème 27 - Si E est un espace compact, pour au'une suite d'éléments de E soit convergente vers a , il faut et **il** suffit qu'elle admette a, comme seul point d'accumulation.

Démonstration - La condition est manifestement nécessaire. Si la suite converge vers a elle l'admet comme point d'accumulation, et ne peut manifestement pas en admettre un autre-& **; il** existe en effet un voisinage $\mathcal{U}$ **de** a et un voisinage $\mathcal{V}$ de b **sans** point commun; or on devrait avoir $x_n \in \mathcal{U}$ pour tous les n , sauf un nombre fini, et $x_n \in \mathcal{V}$ pour une infinité de n**,ce** qui serait absurde. Montrons que la condltinn est suffisante.

Soit donc $x_0, x_1, x_2, \dots x_n, \dots$ une suite d'éléments de E admettant& comme seul point d'accumulation. SI cette suite n'était pas convergente, il existerait au moins un ouvert $\mathcal{O}$ contenant a et une suite **partielle** x_{p_n} de la suite donnée, tels que tous les x_{p_n} soient dans le complémentaire de $\mathcal{O}$. Comme ce **complémentaire** est fermé, il est compact d'après le théorème 22. La suite partielle des x_{p_n} devrait alors avoir sur $\complement\mathcal{O}$ au moins un point d'accumulation, et par conséquent aussi la suite **initiale,ce** qui contredit l'hypothèse qu elle admet a comme seul point d'accumulation.

Remarque, Le même **résultat** serait manifestement faux sur la droite réelle $\mathbb{R}$. Par exemple, la suite $1, 1, 2, \frac{1}{2}, 3, \frac{1}{3}, \dots n, \frac{1}{n}, \dots$ admet 0 comme seul point d'accumulation, et elle n'est **manifestement** pas convergente.

Limite supérieure et limite inférieure d'une suite réelle

soit $x_0, x_1, x_2, \dots x_n, \dots$ une suite d'éléments de $\overline{\mathbb{R}}$. Comme $\overline{\mathbb{R}}$ est compact, l'ensemble de ses points d'accumulation n'est pas vide; c'est une partie F de $\overline{\mathbb{R}}$. F est toujours fermé: car, si $A_n = \{x_n, x_{n+1}, \dots\}$, F est l'intersection des ensembles fermés $\overline{A}_n$ (voir page 75). Alors F est un compact de $\overline{\mathbb{R}}$. Il a un maximum L et un minimum ℓ. L s'appelle la limite supérieure de la suite, ℓ sa limite inférieure; on écrit $L = \lim\limits_{n\to\infty} \sup x_n$, $\ell = \lim\limits_{n\to\infty} \inf x_n$. La limite supérieure est caractérisée par la propriété suivante : quels que soient L_1 et L_2 tels que $L_1 < L < L_2$, tous les x_n sont $\leq L_2$, sauf au plus pour un nombre fini de valeurs de n, et il y a une infinité de valeurs de n pour lesquelles $x_n \geq L_1$.

(2)

On ne doit pas confondre limite supérieure et borne supérieure. Par exemple $1, \frac{1}{2}, \frac{1}{3}, \dots \frac{1}{n}, \dots$ a pour borne superieure 1, mais converge vers 0 , qui est donc son seul point d'accumulation et aussi sa limite supérieure.

Pour qu'une suite de $\overline{\mathbb{R}}$ converge, il faut et il suffit que sa limite supérieure soit égale à sa limite inférieure (car cela exprime qu'elle a un seul point d'accumulation).

§ 8 PROPRIÉTÉS DES FONCTIONS CONTINUES SUR UN ESPACE COMPACT

Théorème 28 - L'image directe d'un compact par une application continue est compacte.

Démonstration Soit f une application continue d'un espace topologique E dans un espace topologique F, E est supposé compact. Bien entendu il n'est pas question de dire que F lui aussi est compact, mais nous devons demontrer que l'image directe $f(E)$ de E par f, en tant que sous-espace de F muni de la topologie induite, est compacte. Soit donc $\mathcal{R}$ un recouvrement ouvert de $f(E)$. Les images réciproques $f^{-1}(\mathcal{O}_i)$ des ouverts $\mathcal{O}_i$ de **R** forment un recouvrement ouvert de E. Si en effets est un point quelconque de E, son image $f(x)$ appartient à l'un au moins des ouverts, soit $\mathcal{O}_i$, et par conséquents appartient à $f^{-1}(\mathcal{O}_i)$; par ailleurs ce sont des ouverts puisque f est continue

Comme alors E est supposé compact, il suffit d'un nombre fini des ouverts $f^{-1}(\mathcal{O}_i)$, par exemple $f^{-1}(\mathcal{O}_1), f^{-1}(\mathcal{O}_2), \dots f^{-1}(\mathcal{O}_n)$, pour recouvrir E. Mais alors cela signifie que les ouverts $\mathcal{O}_1, \mathcal{O}_2, \dots \mathcal{O}_n$ forment un recouvrement de $f(E)$. Si en effet $y \in f(E)$, l'image réciproque $f^{-1}(\{y\})$ n'est pas vide; soit x un élément de cette image réciproque, l'un au moins des ouverts $f^{-1}(\mathcal{O}_1), f^{-1}(\mathcal{O}_2), \dots f^{-1}(\mathcal{O}_n)$, par exemple $f^{-1}(\mathcal{O}_k)$, contient x, et par conséquent $\mathcal{O}_k$ contient $f(x) = y$, ce qui prouve bien ce que nous avancions. Il en résulte que la définition des espaces compacts est satisfaite par $f(E)$: tout recouvrement ouvert admet un sous-recouvrement fini.

Corollaire - Toute bijection continue d'un espace compact E sur un espace topologique F est un homéomorphisme.

En effet l'image directe de toute partie fermée de E est alors l'image directe d'une partie compacte, d'après le théorème 22; par conséquent elle est compacte d'après ce que nous venons de voir, donc fermée d'après le théorème 21. Or nous avons vu au théorème 11 que toute application bijective et continue pour laquelle l'image directe d'un fermé est un fermé, est un homéomorphisme.

Remarque Par contre il serait faux de croire que l'image réciproque d'un compact par une application continue soit un compact. Soit par exemple f une application constante d'un espace non compact E dans un espace quelconque F ; l'image de E tout entier est une partie réduite à un point b de F, c'est-à-dire à un compact; alors l'image réciproque par f de $\{b\}$ n'est pas un compact puisque c'est E tout entier. Nous remarquons ainsi que, si f est une application continue de E dans F, ce sont les images réciproques des parties ouvertes ou des parties fermées, qui sont des parties ouvertes ou des parties fermées, et les images directes des parties compactes, qui sont des parties compactes. Il y a là une distinction très importante. Naturellement en mélangeant les deux sortes de résultats, on obtient le résultat suivant : Si E est compact, et si f est une application continue de E dans F, l'image directe de toute partie fermée de E est fermée dans F ; en effet, une partie fermée de E est alors compacte puisque E est compact, donc son image directe est compacte dans F donc fermée dans F.

Théorème 29 - Toute application continue d'un espace compact non vide dans $\overline{\mathbb{R}}$ admet un maximum et un minimum.

Démonstration - Rappelons qu'on appelle borne supérieure d'une fonction à valeurs dans $\overline{\mathbb{R}}$ définie sur un ensemble E, la borne supérieure de l'ensemble de ses valeurs; on dit que c'est un maximum si cette borne est atteinte pour une valeur particulière de la variable. Soit donc f une application continue d'un espace compact E dans la droite achevée $\overline{\mathbb{R}}$. D'après le théorème précédent, l'image directe $f(E)$ est un compact non vide de $\overline{\mathbb{R}}$. D'après le théorème 23, c'est donc un ensemble fermé non vide de $\overline{\mathbb{R}}$. Un ensemble non vide de $\overline{\mathbb{R}}$ a une borne supérieure, d'après le théorème 2 du chapitre 1 (modifié comme il est indiqué page 22),et en outre, d'après ce même théorème, cette borne est adhérente à l'ensemble. Comme l'ensemble est fermé, cette borne supérieure appartient donc à l'ensemble et c'est bien un maximum. Même démonstration pour le minimum.

Ce théorème 23 est essentiel. Il est bon de s'exercer à en donner plusieurs démonstrations. On en trouvera une, basée sur la propriété de REINE- BOREL-LEBESGUE, page 84 . En voici encore une autre. Soit M la borne supérieure de f. soit $M_0, M_1, M_2, \ldots M_n, \ldots$ une suite croissante de nombres $< M$, tendant vers M. Pour tout n, l'ensemble $F_n : \{x ; x \in E, M_n \leq f(x) \leq M\}$ est non vide, d'après la caractérisation de la borne supé-rieure (théorème 2 du Chapitre I), et fermé, comme image réciproque, par la fonction continue f, de l'intervalle fermé $[M_n, M]$ de $\mathbb{R}$ (théorème 8). La suite des F_n est donc une suite décroissante d'ensembles fermés non vides d'un compact E ; d'après le corollaire 2 du théorème 20, leur intersection n'est pas vide, or cette Intersection est l'ensemble des points x de E où $f(x) = M$, donc il existe au moins un tel point, et M est un maximum.

On peut encore dire ceci. Les notations étant les mêmes que plus haut, quel que soit n, il existe au moins un point x_n de E tel que $M_n \leq f(x_n) \leq M$. La suite des x_n admet au moins un point d'accumulation (WEIERSTRASS-BOLZANO) a, et on peut extraire une suite partielle $x_{p_1}, x_{p_2}, \ldots x_{p_n}, \ldots$ convergeant vers a (théorème 24). Comm"e f est continue en a, la suite des $f(x_{p_n})$ converge vers $f(a)$ (théorème 16); comme $M_{p_n} \leq f(x_{p_n}) \leq M$, $f(x_{p_n})$ converge vers M, donc $f(a) = M$, et M est un maximum.

<u>Corollaire 1 - Si f est une fonction réelle (c'est-à-dire à valeurs dans $\mathbb{R}$) continue sur un compact E, elle est bornée; si en tout point x de E, $f(x)$ est > 0, il existe un nombre fixe $\delta > 0$ tel que, pour tout x de E, $f(x) \geq \delta$.</u>

En effet f peut être considérée comme à valeurs dans $\overline{\mathbb{R}}$. Elle a une borne supbrieure M, mais comme c'est un maximum, c'est la valeur $f(a)$ pour un point a convenable de E, donc $M < +\infty$, et f est bornée supérieurement. Elle est aussi bornée inférieurement. Si f est partour 7 0, son minimum δ est sa valeur en un point b convenable donc $\delta > 0$, et on a bien toujours $f(x) \geq \delta$.

<u>Corollaire 2 - soit f une fonction réelle continue sur un espace métrique E .et soit K un compact de F . Il existe tout un voisinage de K sur lequel f est bornée, Si $f(x)$ est >0 pour tout x de K, il existe'un nombre $\delta > 0$ et tout un voisinage de K sur lequel $f(x) \geq \delta$.</u>

En effet, d'après le corollaire 1, $f(K)$ est un ensemble borné de $\mathbb{R}$. 11 existe donc un nombre M tel que $f(K) \subset]-M, +M[$; alors $K \subset f^{-1}(]-M, +M[)$. Or c'est là un ouvert (puisque f est continue et $]-M, +M[$ ouvert), il contient K, et f y reste bornée en module par M.

Si maintenant $f(x) > 0$ pourtoutx de K, le corollaire 1 indique qu'il existe un nombre $\delta' > 0$ tel que $f(x) \geq \delta'$ pour tout x de K. Posons $\delta = \frac{\delta'}{2}$. Alors $f(K) \subset]\delta, +\infty[$, donc $K \subset f^{-1}(]\delta, +\infty[)$, qui est donc un ouvert contenant K, sur lequel $f(x) \geq \delta$.

<u>Remarque</u> - Ces résultats seraient naturellement complètement inexacts pour une fonction discontinue. Une fonction discontinue peut être partout finie et cependant non bornee; par exemple la fonction $\frac{1}{x}$, définie sur le complémentaire de l'origine dans l'intervalle $[0,1]$ compact, et prolongée par la valeur 0 à l'origine, est une fonction réelle sur un espace compact mais présentant un point de discontinuité; cette fonction est partout finie mais n'est pas bornée. Par ailleurs les résultats seraient également faux pour une fonction continue sur un espace non compact; par exemple, sur la droite réelle, la fonction x est continue, mais n'est pas bornée; la fonction $\operatorname{Arc\,tg} x$ est continue et bornée, mais n'atteint pas sa borne supérieure, celle-ci par conséquent n'est pas un maximum.

Bien entendu, rien ne dit qu'il n'y ait qu'un seul point où la fonction soit maxima ou minima, comme le montre l'exemple d'une fonction constante.

<u>Applications</u> - Soient E un espace métrique, F une partie fermée de E, et a un point quelconque. On appelle distance de a à F, et on note $d(a,F)$, la borne inférieure des distances de a aux points de F; comme toutes ces distances sont ≥ 0, la distance de a à F est elle-même ≥ 0.

Si $a \in F$, $d(a,F) = 0$. D'autre part, si $a \notin F$, $d(a,F) > 0$.

En effet, s'il n'en était pas ainsi, cela signifierait que toute boule de centre a contiendrait au moins un point de F, par conséquent a serait adhérent à F, et, comme F est fermé, a appartiendrait à F, ce qui est contraire à l'hypothèse.

Montrons que $a \longrightarrow d(a,F)$ est une application continue de E dans $\mathbb{R}$, la distance d'un point à un ensemble ferme varie continûment avec le point.

Soient en effet a, a', deux points. Quel que soit $\delta > 0$, il existe un point x de F tel que $d(a,x) \leq d(a,F) + \delta$, d'après les propriétés de la borne inférieure (théorème 2 du Chap. 1); alors $d(a',x) \leq d(a,x) + d(a,a') \leq d(a,F) + d(a,a') + \delta$; donc $d(a',F) \leq d(a',x) \leq d(a,F) + d(a,a') + \delta$. comme δ est quelconque, on a $d(a',F) \leq d(a,F) + d(a,a')$. Mais, en échangeant les rôles de a et a', on a aussi $d(a,F) \leq d(a',F) + d(a,a')$, d'où finalement

(II.8;1) $$|d(a',F) - d(a,F)| \leq d(a,a').$$

Alors $d(a,a') \leqslant \varepsilon$ **entraîne** $|d(a', F) - d(a,F)| \leqslant \varepsilon$, ce qui **entraîne** la **continuité** de l'application $a \longrightarrow d(a,F)$.

On peut alors se **demander, a étant** donne, si la borne **inférieure** $d = d(a,F) = \operatorname*{Inf}_{x\in F} d(a,x)$ **est** un minimum, autrement dit s'il existe un point c de F, tel que $d(a,c) = d$. Nous allons montrer qu'il en est bien ainsi, si dans E toutes les boules **fermées** sont compactes (ce qui est le cas, d'après le théorème 23, si E est un espace vectoriel normé de dimension finie).

Considérons la boule B de centre a et de rayon $d + 1$; cette boule fermée est compacte d'après l'hypothèse. Son Intersection avec le fermé F est une partie fermée de la boule compacte B (théoréme 5), et par conséquent compacte (théorème 22). Il en résulte que la fonction $x \longrightarrow d(a,x)$, qui est une fonction continue sur le compact $B \cap F$, admet un minimum. Si alors c est un point où ce minimum est atteint, il répond bien à la question; on a en effet $d(a,c) \leqslant d(a,x)$ pour tout point x de $B \cap F$, mais aussi a fortiori $d(a,c) \leqslant d(a,x)$ pour tout autre point x de F, **puisque** tous les autres points de F sont à une distance $\geqslant d+1$ de a, et que $d(a,c) \leqslant d+1$.

Finalement $d(a,c) \leqslant d(a,x)$ pour tout $x \in F$, donc $d(a,c)$ est bien un minimum de $d(a,x)$ **pour** $x \in F$ (et $d(a,c) = d$).

Soit maintenant E un espace **métrique** quelconque, et soient F_1, F_2, deux parties fermées de E. On appelle distance de F_1 et F_2, et on note $d(F_1, F_2)$, la borne Inférieure des distances $d(x_1, x_2)$, **pour** $x_1 \in F_1$, $x_2 \in F_2$. Même si $F_1 \cap F_2 = \emptyset$, il peut arriver que $d(F_1, F_2) = 0$, comme le montre l'exemple d'une hyperbole et de son asymptote dans le plan euclidien; cet exemple montre aussi que la borne inférieure n'est pas un minimum.

Nous allons voir que, si F_1 est compact, et **si** $F_1 \cap F_2 = \emptyset$, alors $d = d(F_1,F_2) > 0$; en outre, si les boules fermées de E sont toutes compactes, cette borne inférieure est un minimum.

On a en effet $d = \operatorname*{Inf}_{x_1 \in F_1} d(x_1, F_2)$. Mals nous avons vu **que** $x_1 \longrightarrow d(x_1, F_2)$ est **une fonction** continue; sur **le** compact F_1, elle est toujours > 0, elle admet donc un **minimum** > 0. autrement dit il existe un point c_1 de F_1 tel que $d = d(c_1, F_2) > 0$.

Si maintenant toutes les boules fermées de E sont compactes, nous avons vu plus haut qu'il existe un point c_2 de F_2 tel que $d(c_1,c_2)=d(c_1,F_2)$, alors $d(c_1,c_2) = d = d(F_1,F_2)$ est bien un minimum.

Une application qui sera souvent utilisée est la suivante. Soit Ω un ouvert d un espace métrique E, K un compact $\subset \Omega$.

Alors $d=d(K,\complement\Omega)>0$. Cela résulte de ce que K est un compact sans point commun avec le fermé $\complement\Omega$.

<u>Théorème de d'Alembert 30.</u> <u>Tout polynôme d'une variable complexe à coefficients complexes de degré m admet m racines complexes.</u>

<u>Démonstration</u> Naturellement, quand nous disons m racines complexes, nous entendons que nous comptons chaque racine autant de fois que l'indique son ordre de multiplicité.

Pour démontrer le théorème, il suffit naturellement de montrer que le polynôme admet au moins une racine a, pour $m \geq 1$. Si en effet alors on divise par $z-a$, on est ramené à un polynôme de degré $m-1$, auquel on peut à nouveau appliquer même raisonnement. On démontre ainsi le théorème par récurrence sur le degré du polynôme. Supposons donc que le polynôme

(II.8;2) $$P(z) = a_0 + a_1 z + \cdots + a_m z^m$$

n'admette aucune racine, nous allons démontrer que nous aboutissons à une contradiction. On sait que $|P(z)|$ tend vers $+\infty$ lorsque $|z|$ tend vers $+\infty$ donc il existe un nombre R tel que l'on ait, à l'extérieur du cercle de centre 0 et de rayon R du plan complexe, l'inégalité $|P(z)|\geq|P(0)|$.

Appelons μ le minimum >0 (corollaire du théorème 29) de la fonction continue $|P|$ dans le compact $|z|\leq R$, et soit z_0 un point tel que $|P(z_0)|=\mu$. Comme on a, pour $|z|\geq R$, $|P(z)|\geq|P(0)|\geq\mu$, l'inégalité $|P(z)|\geq\mu$ est vérifiée pour tout z ; on en déduit que μ est le minimum du module du polynôme P dans tout le plan complexe. Considérons le développement de TAYLOR du polynôme P au point z_0 :

(II.8;3) $$P(z) = P(z_0) + c_k(z-z_0)^k + c_{k+1}(z-z_0)^{k+1} + \cdots + c_m(z-z_0)^m .$$

Nous mettons en évidence le premier terme variable non nul de ce développement. On a $|P(z_0)|=\mu$. Il existe un nombre ρ assez petit pour que, sur le cercler de centre z_0 et de rayon ρ, on ait l'inégalité

(II.8;4) $$|c_{k+1}(z-z_0)^{k+1} + \cdots + c_m(z-z_0)^m| < |c_k(z-z_0)^k| = |c_k|\rho^k .$$

On peut supposer ρ tel que $|c_k|\rho^k<\mu$.

Alors si z parcourt la circonférence Γ, la quantité $c_k(z-z_0)^k$ parcourt toute la circonférence de centre origine et de rayon $|c_k|\rho^k$, donc $P(z_0)+c_k(z-z_0)^k$ parcourt toute la circonférence de centre $P(z_0)$ et de rayon $|c_k|\rho^k$; et par suite il existe un point z_1 de Γ tel que $P(z_0)+c_k(z_1-z_0)^k$ soit sur le segment $[0, P(z_0)]$ du plan complexe.

On a alors :

$$|P(z_0)+c_k(z_1-z_0)^k| = \mu - |c_k|\rho^k . \tag{II.8;5}$$

Il en résulte que l'on a la majoration

$$|P(z_1)| \leqslant |P(z_0)+c_k(z_1-z_0)^k| + |c_{k+1}(z_1-z_0)^{k+1}+\cdots+c_m(z_1-z_0)^m| < (\mu-|c_k|\rho^k)+|c_k|\rho^k = \mu, \tag{II.8;6}$$

mais $|P(z_1)|<\mu$ est contradictoire avec le fait que μ est le minimum du module du polynôme? .

Généralisation du théorème 29 - Soit f une fonction à valeurs dans% sur un espace topologique E . On dit que cette fonction est semi-continue supérieurement en un point a de E si, quel que soit $b_1 > f(a)$, il existe un voisinage $\mathcal{V}'$ de a dans E tel que $x \in \mathcal{V}'$ entraîne $f(x) \leqslant b_1$. On dit qu'elle est semi-continue inférieurement si, quel que soit $b_2 < f(a)$, il existe un voisinage $\mathcal{V}''$ de a dans E , tel que $x \in \mathcal{V}''$ entraîne $f(x) \geqslant b_2$.

Une fonction réelle est continue si et seulement si elle est à la fois semi-continue supérieurement et semi-continue inférieurement. Le théorème 29 admet alors la généralisation suivante :

Théorème 30 bis - Sur un espace compact, toute fonction semi-continue supérieurement admet un maximum, et toute fonction semi-continue Inférieurement admet un minimum.

Démonstration - Soit en effet M la borne supérieure dans $\overline{\mathbb{R}}$ de f , supposée semi-continue supérieurement sur le compact E. Supposons que ce ne soit pas un maximum. Pour tout x de E , on aura $f(x)<$ M ; d'après la semi-continuité supérieure, si $\varepsilon_x < M - f(x)$, on pourra trouver un voisinage ouvert $\mathcal{V}_x$ de x dans lequel f reste majorée par $M_x = f(x)+\varepsilon_x < M$. Quand on fait varier x , on voit que le système de tous les $\mathcal{V}_x$ est un recouvrement ouvert de E . Comme E est compact, il en existe un sous-recouvrement fini; autrement dit, il existe un ensemble fini de points $x_1, x_2, \dots x_n$, de E , tel que E soit la réunion de $\mathcal{V}_{x_1}, \mathcal{V}_{x_2}, \dots \mathcal{V}_{x_n}$. Comme, dans $\mathcal{V}_{x_i}$, f est majorée par $M_{x_i} < M$,

elle est majorée dans E tout entier par le maximum $\mathcal{M}$ des nombres M_{x_1}, M_{x_2} . M_{x_n}, qui est $< M$. Ceci est contradictoire avec le fait que M est la borne supérieure de f .
On voit donc que M est bien un maximum .

Continuité uniforme

<u>Les propriétés que nous allons voir sont absolument spéciales aux espaces métriques</u> et ne peuvent pas s'étendre aux espaces topologiques généraux. D'ailleurs deux métriques équivalentes ne donnent pas le même résultat *

<u>Définition</u> - Soient E et F deux espaces metriques. On dit qu'une application f de E dans F est uniformément continue, si, quel que soit $\varepsilon > 0$, il existe $\eta > 0$ tel que $d(x', x'') \leqslant \eta$ entrafne $d(f(x'), f(x'')) \leqslant \varepsilon$.

En abrégé :

(II.8;7) $$(\forall \varepsilon > 0)(\exists \eta > 0)(\forall x' \in E, \forall x'' \in E, d(x', x'') \leq \eta) : d(f(x'), f(x'')) \leq \varepsilon .$$

Toute fonction uniformément continue est évidemment continue, mals la réciproque n'est pas exacte. Si f est une fonction partout continue, alors, pour tout point a de E , quel que soit $\varepsilon > 0$, il existe $\eta > 0$ tel que $d(x, a) \leqslant \eta$ entraîne $d(f(x), f(a)) \leq \varepsilon$, mals le nombre η que nous déterminons ainsi dépend à la fols de a et de ε . Dire que la fonction est uniformément continue, c'est dire qu'il est possible de choisir η dépendant seulement de ε , (Voir à ce sujet chapitre 1, page35).

Sur la droite réelle, la fonction x est uniformément continue, mais la fonction x^2 ne l'est pas. En effet on a

$|(x+h)^2 - x^2| = |2hx + h^2| \geqslant 2hx$ si $h \geqslant 0, x \geqslant 0$; alors, pour ε donné, si nous cherchons h tel que l'on ait $|2hx + h^2| \leqslant \varepsilon$, on doit nécessairement prendre

* - Autrement dit une application d'un espace métrique E dans un espace métrique F pourra être uniformément continue, et cesser de l'être quand on remplace les métriques par des métriques équivalentes. Toutefois si ce sont des métriques définies par des normes sur des espaces vectoriels, l'uniforme continuité subsistera si on remplace les normes par des normes équivalentes, à cause du théorème 12.

$|h| \leq \frac{\varepsilon}{2|x|}$, et il est par suite impossible de choisir η indépendant de x .

On dit qu'une application f d'un espace métrique E dans un espace **métrique** F <u>vérifie une condition de **HÖLDER** ou de **LIPSCHITZ** d'ordre α</u> , $0 < \alpha \leq 1$ * , **s'il** existe une constante k telle que, quels que soient x' et x'' dans E , on ait

$$d(f(x'), f(x'')) \leq k\left[d(x', x'')\right]^{\alpha}. \tag{II.8;8}$$

Pour $\alpha = 1$, on dit simplement que f <u>vérifie une condition de **LIPSCHITZ**</u> (sans **spécifier** : d'ordre **1**), <u>ou qu'elle est **lipschitzienne**</u>.

<u>Une fonction vérifiant une condition de HÖLDER **d'ordre** α, a fortiori une fonction **lipschitzienne**, est uniformément continue.</u> Bien entendu, la réciproque n'est pas vraie, comme le montre l'exemple de la fonction égale à $\frac{1}{\log x}$ sur l'intervalle $\left]0, \frac{1}{2}\right]$, nulle pour $x = 0$; elle estcontinue, donc uniformément continue comme le montrera le théorème 31, **et** $|f(x) - f(0)| = \frac{1}{|\log x|}$ n'est pas majore par une expression $k|x|^{\alpha}$.

<u>Théorème 31</u> - <u>Toute application d'un espace métrique **compact** E dans un espace métrique F , si elle est continue, **est uniformément** continue.</u>

<u>Démonstration</u> - Supposons que f ne soit pas uniformément continue, et montrons que nous aboutissons **à** une contradiction. Si f n'est pas uniformément continue, c'est que

$$(\exists \varepsilon > 0)\,(\forall \eta > 0)\,\big(\exists x' \in E,\ \exists x'' \in E,\ d(x', x'') \leq \eta\big) : d\big(f(x'), f(x'')\big) > \varepsilon .$$

Soit ε le nombre Intervenant au début de cette relation. **Alors, quel** que soit l'entiern ≥ 1 , on peut trouver deux points x'_n , x''_n tels que l'on ait $d(x'_n, x''_n) \leq \frac{1}{n}$ et $d\big(f(x'_n), f(x''_n)\big) > \varepsilon$.

Comme E est **<u>compact</u>**, on peut extraire de la suite des x'_n une suite partielle $n \to x'_{p_n}$ convergeant vers un **point** x **(th. 24 et 25)** Alors la relation $d(x'_n, x''_n) \leq \frac{1}{n}$ montre que la suite partielle des x''_{p_n} est elle aussi convergente vers x . Comme f est

* Le cas $\alpha > 1$ est sans **intérêt.** Par exemple, pour $E = F = \mathbb{R}$, **si** $\alpha > 1$, **on voit** que f a une **dérivée** partout nulle, car $\left|\frac{f(x+h)-f(x)}{h}\right| \leq k|h|^{\alpha-1}$, donc elle est constante.

continue, les suites $f(x'_{p_n})$, $f(x''_{p_n})$ sont nécessairement convergentes vers $f(x)$ (théorème 16) . Alors, pour n suffisamment grand, on a nécessairement $d(f(x'_{p_n}), f(x)) \leq \frac{\varepsilon}{2}$, $d(f(x''_{p_n}), f(x)) \leq \frac{\varepsilon}{2}$, donc $d(f(x'_{p_n}), f(x''_{p_n})) \leq \varepsilon$, ce qui est bien en contradiction avec la construction des x'_n et des x''_n .

Comme nous l'avons vu plus haut avec l'exemple de la droite réelle $\mathbb{R}$, une fonction peut être continue sur un espace métrique non compact, sans être uniformément continue. On a souvent besoin d'un théoréme 31 amélioré; voir ** page 355.

§ 9 ESPACES CONNEXES

On voit Intuitivement que certains espaces topologlques peuvent être considérés comme d'un seul tenant, par exemple une sphère, une boule, dans un espace $\mathbb{R}^n$, alors que d'autres sont composes de plusieurs 'morceaux" distincts, par exemple l'espace formé par la réunion de deux sphères sans point commun,ou le complémentaire d'une sphère dans $\mathbb{R}^n$. Il s'agit de préciser cette notion intuitive.

<u>Définition</u> - <u>On dit qu'un espace topologique E est connexe, s'il n'admet pas de partition formee de deux parties ouvertes, ou encore s'il n'admet pas de partition</u> (*) <u>formée de deux parties fermées, ou encore s'il n'existe pas dans E d'autres parties qui soient à la fols ouvertes et fermées que E et $\emptyset$</u>.

Ces trois définitions sont manifestement équivalentes d'après la définition des fermés comme complementalres des ouverts. Comme la compacité, la connexité est une propriété de l'espace topologlque lui-même; cependant, si F est une partie de E , on dira que F est une partie connexe, si en tant qu'espace muni de la topologie induite, F est connexe.

<u>Théorème 32</u> - <u>Pour qu'une partie E de la droite réelle achevée $\overline{\mathbb{R}}$ soit un espace topologique connexe, Il faut et Il suffit qu'elle soit un intervalle ouvert, semi-ouvert ou fermé</u>.

* Définition page 13 du chap. 1 -

Démonstration - Soit donc E une partie connexe de $\overline{\mathbb{R}}$.
Soient x et y deux points distincts de E ; montrons que tout l'intervalle fermé $[x,y]$ est contenu dans E . S'il n'en était pas ainsi, il existerait au moins un point z de cet intervalle, qui n'appartiendrait pas à E . Alors, sur $\overline{\mathbb{R}}$, les deux ensembles $[-\infty, z[$, $]z, +\infty]$ sont tous les deux ouverts, et leur intersection avec E définirait sur ce dernier une partition (parce que $z \notin E$) formée de deux parties ouvertes. Ainsi E ne serait pas connexe. Si alors nous appelons a (resp. b) la borne inférieure (resp. supérieure) des points de E , ce que nous venons de voir montre que E est nécessairement identique à l'un des quatre intervalles $[a,b]$, $[a,b[$, $]a,b]$, $]a,b[$.

Réciproquement, soit E un intervalle de $\overline{\mathbb{R}}$. Soit **A** une partie non vide de E , à la fois ouverte et fermée dans E ; montrons que **A = E** , ce qui prouvera que E est connexe.
Soit c un élément de **A** . Considérons l'ensemble H de tous les x de E tels que $[c,x] \subset A$; soit γ la borne supérieure de H dans $\overline{\mathbb{R}}$. Pour tout γ' tel que $c \leqslant \gamma' < \gamma$, il existe un élément de **H** qui majore γ' , donc $\gamma' \in$ **A** et $[c, \gamma'] \subset A$; comme A est fermé, on en déduit que $\gamma \in$ **A** et que $[c,\gamma] \subset A$, sauf si $\gamma = b$ et que $b \notin E$. Si on avait $\gamma < b$, alors, A étant aussi ouvert, il existerait un $\gamma'' > \gamma$ tel que $[\gamma, \gamma''] \subset$ **A** donc $[c, \gamma''] \in$ **A** et $\gamma'' \in H$, ce qui est absurde puisque γ est la borne supérieure de H . On a donc $\gamma = b$, $[c, b[\subset$ **A** , et en outre $b \in A$ si $b \in E$.
En faisant à gauche de c le raisonnement que nous venons de faire à droite, on voit de même que $]a,c] \subset A$ et $a \in A$ si $a \in E$; donc **A =** E , et E est bien connexe.

Corollaire - L'ensemble $\mathbb{Q}$ des nombres rationnels n'est pas connexe.

Théorème 33 - L'image directe par une application continue d'un espace topologique connexe, est connexe.

Démonstration Soit f une application continue de **E** dans F nous supposons E connexe. Naturellement, il n'est pas question de dire que F soit connexe; mais nous allons montrer que l'image directe $f(E)$ de E est connexe. Si en effet, il n'en était pas ainsi, il existerait sur $f(E)$ une partition formée de deux parties ouvertes **A** et B ; alors leurs images réciproques formeraient une partition de E (d'après (I,2;3) , leur intersection est vide; si $x \in E$, $f(x)$ appartient à A ou à B, donc x à $f^{-1}(A)$ ou $f^{-1}(B)$, donc leur réunion est E ; aucun n'est vide, car **A**, par exemple, n'est pas vide, et comme il est contenu dans $f(E)$ son image réciproque n'est pas vide) et seraient des ouverts d'après le théorème 7, ce qui serait contradictoire avec l'hypothèse que E est connexe.

Corollaire - Si f est une fonction continue sur un espace connexe E, à valeurs dans $\overline{\mathbb{R}}$, l'ensemble de ses valeurs est un intervalle ouvert, semi-ouvert ou fermé de $\overline{\mathbb{R}}$.

Démonstration En effet ce doit être une partie connexe de $\overline{\mathbb{R}}$, il suffit alors d'appliquer le précédent théorème. On exprime souvent ce corollaire en prenant le cas particulier où E est lui-même un intervalle ouvert, semi-ouvert ou fermé de $\overline{\mathbb{R}}$. Enfin on dit souvent simplement qu'une fonction réelle continue sur un espace connexe ne peut prendre deux valeurs sans prendre toutes les valeurs intermédiaires. Cette propriété s'appelle la propriété des valeurs intermédiaires. Il est bon de noter que cette propriété ne caractérise pas les fonctions continues, il existe des fonctions discontinues sur un espace connexe et qui possèdent la même propriété. Par exemple la fonction définie par $f(x) = \sin\frac{1}{x}$ pour $x \neq 0$, $f(0)=0$, définie sur la droite réelle, est discontinue à l'origine, et pourtant possède cette propriété que l'ensemble de ses valeurs est l'intervalle fermé $[-1, +1]$; dans tout intervalle (a, b), elle ne peut prendre 2 valeurs sans prendre toutes les valeurs intermédiaires.

Par contre on peut donner du théorème la réciproque suivante :

Théorème 34 - Si un espace topologique E est tel que toute fonction reelle continue sur E ne puisse prendre deux valeurs sans prendre aussi toutes les valeurs intermédiaires, alors E est connexe.

Démonstration - Si en effet E n'était pas connexe, il possèderait une partition formée de deux parties ouvertes A et B ; alors la fonction réelle prenant la valeur 0 sur A et la valeur 1 sur B serait continue, parce que l'image réciproque de tout ouvert de $\mathbb{R}$ serait l'une des quatre parties ouvertes $\emptyset$, A, B, E • elle prendrait les valeurs 0 et 1 sans prendre aucune'des valeurs intermédiaires, ce qui serait contraire à l'hypothèse. Donc E est connexe.

Espaces connexes par arcs

Il est utile d'avoir des critères pour reconnaître qu'un espace est connexe. C'est pourquoi nous introduirons la notion d'espace connexe par arcs.

On appelle arc ou chemin joignant un point a à un point b d'un espace topologique E, toute application continue f d'un intervalle $[\alpha, \beta]$ de la droite réelle $\mathbb{R}$ dans E, telle que $f(\alpha) = a$, $f(\beta) = b$. On dit aussi que a et b sont l'origine et l'extrémité du chemin.

Il est bon de ne pas confondre cette application, qui constitue la définition même du chemin, avec l'image par de l'intervalle $[\alpha, \beta]$, qu'on appelle l'image du chemin.

Par exemple, si l'application est constante, cette image se réduit à un point, on pourra dire d'ailleurs dans ce cas que le chemin se réduit à un point, mais le chemin n'en est pas moins l'application elle-même.

De la même manière si nous considérons une lemniscate de BERNOULLI, elle peut être "parcourue" de deux manières différentes, alors que la lemniscate en tant qu'ensemble est la même dans les deux cas; les deux manières de la parcourir correspondent a deux chemins différents, c'est-à-dire à deux applications différentes d'un intervalle de $\mathbb{R}$ dans le plan.

On dira qu'un chemin f passe par un point c de E si l'image $f([\alpha, \beta])$ contient c ; on dira que le chemin rencontre une partie A de E si $f([\alpha, \beta]) \cap A \neq \emptyset$.

On voit immédiatement que si les deux points a et b peuvent être joints par un chemin, et si les deux points b et c peuvent également être joints par un chemin, alors les deux points a et c peuvent encore être joints par un chemin.

Théorème 35 - Si E est un espace topologique, tel que 2 quelconques de ses points puissent être joints par un chemin, alors E est connexe.

Démonstration - Supposons qu'il n'en soit pas ainsi. Alors il existerait une partition de E formée de deux ensembles ouverts A et B. Soit a et b des éléments respectifs de A et B. Par hypothèse, il serait possible de joindre par un chemin a et b ; soit K l'image de ce chemin. Alors $K \cap A$ et $K \cap B$ seraient deux parties ouvertes complémentaires de K (théorème 5) et aucune d'elles ne serait vide puisque a et b seraient respectivement dans l'une et dans l'autre.

On aurait ainsi obtenu une partition de K formé de deux parties ouvertes, ce qui est contraire au théorème 33 qui dit que K , image par une application continue de l'espace connexe $[\alpha, \beta]$, est connexe. Nous aboutissons donc bien à une contradiction, et E est connexe.

La réciproque de ce théorème est Inexacte. Si par exemple, nous considérons, dans le plan euclidien $\mathbb{R}^2$ muni de deux axes de coordonnées, l'ensemble E formé de la réunion de la courbe représentative de la fonction $y = \sin \frac{1}{x}$, $x \neq 0$, et du segment vertical

$x = 0$, $|y| \leqslant 1$, on vérifie que E est connexe, et que cependant les deux points $(0,0)$, $(a, \sin \frac{1}{a})$ ne peuvent pas être joints par un chemin.

On dit qu'un espace E est <u>connexe par arcs,</u> si deux quelconques de ses points peuvent être joints par un arc ou chemin. C'est une propriété plus forte que la connexité.

<u>Exemple</u> - Tout espace vectoriel normé est connexe par arcs. En effet deux quelconques de ses points peuvent être joints par un chemin, défini exactement par le segment qui les a comme origine et comme extrémité.

On vérifie également qu'une boule, une sphère dans un espace vectoriel normé sont connexes par arcs.

<u>Théorème du passage des douanes 36</u> - <u>Si E est un espace topologique, A une partie de E, tout chemin joignant un point de l'intérieur de A à un point de l'extérieur de A rencontre nécessairement la frontière de A</u> .

<u>Démonstration</u> Comme l'image K du chemin est connexe d'après le théorème 32, le présent théorème est un cas particulier du suivant :

<u>Toute partie B connexe de E , et rencontrant à la fois l'intérieur et l'extérieur de A , rencontre nécessairement sa frontière.</u> Cette propriété est évidente car, s'il n'en était pas ainsi, B serait contenu dans la réunion de l'intérieur et de l'extérieur de A , et les intersections de B avec ces deux parties définiraient une partition de B formée de deux parties ouvertes, ce qui serait Impossible.

§ 10 COMPLÉMENTS DE TOPOLOGIE GÉNÉRALE SUR LES ESPACES CONNEXES

Théorème 36 bis - Tout ensemble $(A_i)_{i\in I}$ de parties connexes d'un espace topologique E, ayant deux à deux des intersections non vides, a une réunion A qui est encore connexe.

Démonstration - SI en effet il n'en était pas ainsi, A pourrait être considérée comme réunion de deux parties B' et B'', disjointes, et toutes les deux ouvertes relativement à A. Si alors nous considérons chaque partie A_i, les intersections $A_i \cap B'$ et $A_i \cap B''$ sont ouvertes relativement à A_i, disjointes et de réunion A_i · comme A_i est supposée connexe, l'une de ces deux parties est nécessairement vide, et l'autre est A_i; A_i est toute entière contenue dans B', ou toute entière contenue dans B''. Comme alors l'intersection de deux parties quelconques A_i et A_j est non vide, elles sont nécessairement toutes les deux contenues dans B', ou toutes les deux contenues dans B''; de sorte que, finalement, les parties A_i sont toutes oontenues dans B' ou toutes contenues dans B'' ; donc $A \subset B'$ ou $A \subset B''$. Ceci est contraire à l'hypothèse suivant laquelle A est réunion des parties disjointes non vides B' et B''.
Nous avons donc abouti a une contradiction, et il est ainsi démontré que A est connexe.

Théorème 36 ter - Si E est un espace topologique, A une partie connexe de E, l'adhérence $\bar{A}$ de A dans E est encore connexe.

Démonstration - Supposons en effet qu'il n'en soit pas ainsi, et que $\bar{A}$ soit réunion de deux parties B' et B'' disjointes, non vides, et tnutes les deux fermées relativement à $\bar{A}$. Comme alors $\bar{A}$ est fermée, elles sont aussi fermées relativement à E (théorème 6). Nécessairement A est alors la réunion des deux parties $A\cap B'$ et $A \cap B''$ qui sont disjointes et fermées relativement à A. Commé A est supposée connexe, l'une des deux est vide, et l'on a, par exemple, $A \cap B' = \emptyset$ et $A = A \cap B''$: $A \subset B''$

On en déduit, puisque B est fermée dans E : $\bar{A} \subset B''$, et ceci est contraire à l'hypothèse suivant laquelle $\bar{A}$ est réunion des parties disjointes non vides B' et B''. Nous avons donc bien abouti à une contradiction, et $\bar{A}$ est bien connexe.

Définition - On dit que deux points x et y d'un espace topologique E sont connectés, s'il existe une partie connexe de E contenant à la fois x et y. Deux points qui peuvent être joints par un chemin sont connectes.

Théorème 36 quarto - La relation "x et y sont connectés dans E" est une relation d'équivalence dans E.

Démonstration - Cette relation est évidemment réfléxive et symétrique, il suffit donc de montrer qu'elle est transitive. Or, six et y sont connectés, si d'autre part y et z sont connectés, il existe, d'une part, une partie connexe contenant x et y, et, d autre part, une partie connexe contenant y et z. La réunion de ces deux parties est nécessairement connexe, puisqu'elles sont toutes les deux connexes, et que leur intersection, contenant y, n'est pas vide (théorème 36 bis); par conséquent x et z sont contenus dans une même partie connexe, et, par conséquent, connectés, ce qui démontre le théorème.

Définition - Une classe d'équivalence de E par rapport à la relation d'équivalence "x et y sont connectés dans E" s'appelle une composante connexe de E. E est alors la réunion de ces composantes connexes, qui sont deux à deux disjointes, On appelle composante connexe d'un point x de E la composante connexe qui le contient.

Théorème 36 qulnto - La composante connexe d'un point x de E, identique à l'ensemble des points de E qui sont connectés à x, est la plus grande partie connexe de E contenant x. Les composantes connexes de E sont fermées.

Démonstration - 1°/ Par définition même des classes d'équivalence, la composante connexe E_x de x dans E est l'ensemble des points de E connectés à x. Toute partie connexe de E contenant x est nécessairement toute entière dans E_x, puisque tous ses points sont connectés à x. Inversement, si un point y est dans E_x, il est connecté à x, donc contenu dans au moins une partie connexe de E contenant x. Ainsi E_x est exactement la réunion de toutes les parties connexes de E contenant x.

Comme ces parties connexes ont deux à deux une intersection non vide, puisque toutes contiennent x , il résulte du théorème 36 bis que cette réunion est nécessairement connexe, donc E_x est bien connexe. C'est donc une partie connexe de E , contenant x , et contenant toute partie connexe de E contenant x ; c'est donc bien la plus grande partie connexe de E contenant x.

2°/ D'après le thdorème 36 ter, l'adhérence $\bar{E}_x$ de E_x dans E est encore connexe; comme alors E_x est la plus grande partie connexe de E contenant x , on a nécessairement $\bar{E}_x = E_x$, et par suite E_x est fermée.

Définition - On dit qu'un espace E est localement connexe, si, quel que soit le point a de E et le voisinage9 de a dans E , il existe un voisinage $\mathcal{V}'$ de a contenu dans $\mathcal{V}$, et qui soit connexe. Comme l'indique son nom, le fait pour un espace d'être localement connexe, est une propridté locale, alors que le fait d'être connexe est une propriété globale. Ces deux propriétés n'ont donc aucun rapport l'une avec l'autre * :

1°) Si, par exemple, nous considérons dans le plan $\mathbb{R}^2$, l'ensemble E formé des deux droites d'équations $y = 0$ et $y = 1$, c'est une partie non connexe de R , elle est cependant localement connexe, car, si $\mathcal{V}$ est un voisinage d'un point a de E , ce voisinage contient un Intervalle horizontal de centre a , qui est une partie connexe. Cet exemple montre qu'un espace peut être localement connexe, sans être connexe.

2°) Si nous appelons maintenant E l'ensemble du plan $\mathbb{R}^2$ constitué de toutes les parallèles à l'axe desx d'ordonnées rationnelles, et de l'axe des y tout entier, on voit que cet espace topologique E n'est pas localement connexe : si on considère un point quelconque de E ,de coordonnées a , b, avec $a \neq 0$, et une boule ayant pour centre ce point et un rayon $< |a|$, ce voisinage ne contient aucun voisinage connexe. Par contre E est connexe, et même connexe par arcs , on peut joindre deux quelconques de ses points par un chemin, compose de la succession de trois segments de droite, le premier et le troisième étant des segments parallèles à l'axe des x, et le deuxième étant un segment de l'axe des y . Ainsi cet espace E est connexe sans être localement connexe.

Z

* Ce n'est pas la même chose que pour la compacité : tout espace compact est localement compact (la réciproque n'étant pas vraie), alors qu'un espace connexe n'est pas nécessairement localement connexe.

Théorème 36 sexto - Si E est un espace topologique localement connexe, alors toute composante connexe de E est à la fois ouverte et fermée dans E.

Démonstration - Soit en effet E_x la composante connexe de x dans E. Supposons que y appartienne à E_x ; comme alors y possède un voisinage connexe, tous les points z de ce voisinage sont connectés à y, et par conséquent aussi à x, donc E_x contient tout ce voisinage; ainsi E_x ne peut contenir un point y sans contenir tout un voisinage de y, donc E_x est ouverte; elle est fermée d'après le théorème 36 quinto.

Remarques : 1°/ La réciproque n'est pas exacte. Ainsi, dans l'exemple donné plus haut, où E est connexe sans être localement connexe, il n'y a qu'une composante connexe E ; elle est ouverte et fermée, cependant E n'est pas localement connexe.

2°/ Supposons par exemple que E soit un ouvert d'un espace vectoriel normé E_0 ; alors bien évidemment E est localement connexe. En effet tout voisinage d'un point a de E contient nécessairement une boule de centre a dans E_0, et nous savons qu'une boule est connexe, et même connexe par arcs, On pourra donc dans ce cas, appliquer le théorème. Les composantes connexes de E sont nécessairement à la fois ouvertes et fermées dans E. Elles sont alors aussi ouvertes dans E_0.

3°/ Il existe un autre cas remarquable où toutes les composantes connexes sont à la fois ouvertes et fermées : c est celui où E n'a qu'un nombre fini de composantes connexes. Alors chacune de ces composantes, qui est déjà fermée, est complémentaire de la réunion des autres composantes; or une réunion finie de parties fermées est elle-même fermée; et par conséquent cette composante est nécessairement ouverte.

Définition - On dit qu'un espace topologique est localement connexe par arcs, si, quel que soit le point a et le voisinage $\mathcal{V}$ de a, il existe un autre voisinage $\mathcal{V}'$ de a, contenu dans $\mathcal{V}$, qui soit connexe par arcs. Puisqu'un espace connexe par arcs est connexe, un espace localement connexe par arcs est localement connexe, Un ouvert d'un espace vectoriel normé est localement connexe par arcs, Théorème 36 septimo - soit E un espace topologique localement connexe par arcs. S'il est connexe, il est connexe par arcs; s'il n'est pas connexe, chacune de ses composantes connexes est ouverte, fermee, et connexe par arcs.

Démonstration - Soit x un point de E, et appelons F_x l'ensemble des points de E qui peuvent être joints à x par des chemins. Si y est un tel point, il existe tout un voisinage de y, dont tous les points z peuvent être joints à y par un chemin; mais alors comme y peut être joint à x par un chemin, z peut aussi être joint à x par un chemin; autrement dit F_x ne saurait contenir un point y, sans contenir tout un voisinage de y, F_x est nécessairement ouvert.

Montrons maintenant que F_x est fermé, et, pour cela, montrons que son complémentaire est ouvert. Supposons que y appartienne à ce complémentaire, c'est-à-dire n'appartienne pas a F_x. Il existe un voisinage $\mathcal{V}$ de y, dont tous les points peuvent être joints à y par un chemin; alors aucun point z de $\mathcal{V}$ ne peut appartenir à F_x ; sans quoi on pourrait par un chemin joindre x à z, puis z à y, donc x à y, ce qui est contraire à l'hypothèse. Ainsi le complémentaire de F_x ne peut contenir un point y sans contenir un voisinage de ce point, et par suite il est ouvert, F_x est donc bien fermé.

Si alors E est connexe, F_x est à la fois ouvert et fermé, et non vide puisqu'il contient x, donc $F_x =$ E, et on voit bien que E est connexe par arcs.

De toute façon, même si E n'est pas connexe, soit E_x la composante connexe der dans E ; comme F_x, connexe par arcs, est connexe, on a $F_x \subset E_x$; mais E_x est connexe, et F_x est à la fois ouvert et fermé dans E donc dans E_x, donc $F_x = E_x$. Il en résulte bien que E_x est ouvert et fermé, et connexe par arcs.

Remarques : 1°/ Les 2 théorèmes précédents s'appliqueront en particulier si E est une variété (voir plus loin, chapitre III, § 9). En effet, tout voisinage d'un point de E contient alors un voisinage homéomorphe à une boule, et par conséquent connexe par arcs.

2°/ On pourra démontrer facilement d'autres théorèmes du même genre.

Supposons par exemple que E soit un ouvert d'un espace vectoriel normé. Il est alors localement connexe par lignes polygonales, en ce sens que tout voisinage $\mathcal{V}$ d'un point a contient un autre voisinage $\mathcal{V}'$, (à savoir une boule de centre a), dont deux points quelconques peuvent être joints par une ligne polygonale (c'est-à-dire par un chemin formé de la succession d'un nombre fini de segments de droite : pour une boule, un seul segment suffit). Alors la même démonstration que ci-dessus * montrera que, si E est connexe, il est connexe par lignes polygonales, et que, même s'il n'est pas connexe, toute composante connexe de E est connexe par lignes polygonales.

Si E est un ouvert d'un espace vectoriel normé de dimension finie muni d'un référentiel, on démontrera de même que toute composante connexe de E est connexe 'par lignes polygonales a segments parallèles aux axes", autrement dit que deux quelconques des points d'une composante connexe de E peuvent être jointsparun chemin, formé de la succession d'un nombre fini de segments, parallèles à des axes de coordonnées.

* En considérant l'ensemble F_x des points qu'on peut joindre à x par une ligne polygonale.

Quelques applications de la notion de connexité: critères de non-homéomorphisme

De même que la notion de compacité nous a déjà fourni (pages 70 , 74) des moyens de reconnaître que deux espaces topologiques donnés ne sont pas homéomorphes, la notion de connexité nous donne d'autres exemples; car , si deux espaces sont homéomorphes, et si l'un d'eux est connexe, l'autre l'est aussi. Considérons par exemple la droite réelle $\mathbb{R}$,et la plan $\mathbb{R}^2$. Nous avons vu que ce sont deux ensembles équipotents, donc il existe des bijections de l'un sur l'autre. Il est facile de voir qu'il existe des applications continues surjectives du plan sur la droite : c'est le cas de la projection $(x,y) \longrightarrow x$ de $\mathbb{R}^2$ sur $\mathbb{R}$. PEANO a montré, et c'est beaucoup plus compliqué, qu'il existe aussi des applications continues surjectives de la droite $\mathbb{R}$ sur le plan $\mathbb{R}^2$; mais ces applications, si paradoxal que cela paraisse, ne sont jamais bijectives, et il existe toujours des points du plan qui ont alors plusieurs images réciproques ! On peut montrer très simplement qu'il n'existe pas d'homéomorphisme de la droite sur le plan. La droite possède en effet, d'après le théorème 32, la propriété suivante : si on lui retire un point, la partie complémentaire, en tant qu'espace topologique, n'est pas connexe. Par contre le plan ne possède pas cette propriété: le complémentaire d'un point est manifestement connexe et même connexe par arcs * . Il en résulte bien que la droite et le plan ne peuvent pas être homéomorphes. De la même manière un intervalle $[a, b]$ de la droite $\mathbb{R}$ et une circonférence du plan $\mathbb{R}^2$, bien qu'étant tous les deux compacts, ne sont pas homéomorphes, en effet, si l'on retire à l'intervalle [a,&] son milieu, le complémentaire n'est pas connexe, alors que, si l'on retire de la circonférence un point quelconque, le complémentaire est toujours connexe.

On démontre (mais c'est difficile !) que deux espaces vectoriels normés de dimensions finies différentes ne sont jamais homéomorphes. Une sphère de $\mathbb{R}^m$ et une sphère de $\mathbb{R}^n$, $m \neq n$, ne sont pas homéomorphes.

Existence et Continuité de la fonction réciproque d'une fonction strictement monotone continue

Théorème 37. Si f est une application continue strictement croissante d'un intervalle $|a, b|$ (ouvert, semi-ouvert ou fermé) de $\bar{\mathbb{R}}$ dans $\bar{\mathbb{R}}$ l'image $f(|a, b|)$ est un intervalle de même nature $|\alpha, \beta|$, et f est un homéomorphisme de $|a, b|$ sur $|\alpha, \beta|$. Le prolongement f de f , défini par $f(a) = \alpha$, $f(b) = \beta$, est un homéomorpnisme de $[a, b]$ sur $[\alpha, \beta]$.

* Deux points quelconques peuvent être joints par un chemin formé de la succession de deux segments rectilignes au plus.

Que l'image soit un intervalle (α,β) résulte du corollaire du théorème 33. Que f soit injective résulte de ce qu'elle est strictement croissante; f est donc une bijection continue de $|a,b|$ sur $|\alpha,\beta|$. Comme f conserve la relation d'ordre, et que $a = \operatorname{Inf}(|a,b|)$, $\alpha = \operatorname{Inf}(|\alpha,\beta|)$, on voit que α appartient à l'image si et seulement si l'intervalle $|a,b|$ contient a . De même pour b et β ; ainsi l'intervalle $|\alpha,\beta|$ est de même nature (ouvert, semi-ouvert ou fermé) que $|a,b|$. L'image d'un intervalle $|c,d|$ de $|a,b|$, ouvert dans (a,b), est un intervalle de même nature, donc un ouvert de $|\alpha,\beta|$, alors l'image par f de tout ouvert de $|a,b|$ est un ouvert de $|\alpha,\beta|$, donc f est un homéomorphisme de $|a,b|$ sur $|\alpha,\beta|$ (théorème 11). L'application réciproque f^{-1} , qui est trivialement elle aussi strictement croissante, est donc elle aussi bijective et continue, de $|\alpha,\beta|$ sur $|a,b|$.

Lorsque x tend vers a par valeurs strictement supérieures, $f(x)$ tend vers α (voir page 60). En effet, quel que soit $\alpha' > \alpha$, l'image réciproque de l'intervalle $]\alpha,\alpha'[$ est l'intervalle $]f^{-1}(\alpha), f^{-1}(\alpha')[$, intersection avec $]a,b$ [d'un voisinage de a dans $\overline{\mathbb{R}}$. Donc, si l'on pose $\hat{f}(a) = \alpha$, et de même $\hat{f}(b) = \beta$, $\hat{f}$ est un prolongement de f , qui est une application strictement croissante et continue, donc un homéomorphisme de, $[a,b]$ sur $[\alpha,\beta]$. Si $a = -\infty$, $b = +\infty$, α et β finis, $\hat{f}$ définira un homéomorphisme de $\overline{\mathbb{R}}$ sur un intervalle fermé borné de $\mathbb{R}$.

Remarque. Il était naturel de supposer f strictement monotone : une application f d'un intervalle $|a,b|$ de $\overline{\mathbb{R}}$ dans $\overline{\mathbb{R}}$, continue et injective, est nécessairement strictement monotone. Supposons en effet qu'elle ne le soit pas. Il existerait alors x_1, x_2, x_3 dans $|a,b|$, $x_1 < x_2 < x_3$, tels que $f(x_1) < f(x_2)$ $f(x_2) > (x_3)$, ou tels que $f(x_1) > f(x_2)$, $f(x_2) < f(x_3)$. Prenons par exemple, le premier cas. Soit y un point commun aux intervalles $]f(x_1), f(x_2)[$, $]f(x_3), f(x_2)[$; d'après le théorème des valeurs intermédiaires (corollaire du théorème 33), il existerait un point $x' \in]x_1, x_2[$ et un point $x'' \in]x_2, x_3[$ où f prendrait la valeur y; alors f ne serait pas injective, contrairement à l'hypothèse.

Application : métriques définissant la topologie de $\overline{\mathbb{R}}$.

Théorème 38 $\overline{\mathbb{R}}$ est un espace métrisable; si f est une application strictement croissante, continue et bornée de $\mathbb{R}$ dans $\mathbb{R}$, la métrique ayant pour distance :

$$d(x,y) = |\hat{f}(x) - \hat{f}(y)| , \quad x \in \overline{\mathbb{R}} , \ y \in \overline{\mathbb{R}} , \tag{II,9,1}$$

<u>ou $\hat{f}$ est le prolongement de f en une application de $\overline{\mathbb{R}}$ dans $\overline{\mathbb{R}}$, défini par $\hat{f}(-\infty) = \operatorname{Inf}_{x\in\mathbb{R}} f(x)$, $\hat{f}(+\infty) = \operatorname{Sup}_{x\in\mathbb{R}} f(x)$, définit précisément la topologie de $\overline{\mathbb{R}}$.</u>

D'après le théorème 37, f est un homéomorphisme de $\mathbb{R}$ sur un intervalle $]\alpha,\beta[$ de $\mathbb{R}$; comme f est bornée, $\alpha > -\infty$, $\beta < +\infty$; en outre $\alpha = \operatorname{Inf}_{x\in\mathbb{R}} f(x)$, $\beta = \operatorname{Sup}_{x\in\mathbb{R}} f(x)$, et le prolongement $\hat{f}$ de f défini par $\hat{f}(-\infty)=\alpha$, $\hat{f}(+\infty)=\beta$, est un homéomorphisme de $\overline{\mathbb{R}}$ sur $[\alpha,\beta] \subset \mathbb{R}$.

Comme alors la topologie de $[\alpha,\beta] \subset \mathbb{R}$ est définie par la métrique naturelle de $\mathbb{R}$, celle de $\overline{\mathbb{R}}$ est définie par la <u>métrique transportée</u> de la métrique naturelle de $[\alpha,\beta]$ par $\hat{f}^{-1}$, c'est-à-dire celle dans laquelle la distance de $x \in \overline{\mathbb{R}}$ et $y \in \overline{\mathbb{R}}$ est la distance naturelle de $\hat{f}(x)$ et $\hat{f}(y)$, soit $|\hat{f}(x)-\hat{f}(y)|$ ou (II, 9 ;1), c.q.f.d.

Par exemple, on peut prendre

$$(\mathrm{II},9;2) \qquad f(x) = \operatorname{Arctg} x\,, \qquad \text{avec} \qquad \hat{f}(-\infty) = -\frac{\pi}{2}\,, \quad \hat{f}(+\infty) = +\frac{\pi}{2}\,,$$

ce qui revient à poser $d(x,y) = \left|\int_x^y \frac{d\xi}{1+\xi^2}\right|$. (Plus généralement, si h est une fonction continue >0 sommable sur $\mathbb{R}$, on pourra prendre

$$(\mathrm{II},9;3) \qquad d(x,y) = \left|\int_x^y h(\xi)\,d\xi\right|).$$

On peut prendre

$$(\mathrm{II},9;4) \qquad f(x) = \operatorname{th} x\,, \qquad \text{avec} \qquad \hat{f}(-\infty) = -1\,, \quad \hat{f}(+\infty) = +1\,,$$

On peut prendre

$$(\mathrm{II},9;5) \qquad f(x) = \frac{x}{1+|x|}\,, \qquad \text{avec} \qquad \hat{f}(-\infty) = -1\,, \quad \hat{f}(+\infty) = +1\,.$$

etc.... Toutes ces métriques, très différentes, sont équivalentes sur $\overline{\mathbb{R}}$ (donc aussi sur $\mathbb{R}$) puisqu'elles définissent la même topologie.

§ 11 ESPACES MÉTRIQUES COMPLETS

Soit E un espace métrique et soit : $x_0, x_1, x_2, \ldots x_n, \ldots$ une suite convergente de E, de limite x. Alors cette suite vérifie la propriété appelée <u>critère de CAUCHY</u> :

$$(\mathrm{II},11;1) \qquad (\forall \varepsilon>0)(\exists p\in\mathbb{N})(\forall m\in\mathbb{N}, \forall n\in\mathbb{N}, m \geq p, n\geq p) : d(x_m,x_n) \leq \varepsilon\,,$$

ou encore :

(II,11;2) $d(x_m, x_n)$ tend vers 0 quand m et n tendent vers $+\infty$.

On dit encore que c'est une suite de CAUCHY.

Réciproquement, on a vu en mathématique spéciales que si $x_0, x_1, x_2, \ldots, x_n, \ldots$ est une suite de CAUCHY sur le corps des nombres réels ou des nombres complexes., alors cette suite est convergente. C'est là une propriété fondamentale, car elle permet de reconnaître si une suite est convergente sans avoir besoin de connaitre à l'avance la limite de cette suite, alors que la définition générale de la convergence ne le permet pas. C'est en particulier cette propriété qui aboutit finalement aux différents critères de convergence des séries, et qui permet de reconnaître si une série de nombres réels ou complexes est convergente, sans avoir besoin de connaître à l'avance la somme de la série, ce qui est absolument essentiel en analyse. Il est bien évident que cette propriété, vérifiée sur le corps des nombres réels, n'est pas vraie psur tout espace métrique; c'est précisément cette question qui va faire l'objet du présent paragraphe.

Nous allons d'abord donner quelques propriétés générales des suites de CAUCHY.

Théorème 39 - Toute suite partielle d'une suite de CAUCHY est encore une suite de CAUCHY. Toute suite de CAUCHY est bornée.

Démonstration : La première propriété est évidente. Démontrons la deuxième : Il existe un entier p tel que :

$m \geq p$, $n \geq p$, entraîne $d(x_m, x_n) \leq 1$; alors, pour $n \geq p$, tous les x_n sont contenus dans la boule $B(x_p, 1)$, et par conséquent toute la suite est contenue dans la boule $B(x_p, R)$, où $R = \mathrm{Max}\,[d(x_p, x_0), d(x_p, x_1), \ldots d(x_p, x_{p-1}), 1]$.

Théorème 40 - Si, dans un espace métriaue E . une suite de CAUCHY $x_0, x_1, x_2, \ldots x_n, \ldots$ admet un point d'accumulation a cette suite est convergente vers a.

Démonstration : Etant donné $\varepsilon > 0$, il existe p tel que $m \geq p$, $n \geq p$ entraîne $d(x_m, x_n) \leq \frac{\varepsilon}{2}$; mais il existe une infinité de m, donc au moins une, telles que : $m \geq p$, $d(x_m, a) \leq \frac{\varepsilon}{2}$; on en déduit alors que $n \geq p$ entrafne $d(x_n, a) \leq d(x_n, x_m) + d(x_m, a) \leq \varepsilon$.

Corollaire 1 - Toute suite de CAUCHY, ou bien est convergente, auquel cas elle n a qu'un seul point d'accumulation, ou bien n'a aucun point d accumulation.

Corollaire 2 - Si une suite de CAUCHY est telle qu'une de ses suites partielles est convergente, alors cette suite elle-même est convergente.

En effet 1 hypothése entrafne que la suite donnée admet un point d'accumulation

Définition - On dit qu'un espace métrique E est complet, si toute suite de CAUCHY de E est convergente.

Nous avons vu au début l'utilité de la notion d'espace complet; nous aurons donc besoin de critères permettant de reconnaître qu'un espace métrique donné est complet.

Remarquons tout de suite que la notion d'espace complet n'a aucun sens pour un espace topologique non métrique. D'ailleurs il est facile de voir que deux métriques équivalentes; n'ont pas nécessairement les mêmes suites de CAUCHY et que pour l'une d'entre elles, l'espace peut être complet, sans être nécessairement complet pour l'autre. La notion d'espace complet est une notion métrique et non topologique.

Considérons par exemple la droite réelle $\mathbb{R}$; si nous la munissons de la métrique naturelle, nous verrons plus loin qu'elle est complète (on l'a vu en Mathématiques Spéciales). La suite N des entiers n'est pas une suite de CAUCHY dans cet espace métrique. Si au contraire nous considérons sur la droite achevée $\mathbb{R}$ l'une quelconque des métriques définies page 93, elle induit sur la droite réelle $\mathbb{R}$ une métrique, qui est équivalente à la métrique naturelle. Cependant, pour cette métrique, $\mathbb{R}$ n'est pas complète; si en effet nous considérons la suite des entiers naturels $\mathbb{N}$, elle est une suite de CAUCHY sur $\mathbb{R}$ puisque'elle converge vers $+\infty$, donc elle est une suite de CAUCHY sur $\mathbb{R}$; or elle ne converge pas.

Théorème 41 - Tout espace métrique E , dans lequel toutes les boules fermées sont compactes, est complet. En particulier, tout espace métrique compact est complet, et tout espace vectoriel normé de dimension finie est complet.

Démonstration : Soit $x_0, x_1, x_2, \ldots x_n, \cdots$ une suite de CAUCHY. Cette suite est bornée d'après le théorème 39, donc contenue dans une boule fermée convenable B , c'est-à-dire dans un compact. Alors, d'après le théorème de WEIERSTRASS-BOLZANO, elle Possède au moins un point d'accumulation; donc, d'après le théorème 40, elle est convergente, et E est bien complet.

Remarques 1°) Les hypothèses du théoréme Impliquent que E soit localement compact. Mais un espace métrique localement compact n'a pas nécessairement toutes ses boules fermées compactes * . Par exemple : si nous considérons $\mathbb{R}$ muni de l'une quelconque des métriques définies page 93, le sous-

* Le fait, pour un espace métrique E, d'être localement compact, est une propriété topologique; le fait d'avoir ses boules fermées compactes est une propriété métrique.

espace R muni de cette métrique est localement compact puisque c'est $\mathbb{R}$ muni de sa topologie naturelle; cependant comme nous l'avons vu il n'est pas complet. Il est facile de vérifier en effet que toutes ses boules fermées ne sont pas compactes. Car pour ρ suffisamment grand, une boule de rayon ρ dans $\bar{\mathbb{R}}$ est $\bar{\mathbb{R}}$ tout entière, puisque $\bar{\mathbb{R}}$, compacte, est bornée; alors une boule de rayon ρ dans $\mathbb{R}$ est $\mathbb{R}$ tout entlere, qui n'est pas compacte.

2") Les espaces vectoriels normés de dimension infinie peuvent être complets ou non complets, nous donnerons ultérieurement des exemples de ces deux cas. Naturellement les cas intéressants sont ceux où l'espace est complet.

3") Il existe des espaces métriques non localement compacts qui sont quand même complets, c'est précisément ce qu'indique la remarque 2°).

4°) De même que la propriété de compacité, le fait pour un espace métrique d'être complet est une propriété relative à l'espace lui-même. Cependant, si F est une partie d'un espace métrique E , on dira que F est complet si, en tant qu'espace métrique muni de la métrique induite,11 est complet.

<u>Théorème 42</u> - <u>Soient E un espace métrique, F une partie de E. SI F est complet, il est fermé dans E</u>.

<u>Démonstration</u> Soit a un point de E adhérent à F . D'après le théorème 15, il existe une suite cc,, $x_1, x_2, \dots x_n, \dots$ de F, qui converge vers a ; c'est donc une suite de CAUCHY dans E , donc dans F . Comme F est complet, elle a une limite a' dans F, donc dans E , donc a = a' , ce qui prouve que a appartient à F et que F est bien fermé.

<u>Corollaire 1</u> - <u>Si E est un espace métrique, F une partie dense de E, distincte de E , F n'est sûrement pas complète.</u>

En effet l'adhérence de F est E $\neq$ F , donc F n'est pas fermée. Ceci prouve-par exemple que le corps $\mathbb{Q}$ des nombres rationnels, muni de sa métrique naturelle, n'est pas complet. On forme d'ailleurs très simplement une suite de CAUCHY de $\mathbb{Q}$ qui n'est pas convergente; il suffit de prendre une suite de $\mathbb{Q}$ qui converge dans R vers un nombre irrationnel. Ce qui fait précisément la nécessité d'introduire le corps des réels et de ne pas se contenter du corps des rationnels, ce sont les deux propriétés possédées par les réels et non possédées par les rationnels, à savoir d'une part que toute partie majorée non vide admet une borne supérieure, et d'autre part que toute suite de CAUCHY est convergente.

<u>Corollaire 2</u> - <u>Si E est un espace vectoriel normé, F un sous-espace vectoriel de dimension finie, F est ferme dans E.</u>

En effet, d'après ce qui a été dit avant le théorème 41, F est complet.

On démontre que ce résultat subsiste si E est seulement un espace vectoriel topologique.

Naturellement la réciproque du théorème précédent est inexacte. Par exemple $F = E$ est toujours fermé dans E, et il n'est pas nécessairement complet. Mais :

Théorème 43 - Si E est un espace métrique complet, toute partie fermée F de E est elle aussi complète.

Démonstration : Soit $x_0, x_1, x_2, \ldots x_n, \ldots$ une suite de CAUCHY de F • c'est aussi une suite de CAUCHY de E, et comme E est supposé complet, elle converge vers un point a de E. Mais comme tous les x_n sont dans F, a est nécessairement adhérent à F (théorème 15), donc dans F supposé fermé, et la suite de CAUCHY converge vers un élément de F, qui par conséquent est bien complet.

L'ensemble des deux théorèmes 42 et 43 montre que, si E est complet, il y a identité entre les parties complètes et les parties fermées de E. On notera la ressemblance entre ces propriétés et les propriétés correspondantes des ensembles compacts.

Théorème 44 - Soient E_1 et E_2 deux espaces métriques complets. Alors le produit $E_1 \times E_2$ pour l'une quelconque des métriques définies page 63, est lui aussi complet.

Démonstration plus généralement, appelons d_1 et d_2 les distances sur E_1 et E_2, et soit δ une distance sur le produit $E_1 \times E_2$, possédant les deux propriétés suivantes :

1°/ elle définit sur $E_1 \times E_2$ la topologie produit des topologies définies par d_1 et d_2 sur E_1, et E_2.

2°/ Il existe un nombre k tel que, pour tout couple $(x_1, x_2), (y_1, y_2)$ d'éléments de $E_1 \times E_2$, on ait les inégalités :

$$k\,\delta\big((x_1, x_2), (y_1, y_2)\big) \geq d_1(x_1, y_1) \text{ et } \geq d_2(x_2, y_2) . \quad (\text{II}, 11; 3)$$

Nous allons démontrer que, pour une telle métrique, $E_1 \times E_2$ est complet. Soit en effet $((x_1)_n, (x_2)_n)_{n \in N}$ une suite de CAUCHY de ce produit. D'après l'hypothèse 2°) relative a la métrique, la suite des $(x_1)_n$ est une suite de CAUCHY sur E_1, et la suite des $(x_2)_n$ est une suite de CAUCHY sur E_2. Comme ces espaces sont complets, ces suites convergent respectivement vers des éléments a_1 de E_1 et a_2 de E_2 ; mais alors, pour la topologie produit, c'est-à-dire pour la métrique δ considérée, la suite des $((x_1)_n, (x_2)_n)$ converge vers (a_1, a_2) dans $E_1 \times E_2$, qui est bien complet.

Prolongement des applications uniformément continues

Théorème 45 - Soient E et F des espaces métriques, E_1 un sous-espace dense de E, f_1 une application de E_1 dans F ; on suppose f_1 uniformément continue sur E_1 et F complet.

Alors il existe une application f et une seule de E dans F, qui soit continue et qui prolonge f_1 ; en outre, cette application est uniformément continue.

Démonstration 1°) Même si f_1 est seulement continue et si F n'est pas complet, il ne peut pas exister plus d'une application f ayant les propriétés indiquées. Supposons en effet qu'il en existe une. Soit x un point de E ; comme E_1 est dense, il existe, d'après le théorème 15, une suite $x_0, x_1, x_2, \ldots x_n, \ldots$ de E_1 convergeant vers x ; alors, si f est continue, d'après le théorème 16, la suite des $f(x_n) = f_1(x_n)$ converge vers $f(x)$ dans F. Ainsi $f(x)$ est entièrement connu, puisqu'il est la limite des $f_1(x_n)$, ce qui prouve bien l'unicité de f.

2°) Pour montrer l'existence de f, nous sommes obligés de supposer plus, nous avons supposé la continuité uniforme de f_1, et nous avons supposé F complet. Soit alors de nouveau x un point de E, choisissons une suite $x_0, x_1, x_2, \ldots x_n, \ldots$ de E_1, convergeant vers x. Alors c'est une suite de CAUCHY dans E ; il en résulte immédiatement que la suite des $f_1(x_n)$ est une suite de CAUCHY dans F. En effet, si $\varepsilon > 0$ est donné, il existe, d'après l'hypothèse de continuité uniforme, un nombre $\eta > 0$ tel que $x' \in E_1$, $x'' \in E_1$, $d(x', x'') \leq \eta$ entraîne $d\left(f_1(x'), f_1(x'')\right) \leq \varepsilon$. Comme $x_0, x_1, x_2, \ldots x_n, \ldots$ est une suite de CAUCHY, il existe un entier p tel que $m \geq p$, $n \geq p$, entraîne $d(x_m, x_n) \leq \eta$, et par suite $d\left(f_1(x_m), f_1(x_n)\right) \leq \varepsilon$, ce qui prouve bien ce que nous affirmions.

Comme alors F est supposé complet, la suite des $f_1(x_n)$ admet une limite dans F, appelons la $f(x)$. Il faut d'abord montrer que $f(x)$ est bien déterminé quand x est connu, c'est-à-dire, ne dépend pas du choix de la suite des x_n. Or, si nous considérons deux quelconques de ces suites : x'_n et x''_n, toutes les deux convergentes vers x, la suite "mélangée", c'est-à-dire : $x'_0, x''_0, x'_1, x''_1, x'_2, x''_2, \ldots x'_n, x''_n, \ldots$, est encore une suite de E_1 convergeant vers x, alors la suite : $f(x'_0), f(x''_0), f(x'_1), f(x''_1), \ldots f(x'_n), f(x''_n), \ldots$ doit converger vers un élément de F, ce qui prouve bien que les deux suites $f(x'_n)$, $f(x''_n)$, ont la même limite dans F.

Ayant démontré cela, nous venons bien de définir une application f de E dans F. Cette application prolonge trivialement f_1, car, si x appartient à E_1, nous pourrons considérer la suite $x, x, x, \ldots x, \ldots$ convergeant vers x , l'image $f(x)$ est alors la limite de la suite $f_1(x), f_1(x), f_1(x), \dots f_1(x), \dots$ c'est-à-dire $f_1(x)$. Il nous reste donc à prouver la continuité uniforme de l'application f .

Soit alors $\varepsilon > 0$ un nombre donné, η le nombre qui lui est associé comme il est dit ci-dessus, par la continuité uniforme de f_1 . Soient x et y des points quelconques de E tels que $d(x,y) \leqslant \frac{\eta}{2}$. Nous allons démontrer que l'on a $d(f(x), f(y)) \leqslant \varepsilon$, ce qui prouvera bien la continuité uniforme. Soient x_n et y_n des suites de E_1, convergeant respectivement vers x et y . Alors, d'après la continuité de la fonction distance (théorème 17 bis), il existe un entier p tel que, pour $n \geqslant p$, on ait $d(x_n, y_n) \leqslant d(x,y) + \frac{\eta}{2}$. On en déduira, pour $n \geqslant p$, $d(x_n, y_n) \leqslant \eta$, et par suite $d(f(x_n), f_1(y_n)) \leqslant \varepsilon$ .Comme, dans F aussi, la fonction distance est continue, et que les suites $f_1(x_n)$, $f_1(y_n)$, sont convergentes respectivement vers $f(x)$ et $f(y)$, alors $d(f_1(x_n), f_1(y_n))$ converge vers $d(f(x), f(y))$; on a donc aussi $d(f(x), f(y)) \leqslant \varepsilon$, et ceci achève la démonstration du théorème.

Remarque 1°) Il est essentiel de supposer l'espace F complet. Si par exemple nous prenons $F = E_1$ lui-même, et pour f_1 l'application identique de E_1 , elle ne peut pas se prolonger en une application continue f de E dans E_1.

Soit en effet x un point de E n'appartenant pas à E_1 . Soit x_n une suite de E_1 convergeant vers x . Alors, s'il existait un tel prolongement f la suite des $f_1(x_n) = x_n$ convergerait vers $f(x)$ dans E_1 , donc dans E ; donc on aurait $f(x) = x$, ce qui est absurde, puisque $x \notin E_1$.

2°) Il n'est pas non plus suffisant de supposer f_1 continue. Prenons par exemple $E = \overline{\mathbb{R}}$, muni de l une quelconque des métriques indiquées au théorème 38, et $E_1 = \mathbb{R}$; prenons pour F , la droite $\mathbb{R}$ munie de sa métrique naturelle; alors F est complet, et on voit facilement que l'application Identique f_1 de E_1 dans F est continue, mais n'est pas uniformément continue. Or elle ne peut manifestement pas se prolonger en une application continue f de $E = \overline{\mathbb{R}}$ dans $F = \mathbb{R}$. Si en effet nous considérons le point $x = +\infty$ de $E = \overline{\mathbb{R}}$, il est limite dans E de la suite des entiers naturels $\mathbb{N}$, mais l'image de cette suite, c'est-à-dire $\mathbb{N}$ n'est pas convergente dans $F = \mathbb{R}$, alors qu'elle devrait converger vers $f(+\infty)$. si f existait.

Propriétés particulières aux espaces vectoriels topologiques de dimension finie

Nous avons déjà vu de telles propriétés : le théorème 13, le théorème 23, le corollaire 2 du théorème 41. Il nous reste à démontrer la propriété annoncée page 74 :

<ins>Théorème 45 bis</ins> - (Frédéric RIESZ). <ins>Pour qu'un espace vectoriel topologique soit localement compact, il faut et il suffit qu'il soit de dimension finie.</ins>

Démonstration -

Pour simplifier, nous supposerons que E est un espace vectoriel normé; des modifications infimes donneraient le **résultat** général. Nous savons déju que s'il est de dimension finie, il est localement compact; c'est la réciproque qu'il nous faut montrer, Supposons donc que E ait un voisinage compact $\mathcal{V}$ de 0. Nous emploierons les natations suivantes : si A est une partie de E, λA, λ scalaire, est l'ensemble des $\lambda x, x \in A$; si A et B sont deux parties de E, $A + B$ est l'ensemble des $x + y, x \in A, y \in B$ (attention, $2A$ est contenu dans $A + A$, mals en général distinct). Un sous-espace vectoriel M de E est alors caractérise par $M + M = M$, $\lambda M \subset M$ pour tout λ. Alors $2\mathcal{V}$ est encore un voisinage de 0; et, pour tout $a \in E$, $a + \mathcal{V}$ est un voisinage de a; $a + \mathcal{V}$ est un voisinage ouvert de a. Lorsque a varie dans $2\mathcal{V}$, les $a + \mathcal{V}$ forment un recouvrement ouvert de $2\mathcal{V}$; d'après l'hypothèse de compacité, il en existe un nombre fini $a_1, a_2, \ldots, a_n$, tels que les $a_i + \mathcal{V}$ recouvrent $2\mathcal{V}$. Soit M le sous-espace vectoriel engendré par les a_i ; il est de dimension finie, et $M + \mathcal{V}$ recouvre $2\mathcal{V}$.

Alors $M + \mathcal{V} = M + M + \mathcal{V} \supset M + 2\mathcal{V}$. En multipliant par 2, $M + \mathcal{V} \supset M + 2\mathcal{V} = 2M + 2\mathcal{V} = 2(M + \mathcal{V}) \supset 2(M + 2\mathcal{V}) = 2M + 4\mathcal{V} = M + 4\mathcal{V}$. Et ainsi de suite : pour tout n, $M + \mathcal{V} \supset M + 2^n\mathcal{V} \supset 2^n\mathcal{V}$. Mals la réunion des $2^n\mathcal{V}$ est l'espace entier, donc $M + \mathcal{V}$ est l'espace entier E. Nous allons en déduire que M est déjà l'espace entier E, qui sera donc bien de dimension finie. SI ce n était pas vrai, il existerait un point $a \notin M$. Mals M est fermé (corollaire 2 du théorème 41). Donc il existerait une boule de centre a qui ne rencontrerait pas M ; cette boule pourait s'écrire $a + B$, B boule de centre origine. Alors $M + B$ ne contiendrait pas a. Pour $\rho > 0$ quelconque, $M + \rho B = \rho M + \rho B$ ne contiendrait pas ρa : on aurait $\rho a \notin M + \rho B$. Mais $\mathcal{V}$, compact donc borné, est contenu dans ρB pour un ρ assez grand; alors on aurait $\rho a \notin M + \mathcal{V}$, ce qui serait contraire au résultat antérieurement obtenu $M + \mathcal{V} = E$.

§ 12 THÉORÈME DU POINT FIXE

Définition : Soit E un espace métrique, f une application de E dans E. On dit que f est une contraction, s'il existe une constante positive $k < 1$, telle que l'on ait, pour tout couple d'éléments x, y de E, l'inégalité

(II,12;1) $$d(f(x), f(y)) \leqslant k\, d(x,y) .$$

Cela entraîne évidemment que f soit **lipschitzienne**, et par conséquent uniformément continue. On dit que a est un point fixe pour une application f si l'on a $f(a) = a$.

Théorème **46** - Toute contraction d'un espace métrique complet E dans lui-marne admet un point fixe et un seul.

<u>Démonstration</u> - L 'unicité du point fixe est évidente, même si E n'est pas complet. Si en effet a et b sont deux points fixes, on doit avoir

(II,12;2) $$d(a,b) \leqslant k\, d(a,b) < d(a,b) \quad \text{si} \quad d(a,b) \neq 0 .$$

On a donc nécessairement $d(a,b)=0$, a et b sont confondus.

Démontrons donc l'existence du point fixe. On va utiliser la méthode dite des <u>approximations successives.</u>

Soit x_0 un point quelconque de E, posons :

(II,12;3) $$x_1 = f(x_0) , \quad x_2 = f(x_1), \dots \quad x_n = f(x_{n-1}), \dots$$

Nous formons ainsi une suite Infinie $x_0, x_1, x_2, \dots x_n, \dots$ d'éléments de E . Nous allons montrer que c'est une suite de CAUCHY. Comme f est une contraction, on a la suite d'inégalités :

(II,12;4) $$\begin{cases} d(x_2,x_1) \leqslant k\, d(x_1, x_0) \\ d(x_3,x_2) \leqslant k\, d(x_2, x_1) \leqslant k^2 d(x_1,x_0) \\ \dots \\ d(x_{n+1},x_n) \leqslant k^n d(x_1,x_0) . \end{cases}$$ Alors

(II,12,5) $$\begin{aligned} d(x_{n+p}, x_n) &\leqslant d(x_{n+p}, x_{n+p-1}) + d(x_{n+p-1}, x_{n+p-2}) + \dots + d(x_{n+1}, x_n) \\ &\leqslant (k^{p-1} + k^{p-2} + \dots + k + 1) k^n d(x_1,x_0) \leqslant \frac{k^n}{1-k} d(x_1,x_0) . \end{aligned}$$

On en déduit bien que $d(x_{n+p}, x_n)$ tend vers 0 quand n tend vers $+\infty$, donc que la suite des x_n est une suite de CAUCHY; par suite, elle admet une limite a , puisque E est suppose oomplet. Comme alors x_n tend versa, on voit que $x_{n+1} = f(x_n)$ tend vers $f(a)$ d'après la continuité de f , et, comme x_{n+1} tend aussi versa , on a bien $f(a) = a$, et a est un point fixe. Le procédé précédent donne non seulement l'existence du point fixe, mals une méthode pratique pour le trouver.

Remarquons que la suite des x_n est rapidement convergente. On a en effet :

(II,12;6) $$d(x_n, a) = \lim_{p\to\infty} d(x_n, x_{n+p}) \leqslant \frac{k^n}{1-k} d(x_1,x_0) .$$

<u>Remarques</u> 1°) L'hypothèse $k < 1$ est absolument indispensable, la condition $k \leqslant 1$ n'est pas suffisante pour entrainer ni l'existence ni l'unicité du point fixe.

L'application identique d'un espace métrique dans luimême vérifie toujours 1 inégalité (II,12;1) avec $k = 1$, et tous les points sont des points fixes. Par ailleurs, sur la droite, une translation $x \longrightarrow x+1$ vérifie aussi l'inégalité (II,12;1) avec $k = 1$, mais ne possède aucun point fixe.

2°) Soit f est une application de E dans E qui n'est pas nécessairement une contraction. Si l'une de ses itérées f_p est une contraction, alors l'application f a encore un point fixe et un seul.

Rappelons que les applications itérées d'une application f d'un ensemble E dans lui-même, sont définies par la formule :

$$(II,12;7) \quad f_2 = f \circ f \ , \quad f_3 = f \circ f \circ f \ , \ldots \quad f_p = f \circ f_{p-1} = f_{p-1} \circ f$$

Soit alors a un point fixe de f . Il est aussi un point fixe pour l'une quelconque de ses itérées, mais, l'une d'elles, f_p, est une contraction, elle ne possède pas plus d'un point fixe ce qui prouve que f elle-même n'a pas plus d'un point fixe. Pour montrer l'existence, supposons réciproquement que a soit le point fixe unique de f_p , alors l'élément $f_{p+1}(a)$ peut s'écrire de deux manières, soit $f(f_p(a)) = f(a)$, soit $f_p(f(a))$, ce qui prouve que $f(a)$ est un autre point fixe pour f_p . Comme f_p est une contraction, elle n'a qu'un point fixe, on a nécessairement $f(a) = a$, et a est un point fixe de f . Ainsi f a bien le point fixe unique a . En outre on peut toujours l'obtenir par les mêmes approximations successives : la suite $x_0, x_1, x_2, \ldots x_n, \ldots$ définie par (II,12;3) converge vers a . En effet chacune des suites partielles

$$x_0, x_p, x_{2p}, \ldots x_{np}, \ldots \ ; \quad x_1, x_{p+1}, x_{2p+1}, \ldots x_{np+1}, \ldots ; \ldots ; x_{p-1}, x_{2p-1}, x_{3p-1}, \ldots x_{np+p-1}, \ldots,$$

converge vers a , puisque c'est une suite d'approximations successives pour l'itérée f_p , ce qui signifie exactement que la suite toute entière $x_0, x_1, x_2, \cdots x_n, \cdots$ converge vers a. (Remarquons que l'hypothèse que l'application itérée f_p soit une contraction, n'implique pas nécessairement que f soit continue).

Supposons maintenant que la contraction f dépende de façon suffisamment régulière d'un paramètre λ . Alors, pour toute valeur de λ , elle possède un point fixe et un seul a_λ. Nous nous proposons de chercher si le point fixe a_λ dépend d'une manière continue du paramètre λ .

Théorème 46 bis - Soit E un espace métrique complet, Λ un espace topologique, f une application de $E \times \Lambda$ dans E. Supposons que, pour tout x fixé dans E , l'application partielle $\lambda \longrightarrow f(x, \lambda)$ soit continue de Λ dans E , et que, pour tout $\lambda \in \Lambda$, l'application $f_\lambda : x \longrightarrow f(x, \lambda)$, soit une contraction de E dans E , correspondant à un nombre $k < 1$ (formule (II,12;1)) indépendant de λ . Alors, si l'on appelle a_λ l'unique point fixe de f_λ , il dépend continuement du paramètre λ , autrement dit l'application $\lambda \rightarrow a_\lambda$ de Λ dans E est continue

Démonstration : Appelons λ_0 un point de Λ. Soit $\varepsilon > 0$ donné. On a alors les inégalités :

$$(\text{II},12;8) \quad d(a_\lambda, a_{\lambda_0}) = d(f_\lambda(a_\lambda), f_{\lambda_0}(a_{\lambda_0})) \leq d(f_\lambda(a_\lambda), f_\lambda(a_{\lambda_0})) + d(f_\lambda(a_{\lambda_0}), f_{\lambda_0}(a_{\lambda_0}))$$

$$\leq k\, d(a_\lambda, a_{\lambda_0}) + d(f(a_{\lambda_0}, \lambda), f(a_{\lambda_0}, \lambda_0))$$

d'où
$$(1-k)\, d(a_\lambda, a_{\lambda_0}) \leq d(f(a_{\lambda_0}, \lambda), f(a_{\lambda_0}, \lambda_0)).$$

D'aprks la continuité partielle de f par rapport à λ pour x fixé en a_{λ_0}, il existe un voisinage $\mathcal{V}$ de λ_0 dans Λ tel que $\lambda \in \mathcal{V}$ entraîne : $d(f(a_{\lambda_0}, \lambda), f(a_{\lambda_0}, \lambda_0)) \leq \varepsilon(1-k)$. Alors $\lambda \in \mathcal{V}$ entrainera aussi $d(a_\lambda, a_{\lambda_0}) \leq \varepsilon$, ce qui prouve bien la continuité de l'application considérée, au point λ_0 de Λ.

§ 13 THÉORIE ÉLÉMENTAIRE DES ESPACES VECTORIELS NORMÉS ET DES ESPACES DE BANACH

Soient E et F deux espaces vectoriels normés sur le même corps $\mathbb{K}$, que nous supposerons toujours être ou le corps des réels ou le corps des complexes * . Soit u une application de E dans F. Rappelons que l'on dit que u est linéaire, si l'on a :

$$(\text{II},13;1) \quad u(\vec{x}+\vec{y}) = u(\vec{x}) + u(\vec{y}) \;;\; u(\lambda\vec{x}) = \lambda u(\vec{x})\;;\; \lambda \in K,\ \vec{x} \in E,\ \vec{y} \in E.$$

<u>Si E est de dimension finie, une application linéaire de E dans F est nécessairement continue, et même lipschitzienne, donc uniformément continue.</u> En effet, comme toutes les normes sur E sont alors équivalentes (théorème 13), il suffira, pour la continuité aussi bien que pour la continuité uniforme ** de supposer que E est muni d'une base $\vec{e}_1, \vec{e}_2, \ldots \vec{e}_n$, , et que, pour tout vecteur $\vec{x}$ de coordonnées $x_1, x_2, \ldots x_n$, la norme est définie par $\|\vec{x}\| = \operatorname{Max}_{i=1,2,\ldots n} |x_i|$

* Tout espace vectoriel sur le corps des complexes'est a fortiori un espace vectoriel sur le corps des réels. Si alors, E et F sont des espaces vectoriels, l'un sur les réels, l'autre sur les complexes, on les considérera tous deux comme espaces vectoriels sur les réels.

On convient de noter par le même symbole $\vec{0}$ les éléments neutres (généralement distincts) de E et F . On a alors $u(\vec{0}) = \vec{0}$. On note aussi généralement par le même symbole $\| \|$ les normes dans E et dans F .

** La continuité est une propriété topologique et non métrique. Il n'en est pas de même de la continuité uniforme; mais le théorème 12 assure que, pour la continuité uniforme d'une application d'un espace <u>vectoriel normé</u> dans un autre, on peut remplacer les <u>normes</u> par des <u>normes équivalentes</u>.

On a alors :

$$\|u(\vec{x}') - u(\vec{x}'')\| = \|u(\vec{x}' - \vec{x}'')\| = \left\|u\left(\sum_{i=1}^{n}(x_i' - x_i'')\vec{e}_i\right)\right\|$$

$$= \left\|\sum_{i=1}^{n}(x_i' - x_i'')u(\vec{e}_i)\right\| \leqslant \left(\sum_{i=1}^{n}\|u(\vec{e}_i)\|\right)\operatorname{Max}_{i=1,2,\dots n}|x_i' - x_i''| \leqslant k\|\vec{x}' - \vec{x}''\|$$

ce qui prouve bien notre affirmation.

Si au contraire E est de dimension infinie, il n'en est plus du tout ainsi; si paradoxal que cela paraisse à priori, il existe des applications linéaires discontinues.

Donnons un exemple : Prenons pour E l'espace vectoriel, sur le corps des réels, des polynômes à coefficients réels * . Prenons la norme suivante :

(II,13; 1) $$\|P\| = \operatorname{Max}_{0 \leqslant x \leqslant 1}|P(x)|.$$

Il s'agit bien d'une norme, car toutes les inégalités voulues sont trivialement vérifiées (l'existence du maximum considéré résulte de la continuité de P, et du théorème 29; on a trivialement $\|\lambda P\| = |\lambda|\,\|P\|$, si λ est un nombre réel, et $\|P+Q\| \leqslant \|P\| + \|Q\|$, si P et Q sont 2 polynomes; d'autre part le fait que pour $P \not\equiv 0$, on ait : $\|P\| > 0$ résulte de ce que, si $\|P\| = 0$, alors le polynôme P , nul sur l'intervalle $[0,1]$, est identiquement nul.)

Prenons pour u la fonction réelle sur E , qui, à chaque polynôme P , fait correspondre sa valeur au point $x = 3$: $u(P) = P(3)$.

C'est manifestement une forme linéaire sur E ; montrons que cette forme linéaire est discontinue. Pour cela il nous suffit de considérer la suite de polynômes définie par

(II,13;2) $$P_n(x) = \left(\frac{x}{2}\right)^n.$$

On a bien évidemment $\|P_n\| = \frac{1}{2^n}$, donc cette suite converge vers $\vec{0}$ dans E ; cependant : $P_n(3) = \left(\frac{3}{2}\right)^n$, donc la suite des valeurs u (P_n) tend vers $+\infty$, ce qui prouve bien la discontinuité de u .

Théorème 47 - Toute application linéaire d'un espace vectoriel normé E dans un espace vectoriel normé F , continue à l'origine, est continue partout; elle est même lipschitzienne, donc uniformément continue. Pour qu'il en soit ainsi, il faut et il suffit qu'il existe une constante $k \geqslant 0$ telle que l'on ait :

(II,13;3) $$\|u(\vec{x})\| \leqslant k\,\|\vec{x}\|, \text{ pour tout } \vec{x} \text{ de } E .$$

* L'addition et la multiplication par les réels sont l'addition usuelle des polynomes et leur multiplication usuelle par les réels; E est bien un espace vectoriel, dont l'élément $\vec{0}$ est le polynôme $\equiv 0$. L'espace vectoriel des polynômes de degré $\leqslant m$ a la dimension $m+1$; une de ses bases est constituée par les polynômes $1, x, x^2, \dots x^m$ L'espace vectoriel de tous les polynômes a donc une dimension infinie.

Démonstration : Démontrons d'abord que, si u est continue à l'origine, il existe un nombre k tel que l'on ait (II,13;3) La continuité deu à l'origine entraîne l'existence d'un nombre $\eta > 0$ tel que $\|\vec{x}\| \leq \eta$ entraîne $\|\vec{u}(x)\| \leq 1$.

Alors, par une homothétie de rapport λ, u-étant linéaire, $\|\vec{x}\| \leq \lambda\eta$ entrafne $\|u(\vec{x})\| \leq \lambda$. Mais la première de ces deux inégalités est toujours vérifiée si l'on prend $\lambda = \frac{\|\vec{x}\|}{\eta}$, on en déduit que l'on a toujours l'inégalité $\|u(\vec{x})\| \leq \frac{1}{\eta}\|\vec{x}\|$, ce qui est bien l'inégalité cherchée si l'on prend $k = \frac{1}{\eta}$. Inversement, supposons qu'il existe un nombre k tel que l'on ait l'inégalité (II,13;3). Alors non seulement u est continue à l'origine, mais partout continue, et lipschitzienne, car en vertu de la linéarité de u, on a :

(II,13;4) $$\|u(\vec{x}') - u(\vec{x}'')\| = \|u(\vec{x}'-\vec{x}'')\| \leq k\|\vec{x}'-\vec{x}''\| .$$

Remarque - Ce théorème contient le théorème 12 comme cas particulier. En effet, écrire que 2 normes sur un espace vectoriel E sont équivalentes, c'est écrire que l'application identique de E, muni de l'une quelconque de ces deux normes, sur E, muni de l'autre norme, est continue; alors (II,13;3) donne (II, 4 ; 1).

Noyau et image d'une application linéaire continue

Soient E, F, des espaces vectoriels sur un corps K, u une application linéaire de E dans F.

On appelle noyau de u l'ensemble des $\vec{x}$ de E tels que $u(\vec{x}) = \vec{0}$. C'est encore l'image réciproque ü' $(\{\vec{0}\})$. C'est un sous-espace vectoriel de E ; car, si $\vec{x}$ et $\vec{y}$ sont dans le noyau, on a $u(\vec{x})=\vec{0}$, $u(\vec{y})=\vec{0}$, donc $u(\vec{x}+\vec{y})=u(\vec{x})+u(\vec{y})=\vec{0}$,

donc $\vec{x}+\vec{y}$ est dans le noyau; et, si λ est un scalaire, on a $u(\lambda\vec{x}) = \lambda u(\vec{x}) = \vec{0}$, donc $\lambda\vec{x}$ est dans le noyau.

On appelle image de u l'image $u(E)$ de E par u. C'est un sous-espace vectoriel de F. La dimension de cette image ' ppelle le rang deu Rappelons que, si E et F sont de dimension finie et si l'on y a choisi des bases, u s'exprime par une matrice'; le rang de u est alors le rang le plus élevé des mineurs $\neq$ o de cette matrice. Rappelons que le rang est aussi la différence de la dimension de E et de la dimension du noyau de u. (Voir cours d'Algèbre, Chapitre II).

Si E et F sont normés, et u continue, le noyau est un sous-espace vectoriel fermé car c'est l'image réciproque d'une partie reduite à un point, donc fermée, de F (théorime 8); l'image n'est pas nécessairement fermée * .

Définition : La borne inférieure des nombres k tels que l'on ait l'inégalité (II,13;3) s'appelle la norme de l'application linéaire u ** ; cette norme se note : $\|u\|$ et se définit donc encore par les relations suivantes :

(II,13;5) $$\|u\| = \sup_{\vec{x} \neq \vec{0}} \frac{\|u(\vec{x})\|}{\|\vec{x}\|} = \sup_{\|\vec{x}\| \leq 1, \vec{x} \neq \vec{0}} \|u(\vec{x})\| = \sup_{\|\vec{x}\| = 1} \|u(\vec{x})\| .$$

Notons en particulier que l'on a toujours :

(II,13;6) $$\|u(\vec{x})\| \leq \|u\| \; \|\vec{x}\| .$$

Théorème 48 - L'ensemble $\mathcal{L}(E;F)$ des applications linéaires continues d'un espace vectoriel normé E dans un espace vectoriel normé F, admet lui-même une structure d'espace vectoriel normé, avec la norme définie par (II,13;5) . En outre, si G est un troisième espace vectoriel normé, si

$u \in \mathcal{L}(E;F)$ et si $v \in \mathcal{L}(F;G)$, on sait

que $v \circ u \in \mathcal{L}(E;G)$, et l'on a l'inégalité :

(II,13;7) $$\|v \circ u\| \leq \|v\| \; \|u\| .$$

Démonstration Montrons d'abord que $\mathcal{L}(E;F)$ possède une structure d'espace vectoriel sur $\mathbb{K}$. Soient u_1 et u_2 deux applications linéaires continues de E dans F , nous définirons $u_1 + u_2$ par la relation :

(II,13;8) $$(u_1 + u_2)(\vec{x}) = u_1(\vec{x}) + u_2(\vec{x}) .$$

* L'image directe d'une partie fermée par une application continue n'a pas de raison d'être fermée !

** On peut aussi la définir si u est une application linéaire discontinue, alors $\|u\| = +\infty$.

Noter qu'en fait cette borne inférieure est un minimum.

On vérifie que nous venons bien de définir là une nouvelle application linéaire de **E** dans F ; d'autre part l'inégalité (déduite de II,13;6) :

(II,13;9) $$\|(u_1+u_2)(\vec{x})\| \leqslant \|u_1(\vec{x})\| + \|u_2(\vec{x})\| \leqslant \left(\|u_1\|+\|u_2\|\right)\|\vec{x}\|,$$

montre qu'elle est continue, autrement dit que $u_1+u_2 \in \mathcal{L}(E;F)$, et en outre, d'après (II,13;5) que :

(II,13;10) $$\|u_1 + u_2\| \leqslant \|u_1\| + \|u_2\|.$$

Nous voyons donc bien que nous avons sur $\mathcal{L}(E;F)$ une structure d'addition. Cette addition fait de $\mathcal{L}(E;F)$ un groupe abélien. L'élément neutre de ce groupe est ce qu'on appelle l'application nulle ou application 0 , c'est-a-dire celle qui fait correspondre, à tout élément de E , la valeur $\vec{0}$ dans F. Soit maintenant λ un scalaire; pour $u \in \mathcal{L}(E;F)$, nous définirons .A u par :

(II,13;11) $$(\lambda u)(\vec{x}) = \lambda(u(\vec{x})).$$

On vérifie que nous définissons bien là une nouvelle application linéaire Au ; en outre cette application est continue, autrement dit $\lambda u \in \mathcal{L}(E;F)$, car :

(II,13;12) $$\|(\lambda u)(\vec{x})\| = |\lambda|\ \|u(\vec{x})\| \leqslant |\lambda|\,\|u\|\,\|\vec{x}\|$$

de plus on a :

(II,13;13) $$\|\lambda u\| = \sup_{\|\vec{x}\|=1}\|(\lambda u)(\vec{x})\| = |\lambda| \sup_{\|\vec{x}\|=1}\|u(\vec{x})\|,$$

autrement dit :

(II,13;14) $$\|\lambda u\| = |\lambda|\,\|u\|.$$

Nous avons ainsi défini sur $\mathcal{L}(E,F)$ une multiplication par les scalaires; et il est facile de voir que cette multiplication possède, par rapport à l'addition des éléments de $\mathcal{L}(E;F)$, toutes les propriétés requises pour faire de $\mathcal{L}(E;F)$ un espace vectoriel sur $\mathbb{K}$. En outre (II,13;10) et (II,13;14) montrent que l'application $u \longrightarrow \|u\|$ est bien une norme (le fait que $\|u\|=0$ entraîne $u=0$ est évident, à cause de l'inégalité (II,13;6); ainsi $\mathcal{L}(E;F)$ est bien un espace vectoriel normé. Si maintenant $u \in \mathcal{L}(E;F)$ et $v \in \mathcal{L}(F;G)$, on a trivialement $v \circ u \in \mathcal{L}(E;G)$ * ; d'autre part $\|(v\circ u)(\vec{x})\| \leqslant \|v\|\ \|u(\vec{x})\| \leqslant \|v\|\ \|u\|\,\|\vec{x}\|$, donc $\|v \circ u\| \leqslant \|v\|\ \|u\|$, ce qui est bien l'inégalité (II,13;7) annoncée.

* La composée de deux applications linéaires est linéaire, et la composée de deux applications continues est continue (théoréme 10)

En particulier, si F est le corps des scalaires $\mathbb{K}$, une application linéaire de E dans $\mathbb{K}$ s'appelle aussi une forme linéaire sur E ; l'espace des formes linéaires continues se note aussi E' et s'appelle dual de E, il est lui-même un espace vectoriel normé (*).

Notation - Si $u \in \mathcal{L}(E;F)$ et $\vec{x} \in E$, il est assez commode de noter $u \cdot \vec{x}$ (ou même $u\vec{x}$) l'image $u(\vec{x})$ de $\vec{x}$ par u. Si maintenant $v \in \mathcal{L}(F;G)$, on note souvent $v \cdot u$ ou vu la composée $v \circ u$. Alors $(v \circ u)(\vec{x}) = v(u(\vec{x}))$ peut se noter $(v \cdot u)\cdot \vec{x}$ ou $v\cdot(u\cdot\vec{x})$ ou $v\cdot u\cdot\vec{x}$, ou $v\,u\,\vec{x}$. Si $F = E$, et $u \in \mathcal{L}(E;E)$, l'itérée $u \circ u$ se note alors u^2, l'itérée $u \circ u \circ u$ se note u^3, etc.....

Théorème 49 - Soient E et F des espaces vectoriels normés, E_1 un sous-espace vectoriel dense de E, et u_1 une application linéaire continue de E_1 dans F. Si F est complet, il existe une application et une seule, u, de E dans F, qui prolonge u_1 et qui soit continue; cette application est linéaire, et la norme de u dans $\mathcal{L}(E;F)$ est égale à la norme de u_1 dans $\mathcal{L}(E_1;F)$.

Démonstration : Comme u_1 est linéaire et continue, elle est uniformément continue (théorème 47); donc, d'après le théorème 45, il existe une application continue u et une seule de E dans F, prolongeant u_1. Montrons d'abord que u est linéaire. Soient $\vec{x}$ et $\vec{y}$ des éléments de E, et λ un scalaire; soit $\vec{x}_0, \vec{x}_1, \vec{x}_2, \ldots \vec{x}_n, \ldots$ et $\vec{y}_0, \vec{y}_1, \vec{y}_2, \ldots \vec{y}_n$, des suites d'éléments de E_1 convergeant vers $\vec{x}$ et $\vec{y}$ respectivement; alors, comme E est un espace vectoriel topologique, (page 65), la suite des $\vec{x}_n + \vec{y}_n$ converge vers $\vec{x}+\vec{y}$, et la suite des $\lambda\vec{x}_n$ converge vers $\lambda\vec{x}$.

On a alors :

$$
(\text{II},13;15)\quad \begin{cases} u(\vec{x}+\vec{y}) = \lim\limits_{n\to\infty} u_1(\vec{x}_n+\vec{y}_n) \overset{(1)}{=} \lim\limits_{n\to\infty}\left[u_1(\vec{x}_n)+u_1(\vec{y}_n)\right] \overset{(2)}{=} \lim\limits_{n\to\infty} u_1(\vec{x}_n) + \lim\limits_{n\to\infty} u_1(\vec{y}_n) \overset{(3)}{} \\ \qquad = u(\vec{x}) + u(\vec{y})^{(4)}; \\ u(\lambda\vec{x}) = \lim\limits_{n\to\infty} u_1(\lambda\vec{x}_n) = \lim\limits_{n\to\infty} \lambda u_1(\vec{x}_n) = \lambda \lim\limits_{n\to\infty} u_1(\vec{x}_n) = \lambda u(\vec{x}), \end{cases}
$$

ce qui prouve bien que u est linéaire.

* Dans le cours d'Algèbre et de géométrie on appelle dual d'un espace vectoriel E (page 81) l'espace vectoriel des formes linéaires sur E. Cette définition coïncide avec la notre si E est de dimension finie car alors, pour toute norme sur E, toute forme linéaire est continue (page 104). Mais, si E est de dimension infinie, et normé, les formes linéaires discontinues ne sont pas intéressantes et on préfère réserver le nom de dual à l'espace des formes linéaires continues.

(1) Parce que $\vec{x}_n + \vec{y}_n$ tend vers $\vec{x}+\vec{y}$, et que u est la fonction continue prolongeant u_1.

(2) Parce que u_1 est linéaire.

(3) Parce que F est un espace vectoriel topologique.

(4) Parce que u est continue et prolonge u_1.

Montrons maintenant que la norme de u dans $\mathscr{L}(E;F)$ est égale à la norme de u_1 dans $\mathscr{L}(E_1;F)$; d'après la définition de la norme (formule (II,13;5)) on a trivialement $\|u_1\| \leqslant \|u\|$; par ailleurs, si $\vec{x} \in E$ est la limite des $\vec{x}_n \in E_1$ alors $u(\vec{x})$ est la limite des $u_1(\vec{x}_n)$. De l'inégalité $\|u_1(\vec{x}_n)\| \leqslant \|u_1\| \, \|\vec{x}_n\|$, et de la continuité de la norme (théorème 9), on déduit alors, par passage à la limite, $\|u(\vec{x})\| \leqslant \|u_1\| \, \|\vec{x}\|$ ce qui prouve que $\|u\| \leqslant \|u_1\|$ donc $= \|u_1\|$.

Corollaire : Dans les conditions de l'énoncé du théorème, c'est-à-dire si E_1 est dense dans E, et F complet, on peut identifier les espaces vectoriels normés $\mathscr{L}(E;F)$ et $\mathscr{L}(E_1;F)$.

En effet, à toute application linéaire continue u de E dans F faisons correspondre sa restriction u_1 à E_1 ; le théorème nous dit précisément que la correspondance ainsi définie $u \to u_1$ de $\mathscr{L}(E;F)$ dans $\mathscr{L}(E_1;F)$, est une bijection, conservant la structure d'espace vectoriel, et conservant les normes.

Cela signifie bien que les deux espaces vectoriels normés $\mathscr{L}(E;F)$ et $\mathscr{L}(E_1;F)$ peuvent être identifiés.

Définition - On appelle espace de BANACH un espace vectoriel normé complet, sur le corps $\mathbb{K}$ des réels ou des complexes. Il résulte de ce que nous avons vu page 96 (théorème 41) que tout espace vectoriel normé de dimension finie est un Banach.

Théorème 50 - Si E et F sont des espaces vectoriels normés, et si F est un Banach, alors $\mathscr{L}(E;F)$ est un Banach. En particulier le dual E' de E est un Banach.

Démonstration : Il nous suffit naturellement de démontrer la première affirmation; comme le corps des scalaires $\mathbb{K}$ est un Banach, la seconde affirmation en résultera immédiatement $(E' = \mathscr{L}(E;\mathbb{K}))$. Soit donc $u_0, u_1, u_2, \ldots u_n, \ldots$ une suite de CAUCHY dans $\mathscr{L}(E;F)$

Cela signifie que, lorsque m et n tendent vers $+\infty$, $\|u_m - u_n\|$ tend vers 0.

Pour tout $\vec{x}$ fixé, on a $\|u_m(\vec{x}) - u_n(\vec{x})\| \leqslant \|u_m - u_n\| \, \|\vec{x}\|$, ce qui prouve que la suite des $u_n(\vec{x})$ est une suite de CAUCHY dans F. Comme F est supposé complet, cette suite converge vers un élément de F , que nous appelerons $u(\vec{x})$. Nous venons donc de définir une application $u : \vec{x} \longrightarrow u(\vec{x})$, de E dans F . Kontrons d'abord queu est linéaire. Soient $\vec{x}$ et $\vec{y}$ des éléments de E, et λ un scalaire; on a :

$$\begin{cases} u(\vec{x}+\vec{y}) = \lim\limits_{n\to\infty} u_n(\vec{x}+\vec{y})^{(1)} = \lim\limits_{n\to\infty}\left(u_n(\vec{x})+u_n(\vec{y})\right)^{(2)} = \lim\limits_{n\to\infty} u_n(\vec{x}) + \lim\limits_{n\to\infty} u_n(\vec{y})^{(3)} = u(\vec{x})+u(\vec{y}); \\ u(\lambda\vec{x}) = \lim\limits_{n\to\infty} u_n(\lambda\vec{x})^{(1)} = \lim\limits_{n\to\infty} \lambda\, u_n(\vec{x})^{(2)} = \lambda \lim\limits_{n\to\infty} u_n(\vec{x})^{(3)} = \lambda\, u(\vec{x}), \end{cases} \tag{II,13;16}$$

ce qui prouve la linéarité de u.

Montrons maintenant que u est continue. D'après le critère de CAUCHY, il existe un entier+ tel que $m \geqslant p$, $n \geqslant p$ entraîne $\|u_m - u_n\| \leqslant \varepsilon$. On en déduit que, pour $m \geqslant p$, $n \geqslant p$, on a $\|u_m\| \leqslant \|u_n\| + \varepsilon$

On a donc aussi :

$$\|u_m(\vec{x}) - u_n(\vec{x})\| \leqslant \varepsilon \|\vec{x}\| , \tag{II,13;17}$$

$$\|u_m(\vec{x})\| \leqslant \left(\|u_n\| + \varepsilon\right)\|\vec{x}\| . \tag{II,13;18}$$

Fixons $n \geqslant p$ et $\vec{x}$, et faisons tendre m vers $+\infty$ dans (II,13;18); on obtient

$$\|u(\vec{x})\| \leqslant \left(\|u_n\| + \varepsilon\right)\|\vec{x}\| , \tag{II,13;19}$$

donc u est bien continue, c'est bien un élément de $\mathcal{L}(E;F)$.
Si maintenant, toujours pour $n \geqslant p$ et $\vec{x}$ fixés, nous faisons tendre m vers $+\infty$ dans (II,13;17), nous obtenons

$$\|u(\vec{x}) - u_n(\vec{x})\| \leqslant \varepsilon \|\vec{x}\| , \tag{II,13;20}$$

donc

$$\|u - u_n\| \leqslant \varepsilon . \tag{II,13;21}$$

Cela montre que u_n tend vers u dans l'espace normé $\mathcal{L}(E;F)$, pour n tendant vers $+\infty$, donc $\mathcal{L}(E;F)$ est bien complet.

(1) D'après la définition de u .
(2) D'après la linéarité de u_n .
(3) Parce que F est un espace vectoriel topologique

Produits d'espaces vectoriels normés

Soient E_1 et E_2 deux espaces vectoriels sur le corps $\mathbb{K}$; on sait qu'on peut mettre sur leur produit, $E_1 \times E_2$, une structure d'espace vectoriel. Si en effet $(\vec{x}_1, \vec{x}_2)$ et $(\vec{y}_1, \vec{y}_2)$ sont deux éléments de ce produit, on posera :

(II,13;22) $$(\vec{x}_1, \vec{x}_2) + (\vec{y}_1, \vec{y}_2) = (\vec{x}_1 + \vec{y}_1, \vec{x}_2 + \vec{y}_2) ;$$

et si λ est un scalaire, on posera :

(II,13;23) $$\lambda (\vec{x}_1, \vec{x}_2) = (\lambda \vec{x}_1, \lambda \vec{x}_2) .$$

Les opdrations ainsi définies sur $E_1 \times E_2$ en font manifestement un espace vectoriel sur le corps $\mathbb{K}$ * .

D'ailleurs l'espace vectoriel $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) n'est autre que l'espace vectoriel produit de n espaces, Identiques à $\mathbb{R}$ (resp. $\mathbb{C}$) considéré comme espace vectoriel sur lui-même $\mathbb{R}$ (ou $\mathbb{C}$).

Dans cet espace vectoriel, il existe deux sous-espaces vectoriels remarquables : l'espace $A_1 = E_1 \times \{\vec{0}\}$ des éléments de la forme $(\vec{x}_1, \vec{0})$, et l'espace $A_2 = \{\vec{0}\} \times E_2$ des éléments de la forme $(\vec{0}, \vec{x}_2)$. Ces deux sous-espaces sont supplémentaires, en ce sens que tout élément $(\vec{x}_1, \vec{x}_2)$ du produit s écrit d'une manière et d'une seule comme somme d'un éldment de l'un et d'un élément de l'autre, à savoir $(\vec{x}_1, \vec{x}_2) = (\vec{x}_1, \vec{0}) + (\vec{0}, \vec{x}_2)$.

Lorsque deux sous-espaces vectoriels d'un espace vectoriel sont supplémentaires, on écrit aussi que celui-ci est somme directe des deux premiers, et cette somme directe se note par le signe $\oplus$, nous pouvons donc aussi écrire :

(II,13;24) $$E_1 \times E_2 = E_1 \times \{\vec{0}\} \oplus \{\vec{0}\} \times E_2 .$$

SI E_1 et E_2 sont des espaces vectoriels normés, dont les normes sont notées par $\| \ \|_1$ et $\| \ \|_2$, il n'existe pas, sur le produit $E_1 \times E_2$, une norme qui s'impose plus qu'une autre; on peut prendre par exemple l'une quelconque des normes équivalentes suivantes :

(II,13;25) $$\|\vec{x}_1, \vec{x}_2\| = \mathrm{Max}\left(\|\vec{x}_1\|_1, \|\vec{x}_2\|_2\right) \text{, ou } \|\vec{x}_1\|_1 + \|\vec{x}_2\|_2 \text{, ou } \sqrt{\|x_1\|_1^2 + \|x_2\|_2^2} .$$

* Voir Cours d'Algèbre, Chapitre 1, n°6, page 22.

Pour l'une quelconque de ces normes, la tapologie de $E_1 \times E_2$ est la topologie produit; les sous-espaces supplémentaires $E_1 \times \{0\}$ et $\{0\} \times E_2$ sont fermés dans E . Sauf mention expresse du contraire, $E_1 \times E_2$ sera toujours muni d'une des 3 normes précédentes; et le même symbole $\| \ \|$ désignera les normes dans $E_1, E_2, E_1 \times E_2$.

Théorème 51 - Soit E_1 et E_2 deux espaces vectoriels normés. Toute application linéaire continue u de $E_1 \times E_2$ dans un espace vectoriel normé F s'exprime, d'une manière et d'une seule, sous la forme :

$$u(\vec{x}_1, \vec{x}_2) = u_1(\vec{x}_1) + u_2(\vec{x}_2), \tag{II,13;26}$$

où u_1 (resp. u_2) est une application linéaire continue de E_1 (resp. E_2) dans F . Réciproquement, si u_1 (resp. u_2) est une application linéaire continue de E_1 (resp. E_2) dans F , l'application u définie par (II,13;26) est linéaire et continue de $E_1 \times E_2$ dans F .

Démonstration : Etant donné u , nous définirons u_1 comme suit : la valeur de u_1 sur un élément $\vec{x}_1$ de E_1 est la valeur de u sur l'élément $(\vec{x}_1, \vec{0})$ de $E_1 \times E_2$: $u_1(\vec{x}_1) = u(\vec{x}_1, \vec{0})$.

Nous définirons de la même manière l'application u_2 de E_2 dans F. u_1 et u_2 sont manifestement des applications linéaires. Elles sont d'autre part continues si l on met sur $E_1 \times E_2$ l'une quelconque des normes équivalentes de la formule (II,13;25), c'est-à-dire si l'on met sur $E_1 \times E_2$ la topologie produit.

En effet on a alors l'inégalité :

$$\|u_1(\vec{x}_1)\| = \|u(\vec{x}_1, \vec{0})\| \leq \|u\| \ \|\vec{x}_1, \vec{0}\| = \|u\| \ \|\vec{x}_1\|, \tag{II,13;27}$$

ce qui prouve notre affirmation, et montre en même temps que l'on a les inégalités :

$$\|u_1\| \leq \|u\|, \quad \|u_2\| \leq \|u\|. \tag{II,13;28}$$

On voit alors immédiatement que l'on a :

$$u(\vec{x}_1, \vec{x}_2) = u(\vec{x}_1, \vec{0}) + u(\vec{0}, \vec{x}_2) \overset{*}{=} u_1(\vec{x}_1) + u_2(\vec{x}_2), \tag{II,13;29}$$

ce qui prouve bien que u a l'expression indiquée.

On voit d'ailleurs immédiatement qu'il n'existe pas d'autres, applications u_1 et u_2 possédant la même propriété, parce que, si v_1 et v_2 sont des applications linéaires telles que $u(\vec{x}_1, \vec{x}_2) = v_1(\vec{x}_1) + v_2(\vec{x}_2)$, on a $u(\vec{x}_1, \vec{0}) = v_1(\vec{x}_1)$, donc $v_1 = u_1$, et de même $v_2 = u_2$.

* A cause de la linéarité de u , et parce que $(\vec{x}_1, \vec{x}_2) = (\vec{x}_1, \vec{0}) + (\vec{0}, \vec{x}_2)$.

Inversement, si u_1 (resp. u_2) est une application linéaire continue de E_1 (resp. E_2) dans F , l'application u définie par (II,13;26) est manifestement linéaire et d'autre part elle est continue, car

(II,13;30)
$$\|u(\vec{x}_1,\vec{x}_2)\| \leqslant \|u_1(\vec{x}_1)\| + \|u_2(\vec{x}_2)\| \leqslant \|u_1\| \ \|\vec{x}_1\| + \|u_2\| \ \|\vec{x}_2\|$$
$$\leqslant (\|u_1\| + \|u_2\|) \operatorname{Max}(\|\vec{x}_1\|, \|\vec{x}_2\|) \leqslant (\|u_1\| + \|u_2\|) \|(\vec{x}_1,\vec{x}_2)\| .$$

On en déduit en même temps qu'elle vérifie l'inégalité :

(II,13;31)
$$\|u\| \leqslant \|u_1\| + \|u_2\| .$$

Nous laissons aux élèves le soin de définir la structure d'espace vectoriel normé d'un produit de n espaces vectoriels normés E_1 , $E_2, \dots E_n$, et d'étendre le théorème 51 aux applications linéaires continues de ($E_1 \times E_2, \dots \times E_n$) dans F .

Applications bilinéaires continues d'un produit d'espace vectoriel normé dans un espace vectoriel normé

Soient E , F , G , des espaces vectoriels sur un corps $\mathbb{K}$. Nous avons vu ce qu'étaient les applications linéaires de l'espace vectoriel produit E x F dans G ; nous allons considérer maintenant une notion tout à fait différente, celle d'application bilinéaire de E x F dans G .

Une application u de E x F dans G est dite bilinéaire si, lorsqu'on fixe chacune des variables, elle est linéaire par rapport à l'aütre. Fixons $\vec{x}$, alors u définit une application partielle $\vec{y} \longrightarrow u(\vec{x},\vec{y})$ de F dans G ; cette application se note $u_{\vec{x}}$, de sorte que l'on a :

(II,13;32)
$$u_{\vec{x}}(\vec{y}) = u(\vec{x},\vec{y}) .$$

Il est également assez commode de noter $u(\vec{x}, \ )$ cette application partielle, en omettant la variable $\vec{y}$* .

Alors, si u est bilinéaire, $u_{\vec{x}}$ doit être une application linéaire de F dans G , autrement dit l'on a :

(II,13;33)
$$\begin{cases} u(\vec{x},\vec{y}_1+\vec{y}_2) = u(\vec{x},\vec{y}_1) + u(\vec{x},\vec{y}_2) \\ u(\vec{x}, a\,\vec{y}) = \lambda\, u(\vec{x},\vec{y}) \end{cases}$$

* Si E et F sont différents, la notation $u_{\vec{x}}$, pour un élément $\vec{x}$ donné de E , n'est pas douteuse. Mais si E = F , et si u est une application bilinéaire de E x E dans G , alors $u_{\vec{x}}$, pour un élément $\vec{x}$ de E , ne signifie rien, car on ne sait pas si c'est la première fonction partielle ou la deuxième, et il faut écrire $u(\vec{x}, \ )$ ou $u(\ ,\vec{x})$.

De la même manière, l'application partielle $u_{\vec{y}}$ ou $u(\ ,\vec{y}) : \vec{x} \longrightarrow u(\vec{x},\vec{y})$, doit être une application linéaire de **E** dans G , autrement dit :

(II,13;34) $$u(\vec{x}_1 + \vec{x}_2 , \vec{y}) = u(\vec{x}_1 , \vec{y}) + u(\vec{x}_2 , \vec{y})$$
$$u(\lambda x , \vec{y}) = \lambda u(\vec{x} , \vec{y}) .$$

Les relations concernant l'addition s'expriment souvent en disant que u est distributive relativement à l'addition. Des deux formules précédentes on tire :

(II,13;35) $$u(\vec{x}_1 + \vec{x}_2 , \vec{y}_1 + \vec{y}_2) = u(\vec{x}_1 , \vec{y}_1) + u(\vec{x}_1 , \vec{y}_2) + u(\vec{x}_2 , \vec{y}_1) + u(\vec{x}_2 , \vec{y}_2)^*$$
$$u(\lambda \vec{x}, \mu \vec{y}) = \lambda \mu u(\vec{x} , \vec{y}) .$$

Par exemple, le produit scalaire des vecteurs est une application bilinéaire de $\mathbb{R}^3 \times \mathbb{R}^3$ dans $\mathbb{R}$, le produit vectoriel une application bilinéaire de $\mathbb{R}^3 \times \mathbb{R}^3$ dans $\mathbb{R}^3$.
Le prototype des applications bilinéaires est le produit usuel des nombres réels ou complexes, application bilinéaire de $\mathbb{R} \times \mathbb{R}$ dans $\mathbb{R}$ ou de $\mathbb{C} \times \mathbb{C}$ dans $\mathbb{C}$. D'ailleurs toute application bilinéaire de $\mathbb{R} \times \mathbb{R}$ dans $\mathbb{R}$ est de la forme $(x,y) \to cxy$ ou c est une constante réelle; alors qu'une application linéaire de $\mathbb{R} \times \mathbb{R}$ dans $\mathbb{R}$ est de la forme $(x,y) \to ax + by$, où a et b sont des constantes réelle.

Théorème 52 - Toute application bilinéaire d'un produit d'espaces vectoriels normés $E \times F$ dans un espace vectoriel normé G, continue à l'origine, est continue partout (mais non uniformément continue); pour qu'il en soit ainsi, il faut et il suffit qu'il existe une constante $k \geqslant 0$ telle que l'on ait pour tout $\vec{x}$ de E et tout $\vec{y}$ de F :

(II,13;36) $$\| u(\vec{x},\vec{y}) \| \leqslant k \| \vec{x} \| \, \| \vec{y} \| ^{**} .$$

Démonstration : Soient **E** , F , **G** . des espaces vectoriels normes, u une application bilinéaire de $E \times F$ dans G continue à l'origine. Alors, d'après la définition de la topologie produit, il doit exister un nombre $\eta > 0$ tel que $\|\vec{x}\| \leqslant \eta$, $\|\vec{y}\| \leqslant \eta$ entraîne $\|u(\vec{x},\vec{y})\| \leqslant 1$.

* Alors que, si u était linéaire, on aurait, au 2ème membre : $u(\vec{x}_1 , \vec{y}_1) + u(\vec{x}_2 , \vec{y}_2)$ ou $u(\vec{x}_1 , \vec{y}_2) + u(\vec{x}_2 , \vec{y}_1)$.

** Pour une application linéaire, on aurait au 2ème membre, par exemple, $k(\|\vec{x}\| + \|\vec{y}\|)$.

Alors, si nous faisons des homothéties de rapports λ et $\mu \geq 0$, la 2ème formule (II,13;35) montre que $\|\vec{x}\| \leq \lambda\eta$, $\|\vec{y}\| \leq \mu\eta$ entraîne $\|u(\vec{x},\vec{y})\| \leq \lambda\mu$.

Mais les premières inégalités sont toujours vérifiées si l'on prend $\lambda = \frac{\|x\|}{\eta}$, $\mu = \frac{\|\vec{y}\|}{\eta}$, il en résulte que l'on a toujours $\|u(\vec{x},\vec{y})\| \leq \frac{1}{\eta^2}\|\vec{x}\|\|\vec{y}\|$, ce qui est bien la formule (II,13;36) avec $k = \frac{1}{\eta^2}$. Réciproquement, si u est une application bilinéaire vérifiant une inégalité du type (II,13;36), montrons qu'elle est continue partout. Soit donc $(\vec{a},\vec{b})$ un point de $E \times F$, et soit $\varepsilon > 0$ donné. On a la formule :

(II,13;37) $$u(\vec{x},\vec{y}) - u(\vec{a},\vec{b}) = u(\vec{x}-\vec{a},\vec{y}) + u(\vec{a},\vec{y}-\vec{b}),$$

d'où l'on déduit la majoration :

(II,13;38) $$\|u(\vec{x},\vec{y}) - u(\vec{a},\vec{b})\| \leq k\|\vec{x}-\vec{a}\| \, \|\vec{y}\| + k\|\vec{a}\| \, \|\vec{y}-\vec{b}\| .$$

Choisissons alors le nombre $\eta_2 = \frac{\varepsilon}{2k\|\vec{a}\|}$. On voit que $\|\vec{y}-\vec{b}\| \leq \eta_2$ entraîne la majoration $k\|\vec{a}\| \, \|\vec{y}-\vec{b}\| \leq \frac{\varepsilon}{2}$.
Choisissons ensuite le nombre $\eta_1 = \frac{\varepsilon}{2k(\|\vec{b}\|+\eta_2)}$. On voit que, pour $\|\vec{y}-\vec{b}\| \leq \eta_2$, on a la majoration $\|\vec{y}\| \leq \|\vec{b}\| + \eta_2$, d'où l'on déduit que $\|\vec{x}-\vec{a}\| \leq \eta_1$ entraîne alors la majoration $k\|\vec{x}-\vec{a}\|\|\vec{y}\| \leq \frac{\varepsilon}{2}$.

L'ensemble des majorations obtenues montre que $\|\vec{x}-\vec{a}\| \leq \eta_1$, $\|\vec{y}-\vec{b}\| \leq \eta_2$ entraîne $\|u(\vec{x},\vec{y}) - u(\vec{a},\vec{b})\| \leq \varepsilon$, ce qui est précisément la définition de la continuité de u au point $(\vec{a},\vec{b})$.

Notons par contre que, si une application bilinéaire u n'est pas identiquement nulle, elle n'est jamais uniformément continue. Supposons en effet qu'il existe au moins un couple $(\vec{a},\vec{b})$ tel que $u(\vec{a},\vec{b}) \neq \vec{0}$.

Considérons alors dans $E \times F$ la suite des points $\vec{X}_n = (n\vec{a}, n\vec{b})$ et la suite des points $\vec{Y}_n = \left((n+\frac{1}{n})\vec{a}, (n+\frac{1}{n})\vec{b}\right)$. On a manifestement la majoration: $\|\vec{X}_n - \vec{Y}_n\| \leq \frac{1}{n}(\|\vec{a}\| + \|\vec{b}\|)$, ce qui prouve que cette distance converge vers 0 lorsque n tend vers $+\infty$. Néanmoins, la différence :
$u(\vec{X}_n) - u(\vec{Y}_n) = \left(n^2-(n+\frac{1}{n})^2\right)u(\vec{a},\vec{b})$ s'écrit aussi : $-\left(2+\frac{1}{n^2}\right) u(\vec{a},\vec{b})$, et par conséquent, lorsque n tend vers $+\infty$, sa norme tend vers $2\|u(\vec{a},\vec{b})\| \neq 0$, ce qui est contradictoire avec la

possibilité d'une continuité uniforme * .

Définition : La borne inférieure des nombres k tels que l'on ait l'inégalité (II,13;36) s'appelle **la norme de l'application bilinéaire u** et se note $\|u\|$, on a donc :

(II,13,39) $$\|u\| = \sup_{\vec{x}\neq \vec{0},\vec{y}\neq\vec{0}} \frac{\|u(\vec{x},\vec{y})\|}{\|\vec{x}\|\,\|\vec{y}\|} = \sup_{\|\vec{x}\|\leq 1,\|\vec{y}\|\leq 1} \|(\vec{x},\vec{y})\|.$$

Notons qu'on a toujours :

(II,13;40) $$\|u(\vec{x},\vec{y})\| \leq \|u\|\ \|\vec{x}\|\|\vec{y}\|.$$

Théorème 53 - L'ensemble $\mathcal{L}_2(E,F;G)$ des **applications bilinéaires** d'un produit d'espaces vectoriels normés $E \times F$ dans un espace vectoriel normé G, admet lui-même une structure d'espace vectoriel normé, si l'on **définit** la norme par la formule (II,13;39). Si en outre **G** est un espace de BANACH, il est lui aussi un espace de BANACH; en particulier en prenant pour G le corps des scalaires, l'espace vectoriel des formes bilinéaires continues sur un produit d'espacesvectoriels normés est un espace de BANACH.

La démonstration est identique à celle des théorèmes 48 et 50.

Considérons maintenant une application bilinéaire contins u de $E \times F$ dans . nous avons vu qu'elle **définit, pour** x fixé, une application'linéaire $u_{\vec{x}}$ de F dans **G** ; cette application est trivialement continue et de norme $\leq \|u\|\ \|\vec{x}\|$, en vertu de (II,13;40), de sorte que $u_{\vec{x}}$ appartient à $\mathcal{L}(F;G)$. Nous voyons donc que nous venons d'établir une correspondance **qui,** à chaque élément $\vec{x}$ de E , fait correspondre un **élément** $u_{\vec{x}}$ de $\mathcal{L}(F;G)$ autrement dit nous venons de définir une application U de E'dans $\mathcal{L}(F;G)$. Mais E et $\mathcal{L}(F;G)$ sont tous les deux des espaces vectoriels normés; nous allons montrer que U est une application linéaire continue de E dans $\mathcal{L}(F;G)$, et que la norme $\|U\|$ de cette application linéaire, n'est autre que la norme $\|u\|$ de l'application bilinéaire u considérée.

* D'ailleurs on sait bien que le produit $(x,y) \longrightarrow xy$ n'est pas une application **uniformément** continue de $\mathbb{R} \times \mathbb{R}$ **dans** $\mathbb{R}$ et que la fonction $x \longrightarrow x^2$ n'est pas uniformément oontinue **de** $\mathbb{R}$ **dans** $\mathbb{R}$.

Démontrons d'abord que U est linéaire, et ceci naturellement ne fait pas intervenir les topologies.

Nous devons démontrer que l'on a :

(II,13;41) $$u_{\vec{x}_1+\vec{x}_2} = u_{\vec{x}_1} + u_{\vec{x}_2}\ , \quad u_{\lambda\vec{x}} = \lambda u_{\vec{x}}$$

Or toutes les quantités écrites sont des éléments de $\mathcal{L}(F;G)$ des applications linéaires de F dans G. d'après la définition de la somme de deux applications linéaires, ou du produit d'une application linéaire par un scalaire, ces égalités reviennent à écrire que, pour tout élément $\vec{y}$ de F, on a :

(II,13;42) $$u(\vec{x}_1+\vec{x}_2,\vec{y}) = u(\vec{x}_1,\vec{y}) + u(\vec{x}_2,\vec{y}),\quad u(\lambda\vec{x},\vec{y}) = \lambda u(\vec{x},\vec{y}).$$

Mais cette égalité est vraie, elle revient à écrire que l'application partielle $u_{\vec{y}}$ de F dans G, pour $\vec{y}$ fixé, est linéaire.

Démontrons maintenant la continuité de U. La norme de $U(\vec{x}) = u_{\vec{x}}$, application linéaire continue de F dans G, est majorée par $\|u\|\,\|\vec{x}\|$; cela prouve que U est une application linéaire continue de E dans $\mathcal{L}(F;G)$ et que $\|U\| \leq \|u\|$. Mais on a aussi $\|U(\vec{x})\| \leq \|U\|\,\|\vec{x}\|$; alors

$\|u(\vec{x},\vec{y})\| = \|U(\vec{x})\cdot\vec{y}\| \leq \|U(\vec{x})\|\,\|\vec{y}\| \leq \|U\|\,\|\vec{x}\|\,\|\vec{y}\|$, de sorte qu'on a aussi $\|u\| \leq \|U\|$, et par suite $\|u\| = \|U\|$.

Inversement, partons d'une application linéaire continue U de E dans $\mathcal{L}(F;G)$. Alors, pour tout $\vec{x}$ de E, $U(\vec{x})$ est un élément de $\mathcal{L}(F;G)$, c'est-à-dire une application linéaire continue de F dans G, et par conséquent, pour tout élément $\vec{y}$ de F, $U(\vec{x})\cdot\vec{y}$ est un élément de G. Si nous posons $u(\vec{x},\vec{y}) = U(\vec{x})\cdot\vec{y}$, on voit que $u : (\vec{x},\vec{y}) \to u(\vec{x},\vec{y})$, est une application de $E\times F$ dans G ; on démontre sans difficulté que cette application est bilinéaire, et elle est trivialement continue, car $\|u(\vec{x},\vec{y})\| \leq \|U(\vec{x})\|\,\|\vec{y}\| \leq \|U\|\,\|\vec{x}\|\,\|\vec{y}\|$, et $\|u\| \leq \|U\|$.

L'application partielle associée à u, pour $\vec{x}$ fixé dans E, n'est autre précisément que $U(\vec{x})$ et l'application linéaire continue de E dans $\mathcal{L}(F;G)$ associée à u, n'est autre que l'application U de départ. Nous voyons ainsi que <u>la correspondance établie entre u et U est une bijection de l'espace $\mathcal{L}_2(E,F;G)$ des applications bilinéaires continues de $E\times F$ dans G, sur l'espace $\mathcal{L}(E;\mathcal{L}(F;G))$ des applications liné -</u>

aires continues de E dans $\mathcal{L}(F;G)$, et cette bijection conserve les structures vectorielles et les normes. Il existe naturellement une bijection analogue de $\mathcal{L}_2(E, F; G)$ sur $\mathcal{L}(F; \mathcal{L}(E;G))$. **

Théorème 54 - Si E et F sont des espaces vectoriels normés, l'application $(u, \vec{x}) \to u.\vec{x}$ de $\mathcal{L}(E;F) \times E$ dans F est bilinéaire continue et de norme 1 (sauf si E ou F est réduit à son origine).

Si E, F, G, sont trois espaces vectoriels normés, l'application $(u,v) \to v \circ u$ de $\mathcal{L}(E;F) \times \mathcal{L}(F;G)$ dans $\mathcal{L}(E;G)$ est bilinéaire continue et de norme 1 (sauf si E, F, ou G est réduit à son origine).

Démonstration - Le fait que les applications considérées soient bilinéaires est trivial, et résulte toujours des définitions de la somme de deux applications linéaires et du produit d'une application linéaire par un scalaire * .
Le fait que ces applications bilinéaires soient continues et de norme ≤ 1 , résulte trivialement de la majoration

$$\|u.\vec{x}\| \leq \|u\| \, \|\vec{x}\| , \quad \|v \circ u\| \leq \|v\| \, \|u\| \qquad \text{(formules (II,13;6))}$$

et (II,13;7).

Par ailleurs, u étant quelconque et $\varepsilon > 0$ étant donné, il existe, d'après la définition de la norme de u , au moins un élément $\vec{x} \neq \vec{0}$ de E tel que l'on ait $\|u.\vec{x}\| \geq (1-\varepsilon)\|u\| \, \|\vec{x}\|$, ce qui prouve que la norme de la première application bilinéaire est $\geq 1 - \varepsilon$; comme ε est aussi petit qu'on veut, cette norme est ≥ 1 , et par conséquent elle est égale à 1.

Quant au fait que la norme de la deuxième application bilinéaire soit aussi égale à 1, nous l'admettrons sans démonstration.

Applications multilinéaires continues

Soient $E_1, E_2, \ldots E_n$, F des espaces vectoriels normés sur le même corps $\mathbb{K}$ (corps des réels ou corps des complexes). Une application u de $E_1 \times E_2 \times \cdots \times E_n$ dans F est dite multilinéaire (ou n - linéaire) si quand on fixe n-1 des variables dans n - 1 quelconques des espaces vectoriels, elle est linéaire par rapport à la n-ième.

* C'est la bilinéarité de ces applications qui permet d'introduire les notations multiplicatives $u\vec{x}$ et vu et u^n (page 109).

** Voir Cours d'algèbre, Chapitre II, n°1. Ce que nous appelons ici U y est appelé $\mathfrak{s}_u$.

Par exemple, le produit $(x_1, x_2, \dots x_n) \longrightarrow x_1 x_2 \dots x_n$ est une application n-linéaire de $\mathbb{K}^n$ dans $\mathbb{K}$; dans $\mathbb{R}^3$, le produit mixte de 3 vecteurs définit une application trilinéaire de $\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$ dans $\mathbb{R}$.

L'espace des applications n-linéaires continues de $E_1 \times E_2 \dots \times E_n$ dans F admet une structure d'espace vectoriel normé; c'est un Banach si F est un Banach; on le note $\mathcal{L}(E_1, E_2, \dots E_n ; F)$.
Les développements qui précèdent le théorème 54 s'étendent immédiatement comme suit :

Théorème 54 bis -

Il existe des bijections, conservant la structure vectorielle et la norme, entre les espaces :
$\mathcal{L}_n(E_1, E_2, \dots, E_n ; F)$, $\mathcal{L}_1(E_1 ; \mathcal{L}_{n-1}(E_2, \dots, E_n ; F))$,
$\mathcal{L}_2(E_1, E_2 ; \mathcal{L}_{n-2}(E_3, \dots, E_n ; F))$, ...,
$\mathcal{L}_{n-1}(E_1, E_2, \dots, E_{n-1} ; \mathcal{L}_1(E_n ; F))$,
et ceux qu'on obtient en permutant les indices.
Si $u \in \mathcal{L}_n(E_1, E_2, \dots, E_n ; F)$, l'élément U qui lui est associé dans $\mathcal{L}_p(E_1, E_2, \dots, E_p ; \mathcal{L}_{n-p}(E_{p+1}, \dots, E_n ; F))$ est donné par

$$U(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) . (\vec{x}_{p+1}, \dots, \vec{x}_n) = u(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p, \vec{x}_{p+1}, \dots, \vec{x}_n) \tag{II,13;43}$$

Algèbres; algèbres normées.

Une algèbre $\mathcal{A}$ sur le corps $\mathbb{K}$ est un espace vectoriel sur $\mathbb{K}$ muni en outre d'une application bilinéaire de $\mathcal{A} \times \mathcal{A}$ dans $\mathcal{A}$, appelée multiplication, notée $(\vec{x}, \vec{y}) \longrightarrow \vec{x}\vec{y}$, associative : $(\vec{x}\vec{y})\vec{z} = \vec{x}(\vec{y}\vec{z})$ et ayant une unité $\vec{I} \neq \vec{0}$ ($\vec{x}\,\vec{I} = \vec{I}\,\vec{x}$ pour tout $\vec{x}$). Le corps $\mathbb{K}$ lui-même est une algèbre. Une algèbre normée est une algèbre, qui est aussi un espace vectoriel normé, tel que

$$\|\vec{I}\| = 1 , \|\vec{x}\,\vec{y}\| \leq \|\vec{x}\| \, \|\vec{y}\| \tag{II,13;44}$$

Une algèbre de Banach est une algébre normée complète. Si E est un espace vectoriel normé, $\mathcal{L}(E ; E)$ est une algèbre normée, si on définit comme multiplication la composition $(u, v) \longrightarrow u \circ v$; c'est une algèbre de Banach si E est un Banach.

§ 14 SÉRIES DANS LES ESPACES VECTORIELS NORMÉS

Soit $\vec{u}_0, \vec{u}_1, \vec{u}_2, \dots \vec{u}_n, \dots$ une suite d'éléments d'un espace vectoriel normé E , on pose :

$$\vec{S}_n = \vec{u}_0 + \vec{u}_1 + \dots + \vec{u}_n = \sum_{0 \leq m \leq n} \vec{u}_m . \tag{II,14;1}$$

$\vec{S}_n$ s'appelle la somme partielle d'indice n de la série.

C'est un élément de E . <u>*On* dit que la série $\sum_{n=0}^{\infty} \vec{u}_n$</u> <u>est convergente, et de somme $\vec{S}$, si la suite des</u> $\vec{S}_n$ est <u>convergente, et de limite $\vec{S}$</u> * . La notion de série se ramène ainsi à celle de suite : réciproquement d'ailleurs, dans un espace vectoriel normé la notion de suite se ramène à celle de série. Une suite $\vec{x}_0, \vec{x}_1, \vec{x}_2, \dots \vec{x}_n, \dots$ est convergente, et de limite $\vec{a}$, si et seulement si la série $\vec{x}_0 + (\vec{x}_1 - \vec{x}_0) + (\vec{x}_2 - \vec{x}_1) + \cdots + (\vec{x}_n - \vec{x}_{n-1}) + \cdots$ est convergente et de somme $\vec{a}$.

D'après ce que nous avons vu pour les suites, si E est de dimension finie, et si l'on en a choisi une base, la série $\sum_n \vec{u}_n$ est convergente et de somme $\vec{S}$, si et seulement si chacune des série de composantes : $\sum_n (u_n)_i$ est convergente et de somme S_i, où $(u_n)_i$ est la i-ième coordonnée de $\vec{u}_n$, et S_i la i-ième coordonnée de $\vec{S}$.

L'intérêt des espaces de BANACH est que, dans de tels espaces comme dans le corps des complexes lui-même, on peut reconnaître qu'une série est convergente sans connaître à l'avance la somme de la série.

Il faut et il suffit, pour qu'une série soit convergente dans un espace de BANACH E, qu'elle vérifie la critère de CAUCHY : lorsquem et n tendent vers $+\infty$, la quantité $\|\vec{S}_n - \vec{S}_m\|$ tend vers 0 ; ou encore :

(II,14;2) $$(\forall \varepsilon > 0)(\exists n \in \mathbb{N})(\forall m \geq n)(\forall k \in \mathbb{N}) : \|\vec{u}_m + \vec{u}_{m+1} + \cdots + \vec{u}_{m+k}\| \leq \varepsilon .$$

On en déduit le théorème fondamental suivant,qui donne le critère le plus important pour la convergence des séries de vecteurs :

<u>Théorème 55</u> .: <u>Soit E un espace de BANACH, et $\sum_{n=0}^{\infty} \vec{u}_n$ une série</u> ; <u>si la série des normes $\sum \|\vec{u}_n\|$ est convergente, alors la série elle-même est convergente, et en outre on a</u> :

(II,14;3) $$\left\|\sum \vec{u}_n\right\| \leq \sum \|\vec{u}_n\| .$$

<u>Demonstration</u> - Vérifions en effet que l'on a le critère de CAUCHY. Etant donné le nombre $\varepsilon > 0$, on peut déterminer un entier n tel que $m \geq n$ entraine, quel que soit k ,

* La convergence ou la divergence d'une série subsiste si on remplace la norme de E par une norme équivalente; car la convergence est une propriété topologique.

$\|\vec{u}_m\| + \|\vec{u}_{m+1}\| + \cdots + \|\vec{u}_{m+k}\| \leq \varepsilon$, car la série des normes, étant, une série (à termes positifs) convergente, vérifie le critère de CAUCHY; on en déduit à fortiori, pour $m \geq n$, $\|\vec{u}_m + \vec{u}_{m+1} + \cdots + \vec{u}_{m+k}\| \leq \varepsilon$,

autrement dit la série donnée dans E satisfait au critère de CAUCHY ; et comme E est supposé complet, elle est convergente. En outre, pour un nombre fini de termes, on a l'inégalité $\|\vec{S}_n\| \leq \sum_{0 \leq m \leq n} \|\vec{u}_n\|$, d'ou l'on déduit alors, en passant à la limite pour n tendant vers $+\infty$, et en tenant compte de ce que+ la convergence de $\vec{S}_n$ vers $\vec{S}$ entraine la convergence de $\|\vec{S}_n\|$ vers $\|\vec{S}\|$ (théorème 9), l'inégalité cherchée (II,14;3).

Définition - On dit qu'une série d'éléments d'un espace de BANACH est normalement convergente, ou absolument convergente * , si la série des normes, série numérique à termes positifs, est convergente. On dit qu'une série est semi-convergente, si elle est convergente, sans être normalement convergente. Le fait que E soit complet est absolument essentiel. On peut d'ailleurs démontrer la réciproque suivante :

Théorème 56 - Soit E un espace vectoriel normé. Si toute série $\sum \vec{u}_n$ d'éléments de E , dont la série des normes $\sum \|\vec{u}_n\|$ est convergente, est aussi convergente, alors E est complet.

Démonstration - Nous devons démontrer que toute suite de CAUCHY de E est convergente.

Soit donc $\vec{x}_0, \vec{x}_1, \vec{x}_2, \ldots \vec{x}_n, \ldots$ une telle suite de CAUCHY. Quel que soit l'entier $k \geq 0$, on peut trouver un entier p_k , tel que $m \geq p_k$, $n \geq p_k$ entraîne $\|\vec{x}_m - \vec{x}_n\| \leq \frac{1}{2^k}$.

Nous choisirons les entiers p_k les uns après les autres, de façon que la suite des p_k soit strictement croissante. Dans ces conditions, considérons la série : $\vec{x}_{p_0} + (\vec{x}_{p_1} - \vec{x}_{p_0}) + (\vec{x}_{p_2} - \vec{x}_{p_1}) + \ldots$; la série de ses normes est majorée par la série $\|\vec{x}_{p_0}\| + 1 + \frac{1}{2} + \frac{1}{4} + \cdots$, et elle est par conséquent convergente en vertu des propriétés connues des séries numériques à termes positifs. Alors, d'après les hypothèses faites sur E , la série est elle-mêmeconvergente, ce qui signifie que la suite partielle des x_{p_n} est convergente. Alors la suite des $\vec{x}_n$ est une suite de CAUCHY, admettant une suite partielle convergente; d'après le corollaire 2 du théorème 40, elle est convergente, et E est bien complet.

* A cause du théorème 12 la convergence normale subsiste si on remplace la norme de E par une norme équivalente,

Changement d'ordre des termes d'une série

Soit $\vec{u}_0+\vec{u}_1+\cdots+\vec{u}_n+\cdots$ une série de vecteurs de E . Changer l'ordre des termes de la série, c'est considérer une bijection $n \longrightarrow p_n$ de $\mathbb{N}$ sur $\mathbb{N}$, et remplacer la série donnée par la nouvelle série : $\vec{u}_{p_0}+\vec{u}_{p_1}+\cdots+\vec{u}_{p_n}+\cdots$.

Théorème 57 - SI la série $\vec{u}_0+\vec{u}_1+\cdots+\vec{u}_n+\cdots$, dans l'espace vectoriel normé E , est convergente ainsi que la série des normes $\|\vec{u}_0\|+\|\vec{u}_1\|+\cdots+\|\vec{u}_n\|+\cdots$, alors un changement d'ordre des termes n'altère pas la convergence de la série, et n'altère pas non plus sa somme.

Démonstration - Soit $\vec{S}$ la somme de la série donnée. Etant donné $\varepsilon>0$, il existe un entier m tel que $\sum_{n>m}\|\vec{u}_n\| \leqslant \frac{\varepsilon}{2}$. Il existe ensuite un entier m' tel que l'ensemble d'entiers $\{p_0, p_1, p_2, \cdots p_{m'}\}$ contienne l'ensemble $\{0,1,2,\ldots, m\}$. Alors, pour $n' \geqslant m'$, la somme partielle $\vec{u}_{p_0}+\vec{u}_{p_1}+\cdots+\vec{u}_{p_{n'}}$ de la série modifiée est égale à la somme partielle $\vec{u}_0+\vec{u}_1+\cdots+u_m$, augmentde de la somme d'un nombre fini de termes,dont les Indices sont tous $>m$. Alors la somme de ces termes résiduels a une norme majorée par $\sum_{m<n\leqslant p}\|\vec{u}_n\|$, où p est le plus grand des entiers $p_0, p_1, \ldots p_{n'}$, donc , par $\sum_{n>m}\|\vec{u}_n\|$, donc par $\frac{\varepsilon}{2}$. On a donc, pour $n'\geqslant m'$:

(II,14;4) $$\left\|(\vec{u}_{p_0}+\vec{u}_{p_1}+\cdots+\vec{u}_{p_{n'}})-(\vec{u}_0+\vec{u}_1+\cdots+\vec{u}_m)\right\| \leqslant \frac{\varepsilon}{2} .$$

Comme par ailleurs

(II,14;5) $$\left\|(\vec{u}_0+\vec{u}_1+\cdots+\vec{u}_m)-\vec{S}\right\| = \left\|\sum_{n>m}\vec{u}_n\right\| \leqslant \sum_{n>m}\|\vec{u}_n\| \leqslant \frac{\varepsilon}{2} ,$$

on a finalement, pour $n'\geqslant m'$:

(II,14;6) $$\left\|(\vec{u}_{p_0}+\vec{u}_{p_1}+\cdots+\vec{u}_{p_{n'}})-\vec{S}\right\| \leqslant \varepsilon ,$$

ce qui prouve bien que la série modifiée est convergente et de même somme $\vec{S}$.

Corollaire - SI une série à termes réels positifs est convergente, elle reste convergente et garde la même somme quand on modifie l'ordre de ses termes.

<u>Définition</u> Soit I un ensemble d'indices dénombrable et $(\vec{u}_i)_{i\in I}$ une famille de vecteurs d'un espace vectoriel normé E , indexée par I . On dit que <u>la série</u> $\sum_{i\in I}\vec{u}_i$ est <u>commutativement convergente et de somme $\vec{S}$</u> , si, quelle que soit la bijection $n \to p_n$ de $\mathbb{N}$ sur I , la série usuelle $\sum_{n=0}^{\infty}\vec{u}_{p_n}$ est convergente, et si sa somme est S , indépendamment de la bijection considérée.

Si les u_i sont réels ≥ 0 , appelons S la borne supérieure dans $\overline{\mathbb{R}}$ (donc finie ou égale à $+\infty$), des sommes $S_J = \sum_{i\in J} u_i$, correspondant à tous les sous-ensembles <u>finis</u> J de l'ensemble d'indices I . <u>Si S est finie, la série est commutativement convergente et de somme S .</u> En effet, soit $n \to p_n$ une bijection de $\mathbb{N}$ sur I . Alors $u_{p_0}+u_{p_1}+\cdots+u_{p_n} \leq S$. D'autre part, si $\varepsilon > 0$ est donné, il existe un sous-ensemble fini J de I tel que $S_J \geq S-\varepsilon$. Alors, si m est le plus petit entier tel que l'ensemble $\{p_0, p_1, \cdots p_m\}$ contienne l'ensemble J , on a, pour $n \geq m$: $u_{p_0}+u_{p_1}+\cdots+u_{p_n} \geq S_J \geq S-\varepsilon$.
Donc la série $u_{p_0}+u_{p_1}+\cdots+u_{p_n}+\cdots$ est bien convergente et de somme S . <u>Si S est $+\infty$, la série $\sum_{n=0}^{\infty} u_{p_n}$ est divergente, pour toute bijection $n \to p_n$ de $\mathbb{N}$ sur I</u> , car si elle était convergente pour une bijection particulière et de somme σ , on aurait $S_J \leq \sigma$ pour tout J, et $S \leq \sigma$, ce qui serait contraire à l'hypothèse. On peut dire dans ce cas que la série est commutativement divergente, et on écrit $\sum_{i\in I} u_i = +\infty$.
Pour une série à termes ≥ 0 , on n'aura pas besoin de parler de convergence commutative, ce sera automatique, on parlera simplement de convergence.
Si J est un sous-ensemble infini de I , on a toujours $S_J = \sum_{i\in J} u_i \leq S = \sum_{i\in I} u_i$.

Le Théorème 57 montre que, si $\sum_{i\in I}\vec{u}_i$ est une série de vecteurs d'un espace de BANACH E , et si la série des normes $\sum_{i\in I}\|\vec{u}_i\|$ est convergente, alors la série est elle-même commutativement convergente. On peut ajouter à ce résultat les compléments suivants :

1°/ S'il s'agit d'une série de nombres réels il y a une réciproque au résultat précédent :

<u>Une série de nombres réels ne peut être commutativement convergente que si elle est absolument convergente, c'est-à-dire si la série $\sum_{i\in I}|u_i|$ est convergente</u>. Considérons en effet une série de nombres réels.

Supposons choisie une bijection particulière de $\mathbb{N}$ sur L, ou, ce qui revient au même, supposons $I = \mathbb{N}$. Prenons alors les termes de la série dans l'ordre modifié suivant :

Nous prendrons d'abord autant de termes positifs qu'il en faut, dans l'ordre où ils sont indiqués initialement, pour dépasser sûrement 1, puis nous prendrons le premier terme strictement négatif; nous prendrons ensuite, parmi ceux qui n'ont pas encore été pris, autant de termes positifs qu'il faut, dans l'ordre où ils sont donnés, pour dépasser sûrement 2, puis nous prendrons le deuxième terme strictement négatif; nous prendrons ensuite, parmi ceux qui n'ont pas encore été pris, autant de termes positifs qu'il faut, dans l'ordre où ils sont donnés, pour dépasser sûrement 3, puis nous prendrons le troisième terme strictement négatif; et ainsi de suite. Cette opération est toujours possible si la série partielle des termes positifs est divergente. On aura dans ce cas trouvé un ordre pour lequel la série de nombres réels considérée est aussi divergente puisque, pour tout entier $n \geq 0$ il existera des sommes partielles qui dépasseront n . Il en résulte bien qu'une série de nombre réels ne peut être commutativement convergente que si la *Série* partielle des termes positifs est convergente, et de même la série partielle des termes négatifs. Nais alors, la série donnée est absolument convergente, ce qui est la réciproque énoncée.

2°/ <u>La même réciproque est valable si **E** est un espace vectoriel normé de dimension finie</u>. Comme un espace vectoriel de dimension k sur $\mathbb{C}$ est de dimension $2k$ sur $\mathbb{R}$, on peut supposer qu'il s'agit d'un espace vectoriel sur le corps des réels. Comme d'autre. part le résultat est indépendant de la norme choisie, puisque toutes les normes sont équivalentes (Théorème 13), on peut supposer que l'on a choisi une base, ce qui identifie E à un espace $\mathbb{R}^m$, et dans ce cas là, nous choisirons la norme $\|(x_1, x_2, \ldots x_m)\| = \sum_{1 \leq i \leq m} |x_i|$ sur $\mathbb{R}^m$. Si alors la série est commutativemenc convergente, on voit immédiatement que chacune des composantes de la série doit être uhe série commutativement convergente de nombres réels, et par conséquent, d'après ce que nous venons de voir sur les séries à termes réels, chacune des composantes de la série doit être absolument convergente. Il en résulte bien, par définition même de la norme choisie, que la série donnée est absolument convergente.

3°/ On peut au contraire démontrer, par un raisonnement très délicat, que si E est n'importe quel espace vectoriel normé de dimension infinie, on peut trouver dans E une série commutativement convergente mais non absolument convergente *

* Si donc $\sum_{n=0}^{\infty} \vec{u}_n$ est une série de vecteurs d'un Banach, on peut avoir les diverses circonstances suivantes :

a) elle est absolument convergente, donc commutativement convergente;
b) elle est commutativement convergente, mals non absolument (seulement si E est de dimension infinie);
c) elle est convergente, mais non commutativement;
d) elle est divergente.

Lorsqu'une série est commutativement convergente, sa **propriété** de convergence et sa **somme sont** donc en fait **indépendantes** de tout choix d'un ordre des termes; il est donc naturel d'imaginer que, comme dans le cas d'une série à termes positifs, on puisse donner la définition d'une série commutativement convergente et de sa somme, sans même avoir besoin de choisir une seule bijection de $\mathbb{N}$ sur I .

On démontre en effet ce qui suit :

Théorème **58** - La série $\sum_{i\in I} \vec{u}_i$ est commutativement convergente et de somme $\vec{S}$, si et seulement si, quelque soit $\varepsilon > 0$, il existe un sous-ensemble J fini de l'ensemble d'indices I tel que, pour tout sous-ensemble K fini de I , contenant J, on ait l'inégalité :

$$\|\vec{S}_K - \vec{S}\| = \|\sum_{i\in K} \vec{u}_i - \vec{S}\| \leqslant \varepsilon . \tag{II,14;7}$$

Les élèves pourront démontrer ce théorème à titre d'exercice.

Sommation par paquets d'une série commutativement convergente.

Théorème 59 - Supposons que l'ensemble dénombrable d'indices I soit donné comme réunion d'une famille de sous ensembles non vides disjoints

$$I_\alpha , \quad \alpha \in A \ : \ I = \bigcup_{\alpha\in A} I_\alpha , \qquad I_\alpha \cap I_\beta = \emptyset \text{ pour } \alpha \neq \beta .^{*}$$

Si alors la série $\sum_{i\in I_\alpha} \vec{u}_i$ de vecteurs d'un espace de Banach E est commutativement convergente, chacune des séries partielles $\sum_{i\in I_\alpha} \vec{u}_i$, est commutativement convergente, et si nous désignons par $\vec{S}_\alpha$ sa **somme, la série** $\sum_{\alpha\in A} \vec{S}_\alpha$ est commutativement convergente et de somme $\vec{S}$, autrement dit on a la formule de sommation par **paquets** :

$$\sum_{i\in I} \vec{u}_i = \sum_{\alpha\in A} \Big(\sum_{i\in I_\alpha} \vec{u}_i \Big) \tag{II,14;8}$$

Démonstration : Nous nous bornerons à démontrer ce théorème lorsque la série donnée est absolument convergente. Posons alors $M = \sum_{i\in I} \|\vec{u}_i\|$. On a alors bien **évidemment** $\sum_{i\in I_\alpha} \|\vec{u}_i\| \leqslant M$, ce qui prouve bien que chaque série $\sum_{i\in I_\alpha} \vec{u}_i$ est commutativement convergente (théorème **55** et **57**). En outre sa somme $\vec{S}_\alpha$ **vérifie l'inégalité** : $\|\vec{S}_\alpha\| \leqslant \sum_{i\in I_\alpha} \|\vec{u}_i\|$; par conséquent, si l'on considère un ensemble **fini quelconque** B d'éléments de l'ensemble **d'indices** A , **on a l'inégalité** $\sum_{\alpha\in B}\|\vec{S}_\alpha\| \leqslant \sum_{i\in \bigcup_{\alpha\in B} I_\alpha} \|\vec{u}_i\| \leqslant M$, **ce qui** prouve que la série $\sum_{\alpha\in A} \vec{S}_\alpha$ est elle **aussi commutativement** convergente. Il **reste donc à** montrer la formule : $\sum_{\alpha\in A} \vec{S}_\alpha = \vec{S}$.

* **A** est **fini** ou dénombrable, et chaque I_α fini **ou dénombrable**

Tout d'abord, étant donné $\varepsilon > 0$, on peut trouver un sous-ensemble fini J de l'ensemble d'indices I, tel que l'on ait la majoration :

(II, 14 ; 8) $$\sum_{i\in \complement J} \|\vec{u}_i\| \leq \frac{\varepsilon}{2}, \text{ donc } \|\vec{S}-\vec{S}_J\| \leq \frac{\varepsilon}{2}.$$

Soit alors $n \to p_n$ une bijection de $\mathbb{N}$ sur A, si A est dénombrable, ou d'un sous-ensemble $\{0,1,\cdots m\}$ de $\mathbb{N}$ sur A, si A est fini. Dans tous les cas, il existe un entier m tel que la réunion des ensembles $I_{p_0}, I_{p_1}, \dots I_{p_m}$, contienne **J.** Alors, pour $n \geq m$, la somme partielle $\vec{S}_{p_0} + \vec{S}_{p_1} + \cdots + \vec{S}_{p_n}$ est la somme de $\vec{S}_J$ et d'une série commutativement convergente, formée de termes $\vec{u}_i$, pour i dans le complémentaire de J.

On a donc l'inégalité :

(II, 14 ; 9) $$\|(\vec{S}_{p_0} + \vec{S}_{p_1} + \cdots + \vec{S}_{p_n}) - \vec{S}_J\| \leq \frac{\varepsilon}{2},$$

d'où l'on déduit l'inégalité :

(II, 14 ; 10) $$\|(\vec{S}_{p_0} + \vec{S}_{p_1} + \cdots + \vec{S}_{p_n}) - \vec{S}\| \leq \varepsilon,$$

qui prouve bien la formule $\sum_{\alpha\in A} \vec{S}_\alpha = \vec{S}$ annoncée.

Remarques 1°/ si $I = \mathbb{N}$, et si la série donnée $\sum_{n=0}^{\infty} \vec{u}_n$ est convergente, mais non commutativement, le même énoncé serait évidemment faux. Par exemple, si les u_n sont des nombres réels, et si la série $\sum_{n=0}^{\infty} u_n$ est semi-convergente la série partielle des termes ≥ 0 et la série partielle des termes < 0 sont toutes deux divergentes !

2°/ Etudions la réciproque de ce théorème, Soit $(I_\alpha)_{\alpha\in A}$ une partition de l'ensemble d'indices I ; supposons que pour tout a de A, la série $\sum_{i\in I_\alpha} \vec{u}_i$, soit commutativement convergente et de somme $\vec{S}_\alpha$, et supposons d'autre part que la série $\sum_{\alpha\in A} \vec{S}_\alpha$ soit commutativement convergente et de somme $\vec{S}$; peut-on affirmer que la série initiale $\sum_{i\in I} \vec{u}_i$ soit commutativement convergente et de somme $\vec{S}$?

Si elle est commutativement convergente, il résulte du théorème que nécessairement sa somme est $\vec{S}$; mais rien ne prouve que cette série soit commutativement convergente. Il suffit en effet de considérer l'exemple suivant, dans lequel I est égal à l'ensemble $\mathbb{Z}$ de tous les entiers, A à l'ensemble $\mathbb{N}$ de tous les entiers ≥ 0, et dans lequel, pour tout α, c'est-à-dire pour tout entier $n \geq 0$, I_α se compose de l'ensemble $\{+n, -n\}$ *. Alors, si nous considérons la série dans laquelle $u_i = i$, on voit que le processus $\sum_{\alpha\in A}\left(\sum_{i\in I_\alpha} u_i\right) = \sum_{n\in\mathbb{N}} (n-n)$

Z

* réduit à un élément si $n = 0$.

donne des séries commutativement convergentes avec la somme finale 0, alors que manifestement la série $\sum_{i\in\mathbb{Z}} i$ n'est pas commutativement convergente $\left(\sum_{i\in\mathbb{Z}} |i| = +\infty\right)$.

3°/ Par contre, si tous les u_i sont des nombres réels positifs, alors, si le processus : $\sum_{\alpha\in A}\left(\sum_{i\in I_\alpha} u_i\right)$ donne un resultat S, certainement la série $\sum_{i\in I} u_i$ est convergente (parce que toutes ses sommes partielles d'un nombre fini de termes sont bornées), et par conséquent de même somme S.

Donc, dans le cas de séries à termes positifs, si on convient, comme nous l'avons dit plus haut, d'appeler $+\infty$ la somme d'une série divergente, on voit que l'on a toujours, sans aucune hypothèse, l'égalité (II,14;8), les deux membres étant des nombres positifs finis ou infinis.

4°/ On en déduit enfin que, si E est un espace de Banach, et si la série $\sum_{\alpha\in A}\left(\sum_{i\in I_\alpha} \|\vec{u}_i\|\right)$ est convergente, alors on peut certainement affirmer que la série $\sum_{i\in I}\|\vec{u}_i\|$ est convergente, et que par conséquent la série $\sum_{i\in I}\vec{u}_i$ est commutativement convergente, donc le théorème 59 est encore applicable et la formule (II,14;8) valable.

Effet, sur une série, d'une application linéaire continue

Théorème 60 - Soient E et F deux espaces vectoriels normés, et L une application linéaire continue de E dans F. Si $\sum_{n=0}^{\infty}\vec{u}_n$ est une série convergente de E, alors la série $\sum_{n=0}^{\infty} L\cdot\vec{u}_n$ est convergente dans F, et en outre on a :

(II,14;11) $$\sum_{n=0}^{\infty}(L\cdot\vec{u}_n) = L\cdot\left(\sum_{n=0}^{\infty}\vec{u}_n\right).$$

Même énoncé avec la convergence commutative, et la convergence absolue si E et F sont des Banach, avec en outre

(II,14;12) $$\sum_{n=0}^{\infty}\|L\cdot\vec{u}_n\| \leq \|L\|\sum_{n=0}^{\infty}\|\vec{u}_n\|.$$

Démonstration

L étant linéaire, on a

(II,14;13) $$L\cdot\left(\sum_{n\leq m}\vec{u}_n\right) = \sum_{n\leq m}(L\cdot\vec{u}_n).$$

Lorsque m tend vers $+\infty$, $\sum_{n\leq m}\vec{u}_n$ tend vers la somme $\vec{S}$, donc le premier membre tend vers $L\cdot\vec{S}$, à cause de la continuité de L ; donc aussi le deuxième, ce qui signifie que la série

$\sum_{n=0}^{\infty} L \cdot \vec{u}_n$ est bien convergente et a bien pour somme $L \cdot \vec{S}$, d'où (II,14;11).

De (II,13;6) on déduit

(II,14;14) $$\sum_{n \leq m} \| L \cdot \vec{u}_n \| \leq \| L \| \sum_{n \leq m} \| \vec{u}_n \| \,.$$

D'où (II,14;12) par passage à la limite pour m tendant vers $+\infty$ (que les expressions écrites soient finies ou égales à $+\infty$). Si donc $\sum_{n=0} \vec{u}_n$ est absolument convergente, il en est bien de même de $\sum_{n=0}^{\infty} L \cdot \vec{u}_n$.

Produit de deux séries numériques. Effet d'une application bilinéaire continue sur deux séries

Rappelons que la plus simple des applications bilinéaires continues est le produit, application bilinéaire continue de $\mathbb{R} \times \mathbb{R}$ dans $\mathbb{R}$.

Or on a vu en mathématiques spéciales que, pour calculer le produit de deux séries, sous forme d'une série, il fallait faire des hypothèses de convergence absolue, C'est pourquoi ici il n'existera pas de théorème relatif seulement à des *séries* convergentes.

Théorème 61 - Soit B une application bilinéaire continue d'un produit $E \times F$ d'espaces de Banach dans un espace de Banach G. Soient $\sum_{i \in I} \vec{u}_i$ et $\sum_{j \in J} \vec{v}_j$ deux séries absolument convergentes d'éléments de E et F , alors la série $\sum_{(i,j) \in I \times J} B(\vec{u}_i, \vec{v}_j)$ est absolument convergente dans G , et on a la formule :

(II,14;15) $$\sum_{(i,j) \in I \times J} B(u_i, v_j) = B\Big(\sum_{i \in I} \vec{u}_i \,, \sum_{j \in J} \vec{v}_j\Big)$$

Démonstration - Appelons $\vec{U}$ et $\vec{V}$ les sommes $\sum_{i \in I} \vec{u}_i$ et $\sum_{j \in J} \vec{v}_j$. Soit K une partie finie quelconque de $I \times J$; il existe des parties finies L de I et M de J, telles que $L \times M$ contienne K , on a alors la majoration :

(II,14;16) $$\sum_{(i,j) \in K} \| B(\vec{u}_i, \vec{v}_j) \| \leq \sum_{(i,j) \in L \times M} \| B \| \; \| \vec{u}_i \| \; \| \vec{v}_j \|$$

$$= \| B \| \Big(\sum_{i \in L} \| \vec{u}_i \| \Big)\Big(\sum_{j \in M} \| \vec{v}_j \| \Big) \leq \| B \| \Big(\sum_{i \in I} \| \vec{u}_i \| \Big)\Big(\sum_{j \in J} \| \vec{v}_j \| \Big)$$

Ceci prouve que toutes les sommes partielles d'un nombre fini d'éléments de la série $\sum_{(i,j)\in I\times J} \| B(\vec{u}_i , \vec{v}_j) \|$ sont bornées et que par conséquent la série $\sum_{(i,j)\in I\times J} B(\vec{u}_i , \vec{v}_j)$ est bien absolument convergente. Nous pouvons alors lui appliquer le théorème de la sommation par paquets (théorème 59). Le produit $I \times J$ admet en effet une partition remarquable, Il est la réunion des ensembles $\{i\} \times J$ lorsque i parcourt I.

Donc $\sum_{(i,j)\in I\times J} B(\vec{u}_i , \vec{v}_j) = \sum_{i\in I} \left(\sum_{j\in J} B(\vec{u}_i , \vec{v}_j) \right)$.

Considérons donc d'abord la somme $\sum_{j\in J} B(\vec{u}_i , \vec{v}_j)$, pour i fixé. Quand $\vec{u}_i$ est fixé, l'application $\vec{v} \longrightarrow B(\vec{u}_i , \vec{v})$ est linéaire et continue de F dans G. Nous pouvons donc appliquer à la série $\sum_{j\in J} \vec{v}_j$ absolument convergente dans F, et à cette application linéaire continue, le théorème 60, et écrire la formule :

$$\sum_{j\in J} B(\vec{u}_i , \vec{v}_j) = B(\vec{u}_i , \sum_{j\in J} \vec{v}_j) = B(\vec{u}_i , \vec{V}) . \tag{II,14;17}$$

Mais nous pourrons de même considérer ensuite l'application $\vec{u} \longrightarrow B(\vec{u} , \vec{V})$, linéaire et continue de E dans G, et appliquer à la série $\sum_{i\in I} \vec{u}_i$ et à cette application le théorème 60. On en déduira cette fois ci :

$$\sum_{i\in I} B(\vec{u}_i , \vec{V}) = B\left(\sum_{i\in I} \vec{u}_i , \vec{V}\right) = B(\vec{U} , \vec{V}) \tag{II,14;18}$$

donc $\sum_{(i,j)\in I\times J} B(\vec{u}_i , \vec{v}_j) = B(\vec{U} , \vec{V})$, ce qui est (II,14;15).

Dans le cas ou I et J sont tous les deux l'ensemble $\mathbb{N}$ des entiers ≥ 0 il est assez courant, dans la théorie des séries de TAYLOR: de ranger dans un ordre bien déterminé, la série $\sum_{(i,j)\in \mathbb{N}\times\mathbb{N}} B(\vec{u}_i , \vec{v}_j)$ en posant :

$$\vec{w}_n = B(\vec{u}_n , \vec{v}_0) + B(\vec{u}_{n-1} , \vec{v}_1) + \cdots + B(\vec{u}_0 , \vec{v}_n) , \tag{II,14;19}$$

et en considérant la série produit comme étant le série $\sum_{n=0}^{\infty} \vec{w}_n$

<u>Remarque</u>. Il résulte de la démonctration qu'il n'est pas nécessaire de supposer les séries absolument convergentes, si l'on sait d'avance que la série $\sum_{i,j} B(u_i , v_j)$ est commutativement convergente.

<u>Conséquence</u> : Applications inversibles dans les espaces de BANACH.

<u>Définition</u> Soit une application linéaire continue d'un espace vectoriel normé E dans un espace vectoriel normé F ; on dit que u est inversible si c'est une bijection, et si la bijection réciproque u^{-1}, qui est manifestement linéaire, est aussi continue. Si $E = F$, cela signifie que, dans l'algèbre $\mathcal{L}(E;F)$,

l'élément u possède un inverse. On a entre les normes de ces deux bijections réciproques, compte tenu de la relation $u \circ u^{-1}$ = identité, et de la formule (II,13;7) la relation :

(II,14;20) $$1 = \|\text{identité}\| \leq \|u\|\,\|u^{-1}\|, \quad \text{ou} \quad \|u^{-1}\|^{-1} \leq \|u\|.$$

Les propriétés des séries dans les espaces de BANACH vont alors nous permettre de démontrer qu'une application suffisamment voisine d'une application inversible est elle aussi inversible; plus précisément :

Théorème **62** - Soit u une application linéaire continue inversible d'un espace de BANACH E dans un espace de BANACH F ,et soit v une application linéaire continue de E dans F vérifiant la majoration :

(II,14;21) $$\|v\| < \|u^{-1}\|^{-1}$$

Alors l'application linéaire continue $u+v$ de E dans F est elle aussi inversible; en outre $\|(u+v)^{-1}\| \leq \dfrac{1}{\|u^{-1}\|^{-1} - \|v\|}$

Démonstration : Prenons d'abord le cas où $F = E$, et $u = I$, application identique de E dans E . Alors v est une application linéaire continue de E dans E , astreinte à vérifier $\|v\| < 1$. Le c ne peut évidemment pas être remplacé par $\leq$, car pour $v = -I$, $I - I = 0$ n'est pas inversible.

Faisons d'abord un calcul purement formel. Nous calculerons $(I+v)^{-1}$ comme on calcule $(1+z)^{-1} = \frac{1}{1+z}$, par un développement en série géométrique $1 - z + z^2 - \dots$. Ecrivons :

(II,14;22) $$(I+v)^{-1} = I - v + v^2 - v^3 + \dots + (-1)^n v^n \dots .$$

Dans cette formule v^n est l'itéré $v \circ v \circ v \dots \circ v$, composé de n applicationsidentiques à v . Justifions maintenant ce calcul.

La série du 2ème membre est normalement convergente dans l'espace de BANACH $\mathcal{L}(E;E)$ (théorème 50), car

$\|v^2\| = \|v \circ v\| \leq \|v\|\,\|v\| = \|v\|^2$, et de même $\|v^n\| \leq \|v\|^n$,

et par hypothèse $\|v\| < 1$. Donc, d'après le théorème 55, elle est convergente et représente un élément w de $\mathcal{L}(E;E)$.

D'après le théorème 54, $u \longrightarrow uv$ est une application linéaire continue de $\mathcal{L}(E;E)$ dans $\mathcal{L}(E;E)$; on peut alors appliquer le théorème 60 à la série du 2ème membre de (II,14;22) et à cette application, ce qui donne, par un calcul terme à terme :

(II,14;23) $$w(I+v) = w + wv = (I - v + v^2 - v^3 + \dots) + (v - v^2 + v^3 - \dots) = I.$$

Le même raisonnement montre que $(I+v)w = I$. Alors, d'après ce que nous avons vu page 10 du chapitre 1, $I+v$ est une bijection, et w est sa bijection réciproque; comme $w \in \mathcal{L}(E;E)$, $I + v$ est bien inversible, et le théorème est démontré dans ce cas. On a en outre la majoration :

(II,14;23 bis) $$\|w\| = \|(I+v)^{-1}\| \leq 1 + \|v\| + \cdots + \|v\|^n + \cdots = \frac{1}{1-\|v\|} .$$

Remarquons que l'inversibilité de $I+v$ résulte aussi du thdoréme du point fixe. Soit $\vec{y}$ un élément de E. Existe-J-il un point $\vec{x}$ de E tel que $(I+v)\vec{x} = \vec{y}$ ou $\vec{x} + v\cdot\vec{x} = \vec{y}$ ou $\vec{x} = -v\cdot\vec{x} + \vec{y}$? Considérons l'application $f : \vec{x} \to -v.\vec{x}+\vec{y}$ de E dans E . C'est une contraction, car $\|-v\cdot(\vec{x}_1-\vec{x}_2)\| \leq \|v\|\,\|\vec{x}_1-\vec{x}_2\|$, avec $\|v\| < 1$. Donc il existe un $\vec{x}$ et un seul tel que $\vec{x} = f(\vec{x})$ ou $(I+v)\vec{x} = \vec{y}$, parce que E est complet (théorème 46). Donc $(I+v)$ est une bijection; comme elle est linéaire, sa bijection réciproque est linéaire. En outre, le théorème 46 bis montre que la solution $\vec{x}$ dépend continument de $\vec{y}$, donc $(I+v)^{-1}$ est continue, et $I + v$ est inversible.

Remarquons que l'expression de $\vec{x}$ à partir de $\vec{y}$ par la méthode des approximations successives, en partant de $\vec{x}_0 = \vec{0}$, donne $\vec{x}_1 = \vec{y}$, $\vec{x}_2 = -v\cdot\vec{x}_1 + \vec{y} = -v\cdot\vec{y} + \vec{y}$, $\vec{x}_3 = -v\cdot\vec{x}_2 + \vec{y} = \vec{y} - v\cdot\vec{y} + v^2\cdot\vec{y}$, ... et donne, à la limite, $\vec{x} = \vec{y} - v\cdot\vec{y} + v^2\cdot\vec{y} - v^3\cdot\vec{y} + \cdots$; c'est une nouvelle manière d'écrire $(I+v)^{-1} = I - v + v^2 - v^3 + \cdots$

Prenons maintenant le cas général, E et F quelconques, u inversible, et $\|v\| < \|u^{-1}\|^{-1}$. Comme u est inversible, on peut écrire :

(II,14;24) $$u + v = u(I + u^{-1}v) .$$

Ici, $u^{-1}v$ et $I+u^{-1}v$ sont des applications linéaires continues de E dans E ; $\|u^{-1}v\| \leq \|u^{-1}\|\,\|v\| < 1$, en vertu de l'hypothèse faite sur v . Alors, d'après le cas particulier que nous venons de démontrer, $I + u^{-1}v$ est inversible, et son inverse est donné par :

(II,14;25) $$(I+u^{-1}v)^{-1} = I - u^{-1}v + u^{-1}v\,u^{-1}v - u^{-1}v\,u^{-1}v\,u^{-1}v \; . \; .$$

Alors $u + v$ apparaît, d'après (II,14;22), comme composée de deux applications lnversibles, donc elle est aussi inversible, et son application réciproque est la composée des réciproques, avec inversion de l'ordre :

(II,14;26) $$(u+v)^{-1} = (I+u^{-1}v)^{-1}u^{-1}$$

$$= u^{-1} - u^{-1}v\,u^{-1} + u^{-1}v\,u^{-1}v\,u^{-1} - u^{-1}v\,u^{-1}v\,u^{-1}v\,u^{-1} + \cdots$$

On a en outre la majoration :

$$(\text{II},14;27)\qquad \|(u+v)^{-1}\| \leq \|(I+u^{-1}v)^{-1}\|\ \|u^{-1}\| \leq \frac{\|u^{-1}\|}{1-\|u^{-1}v\|}$$

$$\leq \frac{\|u^{-1}\|}{1-\|u^{-1}\|\ \|v\|} = \frac{1}{\|u^{-1}\|^{-1}-\|v\|}$$

Remarque 1 - Soit $\mathcal{U}$ l'ensemble des éléments inversibles de $\mathcal{L}(E;F)$, $\mathcal{U}^{-1}$ l'ensemble des éléments inversibles de $\mathcal{L}(F;E)$. Si $u_0 \in \mathcal{U}$, tout élément de $\mathcal{L}(E;F)$, appartenant à la boule ouverte de centre u_0 et de rayon $\|u_0^{-1}\|^{-1}$ est donc encore dans $\mathcal{U}$: donc $\mathcal{U}$ est un ouvert de $\mathcal{L}(E;F)$ et aussi $\mathcal{U}^{-1}$ par échange des rôles de E et F . Les inégalités précédentes permettent alors aisément de montrer que $u \longrightarrow u^{-1}$ est un homéomorphisme de $\mathcal{U}$ sur $\mathcal{U}^{-1}$. On pourra le faire à titre d'exercice. Nous le reverrons au théorème 27 du chapitre III.

Remarque 2 - Si $\mathcal{A}$ est une algèbre de Banach, la même démonstration prouve que si $\vec{x} \in \mathcal{A}$ est inversible, c'est-à-dire admet un inverse dans l'algèbre, et si $\vec{y}$ est un élément de $\mathcal{A}$ tel que $\|\vec{y}\| < \|\vec{x}^{-1}\|^{-1}$, $\vec{x}+\vec{y}$ est inversible, et que son inverse est donné par

$$(\text{II},14;27\,bis)\qquad (\vec{x}+\vec{y})^{-1} = \vec{x}^{-1} - \vec{x}^{-1}\vec{y}\vec{x}^{-1} + \vec{x}^{-1}\vec{y}\vec{x}^{-1}\vec{y}\vec{x}^{-1}\ldots$$

Si $\mathcal{U}$ est l'ensemble des éléments inversibles de $\mathcal{A}$, $\mathcal{U}$ est ouvert dans $\mathcal{A}$, et $\vec{x} \longrightarrow \vec{x}^{-1}$ est un homéomorphisme de $\mathcal{U}$ sur lui-même.

Critère de semi-convergence

Soit $\vec{u}_0, \vec{u}_1, \vec{u}_2, \cdots \vec{u}_n, \cdots$ une suite d'éléments d'un espace de BANACH E : on dit que cette suite est à variation bornée Si la série :

$$(\text{II},14;28)\qquad \|\vec{u}_1-\vec{u}_0\| + \|\vec{u}_2-\vec{u}_1\| + \cdots + \|\vec{u}_n-\vec{u}_{n-1}\| + \cdots$$

est convergente. La somme (II,14;28) s'appelle la variation totale de la suite. Cela entraîne naturellement que la suite considérée soit convergente; on sait en effet que sa convergence est équivalente à la convergence de la série $\vec{u}_0 + (\vec{u}_1-\vec{u}_0) + (\vec{u}_2-\vec{u}_1) + \cdots$; or l'hypothèse que nous avons faite est équivalente à la convergence absolue de cette série. On voit même en fait qu'une suite à variation bornée est à une suite convergente, ce qu'est une série absolument convergente à une série convergente. Si les u_n sont des nombres réels, et si la suite des u_n est monotone et bornée, elle est à variation bornée. En effet on a exactement, si par exemple elle est croissante et bornée:

$$(\text{II},14;29)\qquad |u_1-u_0| + |u_2-u_1| + \cdots = (u_1-u_0) + (u_2-u_1) + \cdots = -u_0 + \lim_{n\to\infty} u_n < +\infty.$$

Jordan a d'ailleurs démontré qu'inversement, si une suite de nombres rdels est à variation bornée, elle peut s'écrire sous la forme $u_n = a_n - b_n$, dans laquelle $a_0, a_1, a_2, \dots a_n, \dots$ et $b_0, b_1, b_2, \dots b_n, \dots$ sont deux suites croissantes et bornées *

Nous allons maintenant donner le plus Important critère de seml-convergence des séries, auquel tous les critères usuels peuvent pratiquement se ramener (notamment le théorème des séries alternées).

Théorème d'ABEL, 63 - Soient E, F, G, trois espaces de BANACH Soit $\vec{u}_n$ une suite de vecteurs de E, à variation bornée et tendant vers $\vec{0}$ pour $n \to \infty$, et $\vec{v}_n$ une suite de vecteurs de F à sommes partielles bornées, c'est-à-dire telle que les quantites

(II,14;30) $$\vec{\sigma}_{m,n} = \vec{v}_m + \vec{v}_{m+1} + \cdots + v_n \,, \quad n \geq m \,,$$

aient leurs normes bornées. Alors, si B est une application bilinéaire continue de $E \times F$ dans G, la série de terme général $\vec{w}_n = B(\vec{u}_n, \vec{v}_n)$ est convergente.

En outre, si on pose $U_m = \|\vec{u}_{m+1} - \vec{u}_m\| + \|\vec{u}_{m+2} - \vec{u}_{m+1}\| + \cdots$ et $V_m = \operatorname{Sup.}_{n \geq m} \|\vec{\sigma}_{m,n}\|$, le somme $\vec{S}$ et le reste $\vec{R}_m = \vec{w}_{m+1} + \vec{w}_{m+2} + \cdots$ admettent les majorations

(II,14;31) $$\|\vec{S}\| \leq \|B\| \, U_0 V_0 \,, \quad \text{et} \quad \|\vec{R}_m\| \leq \|B\| \, U_{m+1} V_{m+1} \;{}^{**}$$

Dans la plupart des applications, $E = F = G = \mathbb{C}$, corps des complexes, et B est le produit. Les $\vec{\sigma}_{m,n}$ sont surement bornées si la série $\sum_{n=0}^{\infty} \vec{v}_n$ est convergente.

* C'est à peu près évident. Il suffit de poser $a_0 = (u_0)^+$, $a_1 = (u_0)^+ + (u_1 - u_0)^+$, $a^* = (u_0)^+ + (u_1 - u_0)^+ + (u_2 - u_1)^+, \dots$ où $x^+ = x$ si $x \geq 0$ et 0 si $x < 0$. Ensuite : $b_0 = (u_0)^-$, $b_1 = (u_0)^- + (u_1 - u_0)^-$, $b_2 = (u_0)^- + (u_1 - u_0)^- + (u_2 - u_1)^-, \dots$ où $x^- = 0$ SI $x > 0$, et $|x|$ si $x \leq 0$. On a $x = x^+ - x^-$, et $|x| = x^+ + x^-$.

** Ceci tend bien vers 0 pourmtendant vers $+\infty$, parce que V_{m+1} reste borné, et que U_{m+1} tend vers 0 comme reste d'une série à termes ≥ 0 convergente.

<u>Démonstration</u> - On a

(II,14;32) $$\vec{S}_n = B(\vec{u}_0, \vec{v}_0) + B(\vec{u}_1, \vec{v}_1) + \cdots + B(\vec{u}_n, \vec{v}_n)$$

$$= B(\vec{u}_0, \vec{\sigma}_{0,0}) + B(\vec{u}_1, \vec{\sigma}_{0,1} - \vec{\sigma}_{0,0}) + \cdots + B(\vec{u}_n, \vec{\sigma}_{0,n} - \vec{\sigma}_{0,n-1})$$

$$= B(\vec{u}_0, \vec{\sigma}_{0,0}) + \left[B(\vec{u}_1, \vec{\sigma}_{0,1}) - B(\vec{u}_1, \vec{\sigma}_{0,0})\right] + \cdots + \left[B(\vec{u}_n, \vec{\sigma}_{0,n}) - B(\vec{u}_n, \vec{\sigma}_{0,n-1})\right]$$

$$= B(\vec{u}_0 - \vec{u}_1, \vec{\sigma}_{0,0}) + B(\vec{u}_1 - \vec{u}_2, \vec{\sigma}_{0,1}) + \cdots + B(\vec{u}_{n-1} - \vec{u}_n, \vec{\sigma}_{0,n-1}) + B(\vec{u}_n, \vec{\sigma}_{0,n}).$$

La terme différent des autres, $B(\vec{u}_n, \vec{\sigma}_{0,n})$, est major6 en norme par $\|B\| \|\vec{u}_n\| V_0$, donc il tend vers 0 pour n tendant vers $+\infty$, puisque $\vec{u}_n$ tend vers $\vec{0}$.

Il reste donc à montrer que :

(II,14;33) $$B(\vec{u}_0 - \vec{u}_1, \vec{\sigma}_{0,0}) + B(\vec{u}_1 - \vec{u}_2, \vec{\sigma}_{0,1}) + \cdots + B(\vec{u}_{n-1} - \vec{u}_n, \vec{\sigma}_{0,n-1})$$

a une limite pour n tendant vers $+\infty$. Cela revient à montrer que la série de terme général :

(II,14;34) $$\vec{W}_n = B(\vec{u}_n - \vec{u}_{n+1}, \vec{\sigma}_{0,n})$$

converge. Or la série des normes converge, puisque :

(II,14;35) $$\|\vec{W}_n\| \leq \|B\| \|\vec{u}_n - \vec{u}_{n+1}\| V_0 ,$$

et que la suite des $\vec{u}_n$ est supposée à variation bornée.

Cela démontre la convergence de la série donnée, en vertu du théorème 55, puisque G est supposé complet. La convergence une fois démontrée, on a la majoration :

(II,14;36) $$\|\vec{S}_n\| \leq \|B\| \|\vec{u}_n\| V_0 + \|B\| U_0 V_0 ;$$

donc, en passant à la limite pour n tendant vers l'infini, on a la majoration pour la somme $\vec{S} = \vec{w}_0 + \vec{w}_1 + \ldots + \vec{w}_n + \cdots$:

(II,14;37) $$\|\vec{S}\| \leq \|B\| U_0 V_0 .$$

Le même calcul, commencé au terme $\vec{w}_{m+1}$, et compte tenu de ce que $\vec{R}_m = \vec{w}_{m+1} + \vec{w}_{m+2} + \cdots$ donne la majoration du reste (II,14;31).

Exemples 1°) Théorème des séries alternées.

Considérons la série **réelle** $\sum_{n=0}^{\infty} (-1)^n u_n$, dans laquelle la suite $u_0, u_1, \dots u_n, \dots$ est $\geqslant 0$, décroissante, et converge vers 0 pour n tendant vers $+\infty$.

Dans ce cas, le théorème est applicable, avec $v_n = (-1)^n$, $U_m = u_m$, $V_m = 1$; il donne l'inégalité :

(II,14;38) $$|R_m| \leqslant u_{m+1} .$$

On sait même que, dans ce cas, le reste est du signe du premier terme négligé.

2°) Cas des séries trigonométriques.

Considérons la série $\sum_{n=0}^{\infty} u_n e^{ni\theta}$, dans laquelle on suppose que la suite complexe $u_0, u_1, \dots u_n, \dots$ est à variation bornée et converge vers 0 pour n tendant vers $+\infty$, et que θ est réel * . Prenons $v_n = e^{ni\theta}$.

On sait alors, d'après la formule de sommatinn de la série **géométrique,** que l'on a, si $\theta \neq 2k\pi$,

(II,14;39) $$e^{mi\theta} + e^{(m+1)i\theta} + \cdots + e^{ni\theta} = \frac{e^{(n+1)i\theta} - e^{mi\theta}}{e^{i\theta} - 1}$$

et par conséquent la majoration :

(II,14;40) $$|\sigma_{m,n}| \leqslant \frac{2}{|e^{i\theta} - 1|} .$$

Par contre, si $\theta = 2k\pi$, alors on a : $e^{mi\theta} + \dots + e^{ni\theta} = n - m + 1$, qui n'est pas bornée. On en déduit la conclusion que, pour $\theta \neq 2k\pi$, la série $\sum_{n=0}^{\infty} u_n e^{ni\theta}$ est bien convergente. On a les majorations :

(II,14;41) $$|S| \leqslant U_0 \frac{2}{|e^{i\theta} - 1|}, \qquad |R_m| \leqslant U_{m+1} \frac{2}{|e^{i\theta} - 1|}$$

Bien entendu on a un résultat analogue pour les séries $\sum_{n=0}^{\infty} u_n \cos n\theta$ et $\sum_{n=0}^{\infty} u_n \sin n\theta$.

En outre, pour cette dernière , la convergence a lieu même pour $\theta = 2k\pi$, puisqu'alors tous les termes de la série sont nuls.

* Si les u_n sont réels, vont en décroissant et tendent vers 0, et si $\theta = \pi$, on retrouve le théorème des séries alternées.

§ 15 EXEMPLES USUELS D'ESPACES FONCTIONNELS; CONVERGENCE SIMPLE ET UNIFORME

Espaces fonctionnels

On entend par espace fonctionnel un espace dont les éléments sont des fonctions, c'est-a dire des applications d'un ensemble dans un autre.

Soit E et F deux ensembles. Nous avons vu au chapitre 1, page 9, qu'on appelle F^E l'ensemble des applications de E dans F. S il se trouve que F possède certaines structures (espace vectoriel, espace métrique, etc....) alors il est en général possible d'introduire sur F^E une structure analogue.

1°/ Supposons que F soit un espace vectoriel sur un corps $\mathbb{K}$; alors bien évidemment F^E admet-aussi une structure d espace vectoriel sur $\mathbb{K}$. En effet si $\vec{f}$ et $\vec{g}$ sont deux applications de E dans F, et λ un scalaire, on peut définir la somme $\vec{f}+\vec{g}$ et le produit $\lambda\vec{f}$ comme de nouvelles applications de E dans F, par

$$(\vec{f}+\vec{g})(x) = \vec{f}(x)+\vec{g}(x) \text{ , pour tout } x \in E \text{ ;} \tag{II,15;1}$$

$$(\lambda\vec{f})(x) = \lambda\vec{f}(x) \text{ pour tout } x \in E \text{ .} \tag{II,15;2}$$

Ainsi nous avons défini sur F^E une loi d'addition, et une loi de multiplication par les scalaires de $\mathbb{K}$; on vérifie sans peine que ces lois satisfont à tous les axiomes qui font de F^E un espace vectoriel sur $\mathbb{K}$ * .

Par exemple, si F est le corps des scalaires $\mathbb{K}$, et si celui-ci est le corps des réels ou le corps des complexes, on voit que l'ensemble $\mathbb{K}^E$ des fonctions réelles ou complexes définies sur E est un espace vectoriel sur le corps des réels ou des complexes.

* Si E a deux éléments, F^E peut être identifié à $F^2 = F\times F$; la structure d'espace vectoriel ainsi obtenue sur $F\times F$ est celle d'espace vectoriel produit, indiquée page 112 (formule (II,13;22) et (II,13;23)).

2°/ Supposons maintenant que **F** soit un espace métrique. Si f et g sont deux applications de E dans F, on appelera distance de ces fonctions et on notera $d(f,g)$ la quantité définie par :

$$d(f,g) = \sup_{x \in E} d(f(x), g(x)). \tag{II,15;3}$$

Mais cette distance n'est pas nécessairement finie, la borne supérieure doit être prise dans la droite achevée $\overline{\mathbb{R}}$. Il en résulte qu'il n'est pas directement possible de mettre sur F^E une structure naturelle d'espace métrique. C'est pourquoi nous nous contenterons de considérer le sous-espace $(F^E)_b$ de F^E constitué par les applications bornées de E dans **F**. On dit qu'une application f de E dans F est bornée, si l'image $f(E)$ de E est une partie bornée de F. Si alors f et g sont deux applications bornées de **E** dans F, leur distance est nécessairement finie. (En effet, si par exemple $f(E)$ est contenu dans la boule de centre a et de rayon α de F, et si $g(E)$ est contenu dans la boule de centre b et de rayon β de F, donc à fortiori dans la boule de centre a et de rayon $\beta + d(a,b)$ de F, on voit que l'on a, pour tout x de E, l'inégalité $d(f(x), g(x)) \leq \alpha + \beta + d(a,b)$, et que par conséquent $d(f,g)$ est bien fini.)

Vérifions que l'application $d : (f, g) \longrightarrow d(f,g)$ est bien une fonction distance sur $(F^E)_b$ * . Nous devons vérifier les **3** propriétés (II,1;1). La symétrie est évidente. La positivité l'est également, car d'une part $d(f,g) \geq 0$, et d'autre part, si f et g dont distinctes, il existe au moins un point x de E tel que $d(f(x), g(x)) > 0$, et alors on a bien $d(f,g) > 0$

Reste à vérifier l'inégalite triangulaire.

Soient donc f, g, h, trois applications de E dans **F** ; pour tout x de E, on a :

$$d(f(x), h(x)) \leq d(f(x), g(x)) + d(g(x), h(x)) \leq d(f,g) + d(g,h) \tag{II,15;4}$$

Comme ceci est vrai pour tout x, on a :

$$d(f,h) \leq d(f,g) + d(g,h), \tag{II,15;5}$$

ce qui est bien l'inégalité cherchée.

* SI E a deux éléments, F^E peut s'identifier à $F^2 = F \times F$; la distance que nous venons de placer sur $(F^E)_b$ (qui est alors F^E lui-même) est l'une des distances que nous avons choisies pour le produit, page 112:

$$d((x_1, x_2), (y_1, y_2)) = \sup[d(x_1, y_1), d(x_2, y_2)].$$

<u>Remarque</u> : Les notions que nous venons d'introduire sont <u>métriques et non topologiques</u>. On démontre facilement=, si l'on remplace sur F la metrlque donnée par une métrique équivalente, on change d'abord complètement l'espace $(F^E)_b$, car le fait pour une fonction d'être bornée dépend de la métrique et non de la topologie(voir page 56) et en outre même si cet espace ne change pas, la nouvelle métrique que l'on place sur $(F^E)_b$ n'est pas équivalente à la première.

3°/ Supposons maintenant que F ait une structure d'espace vectoriel normé; alors on peut mettre sur l'espace $(F^E)_b$ une structure d'espace vectoriel normé en posant :

(II,15;6) $$|||f||| = \sup_{x \in E} \|\overrightarrow{f(x)}\| \quad *$$

Les mêmes méthodes que précédemment montrent que nous venons de définir là une norme; en outre la métrique associée à cette norme est la métrique que nous avions définie dans 2°/ : $d(f,g) = |||f-g|||$.

Par exemple, si F est le corps des réels ou le corps des complexes lui-même, on voit que l'espace des fonctions réelles ou complexes bornées définies sur un ensemble E est un espace vectoriel normé ** .

Les espaces vectoriels obtenus dans 1°/ et dans 3°/ sont pratiquement toujours de dimension Infinie. Supposons en effet que E soit un ensemble à n éléments, et pour simplifier, nommons 1, 2,...n, ces éléments. Alors l'ensemble F^E , dans le cas particulier ou $F = \mathbb{R}$, n'est autre que l'espace produit $\mathbb{R}^n$, et la norme introduite dans 3°/ n'est autre que :

(II,15;7) $$\|(x_1, x_2, \ldots x_n)\| = \sup(|x_1|, |x_2|, \ldots |x_n|).$$

Z

* Nous noterons par $\| \ \|$ la norme d'un élément de F, et par $||| \ |||$ la norme d'une application bornée de E dans F, pour éviter toute confusion entre les espaces vectoriels normés F et $(F^E)_b$ Alors $\|\vec{f}\|$ désignera <u>la fonction ≥ 0</u> : $x \to \|\overrightarrow{f(x)}\|$, tandis que $|||\vec{f}|||$ désignera la borne supérieure de cette fonction, un nombre 3 0

** Sur le corps des scalaires $\mathbb{K}$, la norme est le module, noté $| \ |$. Alors, si $f \in (\mathbb{K}^E)_b$, $|f|$ est la fonction ≥ 0 : $x \to |f(x)|$. Ici il n'y aura pas d'inconvénient à noter par $\|f\|$ la borne supkrieure de cette fonction, ou norme de f dans $(\mathbb{K}^E)_b$.

Toutes les fois que E contient une infinité d'éléments, (et les cas pratiques les plus importants sont ceux où E est la droite réelle ou un intervalle de la droite réelle) alors l'espace F^E antérieurement considéré est de dimension infinie.

Théorème 64 - Si F est un espace métrique complet, l espace métrique $(F^E)_b$ des applications bornées de E dans F est lui aussi complet *.

Démonstration : Soit $f_0, f_1, f_2, \ldots f_n, \ldots$ une suite de CAUCHY de $(F^E)_b$. D'après la définition même de la distance $d(f_m(x), f_n(x)) \leq d(f_m, f_n)$, on voit donc que, pour tout point x de E, la suite des points $f_n(x)$ est une suite de CAUCHY dans F.

Comme F est supposé complet, cette suite est convergente vers un point de F, que nous appellerons $f(x)$. Nous venons donc de définir une application f de E dans F.

Montrons d'abord que cette application est bornée. Soit $\varepsilon > 0$. Il existe un entier p tel que $m \geq p$, $n \geq p$, entraîne $d(f_m, f_n) \leq \varepsilon$ donc $d(f_m(x), f_n(x)) \leq \varepsilon$, pour tout x de E. Alors, pour x fixé dans E on peut dans l'inégalité précédente passer à la limite pour m tendant vers $+\infty$, et, étant donné la continuité de la fonction distance sur F, on a l'inégalité: $d(f(x), f_n(x)) \leq \varepsilon$, pour $n \geq p$. Il en résulte que, f_p étant bornée, l'ensembl[e] $f_p(E)$ est contenu dans une boule de centre a_p et de rayon R_p ; l'inégalité $d(a_p, f(x)) \leq d(a_p, f_p(x)) + d(f_p(x), f(x))$ montre alors que l'ersemble $f(E)$ est contenue dans la boule de centre a_p et de rayon $R_p + \varepsilon$, ce qui prouve bien que f est bornée. Alors f est elle aussi un élément de $(F^E)_b$; il nous reste à voir que f_n converge vers f pour n tendant vers $+\infty$. Or, ε étant donné, l'entier p étant choisi comme précédemment, l'inégalité $d(f(x), f_n(x)) \leq \varepsilon$ pour tout x de E, pour $n \geq p$, montre que $d(f, f_n) \leq \varepsilon$ pour $n \geq p$; ce qui montre bien cette convergence. Il en résulte que $(F^E)_b$ est un espace métrique complet.

Corollaire. Si F est un espace de BANACH, l'espace vectoriel normé $(F^E)_b$ est lui aussi un eppace de BANACH. En particulier l'espace $(\mathbb{K}^E)_b$ des fonctions, à valeurs réelles ou complexes, bornées, sur un ensemble E, est un espace de BANACH.

* D'après la note * page 138, cela contient comme cas particulier le théorème 44.

Il serait utile de pouvoir utiliser les résultats précédents dans **l'étude** de la convergence des suites de fonctions, et de pouvoir dire qu'une suite de fonctions f_n converge vers une fonction limite f lorsque les points f_n d'un certain espace topologique convergent, dans cet espace, vers le **point** f.

Convergence simple d'une suite de fonctions On dit qu'une suite de fonction f_n, c'est-à-dire d'applications d'un ensemble E dans un espace métrique F, converge simplement, pour n tendant vers $+\infty$, vers **une fonction limite** f, si, pour tout x de E, la suite des points $f_n(x)$ de F converge, pour n tendant vers $+\infty$, vers le point $f(x)$ de F.

Cela **s'écrit** sous la forme logique suivante :

$$(\mathrm{II},15;8) \qquad (\forall x \in E)(\forall \varepsilon > 0)(\exists m \in \mathbb{N})(\forall n \geq m) : d\big(f_n(x), f(x)\big) \leqslant \varepsilon \, .$$

On notera, comme nous l'avons vu au chapitre 1, page 35, que l'entier m, que nous venons de déterminer, dépend en fait **à** la fois de ε et de x. Peut-on construire sur l'ensemble F^E des applications de E dans F une structure **topologique, de** telle manière que des éléments f_n de cet espace' **topologique** convergent vers l'éldment f, si et seulement si les fonctions f_n convergent simplement vers la fonction f au sens que nous venons d indiquer ? C'est en effet possible, mais c'est un peu délicat. L'espace topologique obtenu n'est pas un espace métrisable, et nous n'en parlerons pas ici,

Convergence uniforme d'une suite de fonctions On dit que la suite des fonctions f_n converge uniformément vers la. fonction , pour n tendant vers $+\infty$, si l'entier m déterminé **dans** (II,15;8) peut être choisi indépendamment de x ; c'est-à-dire seulement en fonction de ε, autrement dit, si

$$(\mathrm{II},15;9) \qquad (\forall \varepsilon > 0)(\exists m \in \mathbb{N})(\forall x \in E)(\forall n \geqslant m) : d\big(f_n(x), f(x)\big) \leqslant \varepsilon \, .$$

Cela **s'écrit** encore

$$(\mathrm{II},15;10) \qquad (\forall \varepsilon > 0)(\exists m \in \mathbb{N})(\forall n \geqslant m) : \quad d(f_n, f) \leqslant \varepsilon \, .$$

Cela veut simplement dire que la distance de f_n **à** f tend vers 0 pour n tendant vers $+\infty$.

Il est bien évident que la convergence uniforme entralne la convergence simple, mais, comme nous allons le voir, la réciproque n'est pas vraie: la convergence uniforme est une propriété beaucoup plus forte que la convergence simple.

<u>Exemple 1</u> : **Considérons** la fonction réelle g d'une **variable réelle définie** par :

(II,15;11)
$$g(x) = \frac{1}{1+x^2} .$$

On appelle <u>translatée de cette fonction, par la translation h, celle que l'on obtient en translatant le graphique de</u> cette fonction, parallèlement à l'axe des x, de la translation h, c'est-à-dire par l'application $(x,y) \longrightarrow (x+h, y)$.

Autrement dit, cette nouvelle fonction $\tau_h g$ est telle que sa valeur en un pointa: soit la valeur de la fonction ancienne g au point $x - h$:

(II,15;12)
$$(\tau_h g)(x) = g(x-h) .$$

On a donc

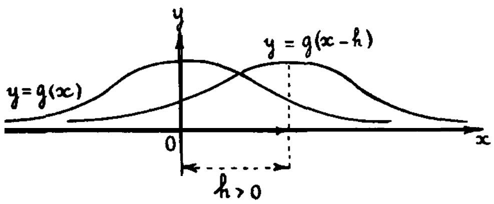
(II,15;13)
$$(\tau_h g)(x) = \frac{1}{1+(x-h)^2} .$$

Considérons alors la suite des translatées $\tau_n g$, $n \in \mathbb{N}$

On voit Immédiatement que cette suite de fonctions converge simplement vers la fonction identiquement nulle, pnur n tendant vers $+\infty$. On a en effet, pour x fixé :

(II,15;14)
$$\lim_{n \to \infty} \frac{1}{1+(x-n)^2} = 0 .$$

D'ailleurs cela revient à dire que la suite des valeurs de g aux points $x - n$ tend vers 0 pour n tendant vers $+ \infty$, ce qui est évident.

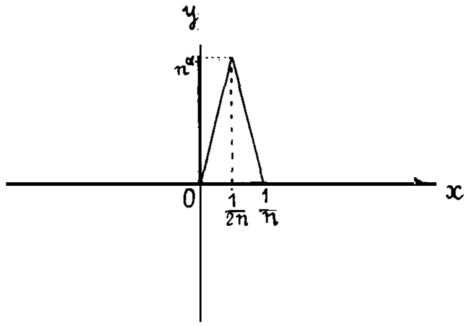


Cependant la suite des fonctions $\tau_n g$, ne converge pas uniformément vers 0 **,lorsque** n tend vers $+ \infty$, car la distance **de** $\tau_n g$ et de 0 est indépendante de n , elle est toujours égale à 1.

<u>Exemple 2</u> : Considérons la fonction f_n $(n \geqslant 1)$ réelle, d'une variable réelle, définie comme suit : elle est égale à 0 , pour x $\leqslant 0$ et $x \geqslant \frac{1}{n}$, elle est égale à n^α , $\alpha > 0$, pour $x = \frac{1}{2n}$, et dans chacun des Intervalles $\left[0, \frac{1}{2n}\right]$, $\left[\frac{1}{2n}, \frac{1}{n}\right]$, **elle** est linéaire affine *

* La fonction $y = ax + b$ est dite linéaire affine, ou affine, le mot **linéaire** devant être réservé conformément aux définitions générales relatives aux espaces vectoriels, à la fonction y $=$ a x .

Elle est représentée par le graphique suivant :



Malgré les apparences, la suite des fonctions f_n converge simplement vers la fonction 0 pour n tendant vers $+\infty$. En effet, quelque soit $x > 0$, pour n suffisamment grand, on a $\frac{1}{n} < x$, et par conséquent, pour n suffisamment grand, on a $f_n(x) = 0$; par ailleurs, pour $x \leq 0$, on a toujours $f_n(x) = 0$, d'où résulte bien ce que nous avons annoncé. Cependant, la distance de f_n et de 0 est égale à n^α, qui tend vers l'infini, donc f_n ns converge pas uniformément vers 0 pour n tendant vers $+\infty$. On voit que la notion de convergence simple, n'est pas en réalité aussi naturelle qu'elle le semblait à priori, car il parait assez paradoxal de dire que les deux fonctions précédentes convergent vers 0 pour n tendant vers $+\infty$. L'idée que l'on se fait de la convergence d'une suite de fonctions est plutôt celle de la convergence uniforme que celle de la convergence simple.

D'après ce que nous avons dit plus haut sur la relation entre la convergence uniforme et distance des fonctions, on voit que 1 espace topologique adapté à la convergence uniforme est l'espace métrique $(F^E)_b$. Dire qu'une suite d'applications bornées f_n de E dans F converge uniformément vers l'application bornée f de E dans F, c'est dire que la suite des points f_n de l'espace métrique $(F^E)_b$ converge vers le point f de cet espace métrique.

Autres emplois de l'expression: convergence uniforme

Considérons une suite d'éléments $x_n(\lambda)$ d'un espace métrique F, dépendant d'un paramètre λ parcourant un ensemble Λ. On dit que cette suite d'éléments converge vers une limite $x(\lambda)$ de F (dépendant évidemment elle aussi du paramètre λ), uniformément quand λ parcourt Λ, si

$$(\forall \varepsilon > 0)(\exists m \in \mathbb{N})(\forall \lambda \in \Lambda)(\forall n \geq m) : d(x_n(\lambda), x(\lambda)) \leq \varepsilon .$$

Cela revient exactement à écrire que la suite de fonctions $\lambda \longrightarrow x_n(\lambda)$, définies sur Λ, à valeurs dans F, converge uniformément vers la fonction $\lambda \longrightarrow x(\lambda)$. Cette notion se ramène donc exactement à la précédente, mais, psychologiquement, on ne se place pas exactement dans la même situation, *en* considérant une suite de fonctions ou en considérant une suite de points dépendant d'un paramètre λ.

Considérons maintenant une suite d'applications f_n de la droite réelle $\mathbb{R}$ dans un espace métrique F. Que" signifie l'expression : la suite f_n converge, pour n tendant vers $+\infty$, vers la fonction limite f, <u>uniformément sur tout intervalle borné de $\mathbb{R}$</u> ? Cela signifie que, quel que soit l'intervalle borné $[a,b]$ de $\mathbb{R}$, la suite des restrictions des f_n à cet intervalle converge uniformément vers la restriction de f ; autrement dit cela signifie :

(II,15;15) $$(\forall a\in\mathbb{R})(\forall b\in\mathbb{R}, b\geqslant a)(\forall\varepsilon>0)(\exists m\in\mathbb{N})(\forall x\in[a,b])$$
$$(\forall n\geqslant m): d(f_n(x), f(x))\leqslant\varepsilon .$$

Le nombre m qui intervient ici n'est alors pas fonction des , mais il est fonction d'une part de ε et d'autre part de l'intervalle $[a,b]$. On peut naturellement ici remplacer $\mathbb{R}$ et les intervalles bornés par $\mathbb{R}^n$ et les parties bornées de $\mathbb{R}^n$; plus généralement, étant donné un ensemble E, un espace métrique F, et une famille de parties $(A_i)_{i\in I}$ de E, on pourra parler d'une suite d'applications f_n de E dans F, qui, pour n tendant vers $+\infty$, converge vers l'application f, uniformément sur toute partie A_i de la famille. SI nous reprenons 1 exemple des $\tau_n g$ de la formule (II,15;13), on voit que la suite des fonctions $\tau_n g$ converge bien vers 0 ,uniformément sur tout intervalle borné de $\mathbb{R}$, et même uniformément sur toute demi-droite]-@, b[. En effet, pour $n\geqslant b$, on a, pour tout $x\leqslant b$, $\frac{1}{1+(x-n)^2}\leqslant\frac{1}{1+(b-n)^2}$, qui tend vers 0 pour n tendant vers $+\infty$. Si nous reprenons maintenant le 2ème exemple,(page 65), on voit que la suite des fonctions f_n converge vers la fonction 0 , uniformément sur le complémentaire de tout intervalle $[-\delta,+\delta]$, $\delta>0$, ayant pour centre l'origine. Elle ne converge uniformément sur aucun intervalle $]0,\delta[$.

Enfin, si E est un espace topologique, on dira que la suite des f_n converge vers f <u>localement uniformément</u> sur E, si tout point a de E admet un voisinage $\mathcal{U}$ sur lequel les Qn convergent uniformément vers f. Cela se traduira par la formule suivante :

II,15;15bis) $$(\forall a\in E)(\exists\mathcal{V}, \text{voisinage de } a)(\forall\varepsilon>0)(\exists m\in\mathbb{N})$$
$$(\forall x\in\mathcal{V})(\forall n\geqslant m): d(f_n(x), f(x))\leqslant\varepsilon .$$

Si E est localement compact, la convergence uniforme locale est équivalente a la convergence uniforme sur tout compact de E. Si en effet les f_n convergent vers f uniformément sur tout compact, tout point a de E admet un voisinage compact , sur lequel les f_n convergent uniformément, et la convergence est uniforme locale. Inversement supposons la convergence uniforme locale. Soit K un compact de E. Pour tout a de K, il existe un voisinage $\mathcal{V}_a$ de a sur lequel la convergence est uniforme. K est recouvert par un nombre fini des $\mathcal{V}_a$, d'où l'on déduit immédiatement aue la convergence est uniforme sur K.

Par ailleurs, tout ce que nous venons de dire sur la convergence simple ou uniforme d'une suite de fonctions s'étend à la convergence d un ensemble de fonctions, au sens indiqué page 60 . Par exemple, si, pour tout $t \in \mathbb{R}$, f_t est une application de E dans un espace métrique F : $x \longrightarrow f_t(x)$, on pourra parler de la convergence simple ou uniforme de f_t vers une application f de E dans F, lorsque t tend vers 0 par valeurs > 0 , ou vers $+\infty$, etc.,.

Espaces faisant intervenir à la fois la structure de E et la structure de F

Jusqu'à présent nous n'avons introduit des espaces fonctionnels qu'à partir de structures algébriques ou topologiques sur F Mais si à la fois F et F ont de telles structures, on peut introduire de nouveaux espaces; par exemple , si E et F sont tous les deux des espaces vectoriels sur le même corps $\mathbb{K}$, on peut considérer l'espace des applications linéaires de E dans F C'est un sous-espace vectoriel de l'espace F^E de toutes'les applications de E dans F . Si maintenant E et F sont des espaces topologlques, on peut introduire l'espace $(F^E)_c$ des applications continues de E dans F si F est métrique, cet espace n'est pas un sous-espace de $(F^E)_b$, car une application continue n'est pas nécessairement bornée; mais on pourra considérer le sous-espace $(F^E)_{bc}$ de $(F^E)_b$ formé des applications continues bornées de E dans F . Supposons enfin que E et F soient tous les deux des espaces vectoriels normés; il est alors possible d'introduire, comme nous l'avons fait, page 107, l'espace $\mathcal{L}(E;F)$ des applications linéaires continues de E dans F Il n'est pas un sous-espace de $(F^E)_{bc}$, parce qu'une application linéaire, à moins d'être Identiquement nulle, n est jamais bornée * . Mais appelons E_0 la boule unité de E ; si une application linéaire de E dans F est connue sur E_0, elle est connue partout, à cause de la formule d'homothétie, $u(\lambda \vec{x}) = \lambda u(\vec{x})$. D'ailleurs, étant donné une application de E_0 dans F on peut reconnaître si elle est ou non la restriction d'une application linéaire de E dans F ; il suffit en effet de la prolonger sur E , en posant : $u(\vec{x}) = \|\vec{x}\| u\left(\frac{\vec{x}}{\|\vec{x}\|}\right)$ pour $\|\vec{x}\| > 1$ et de vérifier si l'application obtenue est linéaire. D'après le théorème 47, pour qu'une application linéaire u de E dans F soit continue, il faut et il suffit que l'image par u de la boule unité E_0 soit bornée dans F . Par suite, on voit qu on peut identifier l'espace des applications linéaires de E dans F à un sous-espace de l'espace de tqutes les applications de E_0 dans F , à savoir le sous-

* Si en effet $\vec{a}$ est un vecteur de E tel que $u(\vec{a}) \neq \vec{0}$, alors la suite $u(n\vec{a}) = n\, u(\vec{a})$ n'est pas bornée, car $\|n\, u(\vec{a})\| = n \|u(\vec{a})\|$ tend vers $+\infty$ avec n .

espace des applications qui sont restrictions d'applications linéaires de E dans F ; et l'espace des applications linéaires continues de E dans F, c'est-à-dire $\mathcal{L}(E;F)$, à un sous-espace de l'espace $(F^{E_0})_b$ des applications bornées de E_0 dans F ; et, d'après la définition que nous avons donnée Pour la norme d'une application linéaire continue de E dans F, et Pour la norme d'une application bornée de E dans F, on voit que $\mathcal{L}(E;F)$ est ainsi exactement Identifié à un sous-espace vectoriel normé de l'espace vectoriel normé $(F^{E_0})_b$.

On voit même sans grande difficulté que $\mathcal{L}(E;F)$ est ainsi identifié à un sous-espace fermé de $(F^{E_0})_b$; si F est complet, $(F^{E_0})_b$ est complet, d'après le théorème 64; alors $\mathcal{L}(E;F)$, sous-espace fermé d'un espace complet, est complet (théorème 43), ce qui redonne une démonstration du théorème 50.

Continuité de la limite uniforme locale d'une suite de fonctions continues

Théorème 65 - Soient E et F deux espace métriques, et $f_0, f_1, f_2, \ldots f_n, \ldots$ une suite d'applications de E dans F, qui converge localement uniformément vers f. On suppose en outre que toutes les f_n sont continues en un point a de E ; alors la limite est elle aussi continue au point a. Si toutes les f_n sont partout continues, alors f est partout continue ; si la convergence est uniforme sur E, et si toutes les f_n sont uniformément continues sur E, alors f est uniformément continue sur E.

Démonstration : Soit $\mathcal{V}_a$ un voisinage de a, sur lequel la convergence est uniforme. D'après l'hypothèse de convergence uniforme, ε étant donné, il existe un entier m tel que l'on ait, pour tout x de $\mathcal{V}_a$, l'inégalité

(II,15;16) $$d\big(f_m(x), f(x)\big) \leq \frac{\varepsilon}{3}.$$

Le nombre m dtant ainsi choisi, la fonction f_m est supposée continue au pointa ; il existe donc un voisinage $\mathcal{V} \subset \mathcal{V}_a$ de a tel que, pour x dans $\mathcal{V}$, on ait :

(II,15;17) $$d\big(f_m(x), f_m(a)\big) \leq \frac{\varepsilon}{3}.$$

Dans ces conditions, pourx dans $\mathcal{V}$, on aura :

(II,15;18) $$d\big(f(x), f(a)\big) \leq d\big(f(x), f_m(x)\big) + d\big(f_m(x), f_m(a)\big) + d\big(f_m(a), f(a)\big) \leq \varepsilon,$$

ce qui prouve bien la continuité de f au point a.

Si les f_n sont partout continues, alors cette démonstration montre que f est partout continue *

* Pour ces deux premiers résultats, E peut être un espace topologique, non nécessairement métrisable.

supposons maintenant les f_n **uniformément** continues, et la convergence uniforme sur E. Alors, ε étant donné, et l'entier m ayant été choisi de manière que (II,15;16) soit vrai pour tout x de E, la fonction f_m est uniformément continue; il existe alors un nombre $\eta > 0$ tel que pour $d(x', x'') \leq \eta$, on ait $d(f_m(x'), f_m(x'')) \leq \frac{\varepsilon}{3}$.

On aura alors, pour $d(x', x'') \leq \eta$,

$$d\big(f(x'), f(x'')\big) \leq d\big(f(x'), f_m(x')\big) + d\big(f_m(x'), f_m(x'')\big) + d\big(f_m(x''), f(x'')\big) \leq \varepsilon, \tag{II,15;19}$$

ce qui prouve l'uniforme continuité de f

Remarque : Au contraire, une suite de fonctions continues f_n peut très bien converger **simplement** vers une fonction **discontinue** f. Ainsi, si nous posons $f_n(x) = x^n$, pour $0 \leq x \leq 1$ f_n est continue pour tout n, et converge simplement, pour n tendant vers $+\infty$, vers la fonction discontinue f égale à 0 pour $0 \leq x < 1$, à 1 pour $x = 1$.

Corollaires 1 - Dans l'espace $(F^E)_b$ des applications bornées de E dans F, muni de la métrique définie par (II,15.3), le sous-espace $(F^E)_{bc}$ des applications bornées continues de E dans F est fermé.

Ceci n'est pas autre chose qu'une autre forme de l'énoncé du théorème.

Corollaire 2 - Si F est complet, l'espace $(F^E)_{bc}$ des applications bornées continues de E dans F, muni de la métrique (II,15;3), est un espace métrique complet.

En effet il est fermé dans l'espace métrique complet $(F^E)_b$ (théorème 64) et il suffit d'appliquer alors le théorème 43.

Corollaire 3 - Si F est un espace de BANACH, l'espace $(F^E)_{bc}$ des applications continues bornées de E dans F, muni de la norme définie à la formule (II,15;6), est un espace de BANACH.

Un cas particulièrement important est le suivant : Prenons pour F le corps des scalaires $\mathbb{K}$, alors l'espace des fonctions réelles ou complexes, continues et bornées sur un espace métrique E, est un espace de BANACH.

En particulier, l'espace des fonctions réelles ou complexes continues et bornées, d'une variable réelle, est un espace de BANACH.

Si E est un compact (qui, dans de nombreuses applications pratiques, sera un intervalle fermé, borné, de la droite réelle), on sait, d'aprbs le thdorème 29, que toute fonction vectorielle, continue sur E , est nécessairement bornée * ; alors l'espace $(F^E)_c$ des fonctions continues sur un espace compact E , à valeurs dans l'espace de BANACH F , muni de la norme (II,15;6), est un espace de BANACH.

Application : quelques contre-exemples.

Nous sommes maintenant en mesure de donner des contre-exemples annoncés dans les paragraphes précédents.

1°/ Appelons $\mathscr{C}([a,b])$ l'espace vectoriel (de dimension infinie) des fonctions réelles ou complexes continues sur l'intervalle fermé, borné, $[a,b]$, $a<b$, de $\mathbb{R}$ ** . Nous allons montrer comment, sur cet espace, il est possible de définir plusieurs normes qui ne sont pas équivalentes les unes aux autres; c'est ce que nous avions annoncé après le théorème 13. Une première norme est celle qui a déjà été définie, par (II,15;6).

En voici maintenant une deuxième, utilisant la notion d'intégrale :

(II,15;20) $$N(\vec{f}) = \int_a^b |\vec{f}(x)|\, dx \quad ^{***}$$

* Nous l'avons dit seulement pour une fonction réelle; mais, si $\vec{f}$ est continue vectorielle, la norme $\|\vec{f}\|$ est une fonction continue réelle, donc bornée.

** $\mathscr{C}([a,b])$ peut aussi s'écrire $(\mathbb{K}^{[a,b]})_c$, $\mathbb{K}$ étant le corps des réels ou des complexes.

*** C'est bien un nombre ≥ 0 . L'inégalité triangulaire et la relation $N(\lambda \vec{f}) = |\lambda|\, N(\vec{f})$ sont évidentes. Si $\vec{f} \neq 0$, c'est-a-dire si $\vec{f} \not\equiv 0$, il existe au moins un point c de $[a, b]$ tel que $\vec{f}(c) \neq 0$. Il existe alors, à cause de la continuité de $\vec{f}$, un intervalle entourant C où $|\vec{f}(x)| > 0$, alors $N(\vec{f}) > 0$.

Ces deux normes ne sont pas équivalentes. On a la majoration :

(II,15;21) $$N(f) \leq (b-a)|||f||| .$$

Par contre, il n'existe évidemment aucune **majoration** du type $|||f||| \leq k\,N(f)$, où **K** serait une constante, **indépendante** de la fonction continue f . Considérons en effet le cas $[a,b]=[0,1]$, et la suite des fonctions f_n indiquée dans **l'exemple 2 de la** page 142 , Pour ces fonctions, on a les normes suivantes :

(II,15;22) $$|||f_n||| = n^{\alpha} \quad , \quad N(f_n) = \frac{1}{2} n^{\alpha-1} \quad , \qquad \frac{|||f_n|||}{N(f_n)} = 2n \ ,$$

ce qui prouve notre assertion. Si par exemple $0 < \alpha < 1$, f_n tend vers 0 pour n tendant vers $+\infty$, pour la **norme** N, mais non pour la norme $|||\ |||$. On démontre facilement qu'il **existe,** sur l'espace **vectoriel** $\mathcal{C}([a,b])$, une **infinité** d'autres normes deux à deux non équivalentes.

2°/ Montrons maintenant que, sur l'espace $\mathcal{C}([0,1])$ **muni** de la norme $|||\ |||$, la boule unité n'est pas compacte.

Il suffit pour cela de considérer la suite des **fonctions** f_n définies par $f_n(x) = x^n$ (voir remarque page **147**).

(II,15;23) $$f_n(x) = \sin nx \ ; \ |||f_n||| = 1 \text{ pour } n \geq 1 \ .$$

Comme cette suite converge simplement vers la fonction égale a0 pour $0 \leq x < 1$, à 1 pour $x = 1$, **une** suite partielle qui serait uniformément convergente ne pourrait converger uniformément que vers cette **limite; comme cette** limite est discontinue, cela ne peut pas se produire. Nous avons donc bien trouvé une suite appartenant à la boule unité de $\mathcal{C}([0,1])$, et dont aucune suite partielle n'est convergente, donc cette boule unité n'est pas compacte, d'après le théorème de Weierstrass Bolzano (théorème 25). Ainsi $\mathcal{C}([0,1])$ n'est pas localement compact. Ceci est vrai, comme nous l'avons vu, pour tout espace vectoriel normé de dimension infinie (théorème 45 bis).

Théorème 66 - Soient E et F deux espaces métriques, A une de E et $f_0, f_1, f_2, \ldots f_n$, une suite d'applications de A dans F, convergeant uniformément vers f. Soit a un point de E adhérent à A *. Si, pour chaque n, $f_n(x)$ a une limite quand x tend vers a par valeurs dans A, et si F est complet, alors $f(x)$ a une limite quand x tend vers a par valeurs dans A, et en outre

$$\lim_{\substack{x \to a \\ x \in A}} f(x) = \lim_{n \to \infty} \left[\lim_{\substack{x \to a \\ x \in A}} f_n(x) \right]. \tag{II,15;25}$$

Démonstration. Considérons, pour tout n, la fonction $\tilde{f}_n$ définie sur $A \cup \{a\}$ par la formule $\tilde{f}_n(x) = f_n(x)$ pour $x \neq a$, $\tilde{f}_n(a) = \lim_{\substack{x \in A \\ x \to a}} f_n(x)$. D'après sa définition même, $\tilde{f}_n$ est continue au point a. D'autre part, quel que soit $\varepsilon > 0$, il existe un entier p tel que $m \geq p$, $n \geq p$, entraîne

$$d(f_m(x), f_n(x)) \leq \varepsilon, \quad \text{et aussi} \quad d(f(x), f_n(x)) \leq \varepsilon,$$

pour $x \in A$, $x \neq a$. En faisant tendre x vers a, on a, à la limite, $d(\tilde{f}_m(a), \tilde{f}_n(a)) \leq \varepsilon$. Cela prouve que les $\tilde{f}_n(a)$ forment une suite de CAUCHY dans F ; comme F est supposé complet, ils ont une limite, que nous appellerons $\tilde{f}(a)$. Alors $\tilde{f}$, définie par $\tilde{f}(x) = f(x)$ pour $x \in A$, $x \neq a$, et prenant la valeur $\tilde{f}(a)$, que nous venons de définir, au point a, est une fonction sur $A \cup \{a\}$, à valeurs dans F. Lorsque n tend vers $+\infty$, $\tilde{f}_n(x)$ tend vers $\tilde{f}(x)$ pour $x \neq a$, mais $\tilde{f}_n(a)$ tend vers $\tilde{f}(a)$, par définition de $\tilde{f}(a)$; donc $\tilde{f}_n$ converge simplement vers $\tilde{f}$. Mais elle converge en outre uniformément; car, pour $m \geq p$, $n \geq p$, on a,

* Aucune hypothèse n'est faite relativement à l'appartenance de a à A.

pour $x \neq a$, $d(\tilde{f}(x), \tilde{f}_n(x)) = d(f(x), f_n(x)) \leq \varepsilon$, **mais** aussi $d(\tilde{f}_m(a), \tilde{f}_n(a)) \leq \varepsilon$, donc, en faisant tendre m **vers** $+\infty$ **et passant à** la limite $d(\tilde{f}(a), \tilde{f}_n(a)) \leq \varepsilon$, donc finalement $d(\tilde{f}, \tilde{f}_n) \leq \varepsilon$.

Comme alors les $\tilde{f}_n$ -sont continues au point a, le théorème 65 montre que f aussi est continue au point a, c'est-a-dire que $f(x)$ à une limite quand x **tend vers** a en restant dans A, et que cette limite est $\tilde{f}(a) = \lim\limits_{n \to \infty} \tilde{f}_n(a)$ $= \lim\limits_{n \to \infty} \left[\lim\limits_{\substack{x \to a \\ x \in A}} f_n(x) \right]$, ce qui prouve le théorème **énoncé.**

<u>Exemple</u> - SI les f_n et f sont des fonctions réelles continues **sur la** droite **réelle** $\mathbb{R}$, **si** les f_n **convergent uniformément** vers f pour n tendant vers $+\infty$ **, et si, pour** tout n, $\lim\limits_{x \to +\infty} f_n(x) = \ell_n$ **existe, alors** $\lim\limits_{x \to +\infty} f(x) = \ell$ existe aussi, et les ℓ_n convergent vers ℓ **.**

Séries de fonctions à valeurs dans un espace vectoriel normé

Soient $\vec{u}_0$, $\vec{u}_1, \vec{u}_2, \dots \vec{u}_n, \dots$ des applications d'un ensemble E dans un espace vectoriel normé F. Alors on peut considérer **la série** $\sum\limits_{n=0}^{\infty} \vec{u}_n$ **.** Dire que cette série est simplement <u>convergente</u> et de somme $\vec{S}$, application de E dans F, veut **dire** que, pour tout point x de E, la série des vecteurs de F : $\sum\limits_{n=0}^{\infty} \vec{u}_n(x)$ est convergente dans l'espace vectoriel normé F, et de **somme** $\vec{S}(x)$ **.Cela revient** exactement **à** dire que la suite des sommes partielles $\vec{S}_n = \vec{u}_0 + \vec{u}_1 + \dots + \vec{u}_n$ est une suite de **fonctions** sur E **, à** valeurs dans F **, convergeant** simplement vers $\vec{S}$. On **dira** que la **série** est **uniformément** convergente si la suite des sommes partielles $\vec{S}_n$ est uniformément convergente.

Par contre, si F est un Banach en ce qui concerne la notion de convergence absolue ou de convergence normale, il peut y avoir une certaine **ambiguité** dans les termes. Quand il s'agissait de séries de vecteurs d'un espace vectoriel normé, nous avons **employé** indifféremment les termes "série absolument convergente" ou "série normalement convergente **".** On fait une distinction lorsqu'il s'agit d'une série de fonctions.

Z

On dira que la série est simplement absolument convergente, si, pour tout x, la série des normes dans l'espace de Banach F, c'est-a-dire $\sum_{n=0}^{\infty} \|\vec{u}_n(x)\|$, est convergente; cela signifie que la série de fonctions $\sum_{n=0}^{\infty} \|\vec{u}_n\|$ * définies sur E, à valeurs réelles positives, est une série simplement convergente. On dira que la série est normalement convergente, si la série des normes dans l'espace vectoriel normé $(F^E)_b$, à savoir $\sum_{n=0}^{\infty} \|\!|\vec{u}_n\|\!|$, est une série à termes positifs convergente. C'est là évidemment la notion la plus forte. Toute série normalement convergente est simplement absolument convergente, sans que la réciproque soit nécessairement vraie. Et, dans l'un quelconque de ces cas, la série est simplement convergente (F est supposé complet). En outre. une série normalement convergente est uniformdment convergente. La convergence normale est même le critère le plus important de convergence uniforme d'une série de fonctions vectorielle On l'exprime encore souvent de la manière suivante :
SI l'on a une série de fonctions $\sum_{n=0}^{\infty} \vec{u}_n$, définies sur E, à valeurs dans un Banach, et s'il existe une suite de constantes réelles $a_n \geq 0$ telles que $\|\vec{u}_n(x)\| \leq a_n$ pour tout x de E, et $\sum_{n=0}^{\infty} a_n < \infty$, alors cette série est normalement, donc uniformément convergente. Il peut être utile de possèder aussi des critères de semi-convergence uniforme. D'une façon générale, quand on a une série de fonctions $\sum_{n=0}^{\infty} \vec{u}_n$, pour démontrer que cette série est uniformément convergente, on pourra commencer par démontrer que, pour touts de E, elle est convergente. On calculera alors le reste $\vec{R}_m = \vec{u}_{m+1} + \vec{u}_{m+2} + \dots$; de la série et on montrera que la suite de fonctions $\vec{R}_m$ converge uniformément vers 0 lorsque m tend vers $+\infty$. On utilisera à cet effet des majorations convenables du reste.

* Rappelons que $\|\vec{u}_n\|$ est la fonction $x \to \|\vec{u}_n(x)\|$, et que $\|\!|\vec{u}_n\|\!| = \sup_{x \in E} \|\vec{u}_n(x)\|$.

Donnons un exemple particulièrement interessant dans la **théorie** des **séries** de Taylor.

Théorème 67 - Si une série de Taylor (a coefficients complexes) $\sum_{n=0}^{\infty} a_n x^n$ est convergente au point $x = R$ de son cercle de convergence, alors, dans tout l'intervalle $[0, R]$, elle est uniformément convergente. En particulier, sa somme est une fonction continue sur l'intervalle $[0, R]$.

Démonstration On a en effet la formule $a_n x^n = (a_n R^n) \left(\frac{x}{R}\right)^n$.

Nous pouvons alors, pour tout x de $[0, R[$ appliquer le critère de convergence d'Abel * . Si en effet, nous posons $u_n = \left(\frac{x}{R}\right)^n$, $v_n = a_n R^n$, l'application bilinéaire B étant le produit, on voit que ces quantités satisfont bien aux conditions d'Abel, et la formule (II,14;31), donne une majoration du reste :

$$U_{m+1}(x) = \left|\left(\frac{x}{R}\right)^{m+1} - \left(\frac{x}{R}\right)^{m+2}\right| + \left|\left(\frac{x}{R}\right)^{m+2} - \left(\frac{x}{R}\right)^{m+3}\right| + \cdots \tag{II,15;26}$$

$$= \left(\left(\frac{x}{R}\right)^{m+1}\left(1 - \frac{x}{R}\right)\right)\left(1 + \frac{x}{R} + \left(\frac{x}{R}\right)^2 + \cdots\right) = \left(\frac{x}{R}\right)^{m+1}$$

$$V_{m+1}(x) = V_{m+1} = \operatorname*{Sup}_{n \geq m+1} \left|a_{m+1} R^{m+1} + \ldots + a_n R^n\right|$$

$$R_{m+1}(x) \leqslant \left(\frac{x}{R}\right)^{m+1} \operatorname*{Sup}_{n \geqslant m+1} \left|a_{m+1} R^{m+1} + \cdots + a_n R^n\right| \tag{II,15;27}$$

* Il peut **paraître absurdé** d'appliquer un critère adapté aux séries **semi-convergentes** dans la région où la série est absolument convergente, majorée par une série géométrique! Mais nous voulons démontrer une convergence uniforme , et comme la série n'est pas **supposée** absolument convergente **pour** $x = R$, **il** n'y aura pas d'autre moyen.

On ne peut plus appliquer le même critère d'ABEL pour $x = R$, parce que la suite des $(\frac{x}{R})^n = 1$, qui est toujours à variation bornée puisque constante, ne converge pas verso; mais, de toute façon la série est supposée convergente et la majoration du reste est encore valable, par définition même du reste. Il en résulte que cette majoration est valable pour tout x de l'intervalle $[0, R]$; comme alors, lorsque m tend vers $+\infty$, la quantité V_{m+1} tend vers 0 (critère de CAUCHY relatif à la série numérique convergente $\sum_{n=0} a_n R^n$), et que $U_{m+1} \leq 1$, on voit bien que le reste R_{m+1} converge uniformément vers 0 lorsque m tend vers $+\infty$, et ceci démontre la convergence uniforme de la série de TAYLOR. La continuité de la somme de la série résulte- alors du théorème 65. On a déjà vu en mathématiques spéciales quelques applications remarquables de ce théorème Par exemple on a le développement de TAYLOR

(II,15;28) $$\log(1+x) = \frac{x}{1} - \frac{x^2}{2} + \cdots + (-1)^{n-1}\frac{x^n}{n} + \cdots \quad \text{pour } |x| < 1 .$$

Or, pour $x = 1$ le théorème des séries alternées montre que la série est'encore convergente; elle représente donc une fonction continue dans l'intervalle $[0, 1]$, et sa somme est $\lim\limits_{\substack{x \to 1 \\ x < 1}} \log(1+x) = \log 2$ pour $x = 1$, on en déduit la formule :

(II,15;29) $$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} + \cdots + \frac{(-1)^{n-1}}{n} + \cdots \; *$$

De la même manière, en considérant le développement de la fonction

(II,15;29) $$\operatorname{Arc\,tg} x = \frac{x}{1} - \frac{x^3}{3} + \frac{x^5}{5} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + \cdots$$

et en faisant le même raisonnement pour $x = 1$, on voit que l'on a la formule :

(II,15;30) $$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} + \cdots + (-1)^n \frac{1}{2n+1} + \cdots$$

* Dans ce cas particulier on peut appliquer directement le théorème des séries alternées. Pour tout x de $[0, 1]$, ce théorème est applicable, donc le reste $R_{m+1}(x)$ est majoré en module par le premier terme négligé $\frac{x^{m+1}}{m+1} \leq \frac{1}{m+1}$, donc converge bien uniformément vers 0 .

§ 16 PRODUITS INFINIS DE NOMBRES OU DE FONCTIONS RÉELS OU COMPLEXES

Définition : Soit u_0, u_1, u_n, ... une suite de-nombres réels ou complexes. Cn dit que le produit infini $\prod_{n=0} u_n$ est convergent, si la suite des produits partiels $\Pi_n = \prod_{0\leq m\leq n} u_m$ converge, pour n tendant vers $+\infty$, vers un nombre fini $\neq 0$.

Dans tous les autres cas, le produit est dit divergent. Il peut paraître paradoxal de-considérer comme divergent un produit dans lequel les produits partiels Π_n convergent vers 0 ; nous en verrons dans la suite de nombreuses raisons. Si un des u_n est nul, le produit est donc surement divergent. Si tous les u_n sont réels ≥ 1, les Π_n forment une suite croissante, donc ont une limite finie ou égale à $+\infty$; dans ce cas, si le produit est divergent, on écrit $\prod_{n=0}^{\infty} u_n = +\infty$. De même, toutes les fois que les Π_n convergent vers 0, quoique le produit soit divergent, on écrit $\prod_{n=0}^{\infty} u_n = 0$.

Thciorème 68 - Pour qu'un produit infini $\prod_{n=0}^{\infty} u_n$ soit convergent, il est nécessaire que son terme général u_n tende vers 1 lorsque n tend vers $+\infty$.

Démonstration En effet, si le produit est convergent, et si sa valeur est $\Pi \neq 0$, alors Π_n et Π_{n-1} convergent tous deux vers Π, ce qui prouve bien que leur quotient u_n tend vers $\frac{\Pi}{\Pi} = 1$ pour n tendant vers $+\infty$. Nous remarquons bien que le résultat ne subsiste pas pour un produit infini, dont les produits partiels Π_n convergent vers 0 ; par exemple, si nous considérons le produit infini(?) $\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)\cdots\left(\frac{1}{n}\right)\cdots$, bien évidemment les produits partiels convergent vers 0, et le terme général $\frac{1}{n}$ ne converge pas vers 1.

Remarque- Si un produit infini est convergent, on appellera reste R_m le produit $\prod_{n\geq m+1} u_n$; R_m tend vers 1 pour m tendant vers + co .

Pour qu'un produit Infini soit convergent, il faut et il suffit qu'il vérifie le critère de CAUCHY : $\frac{\Pi_n}{\Pi_m}$, doit tendre vers 1 lorsque m et n tendent vers $+\infty$ (tous les termes u_n étant supposés $\neq 0$). Il est d'abord Cvldent qu'un produit convergent satisfait à ce critère. Réciproquement, supposons qu'un produit Infini vérifie le critère de CAUCHY. Il existe donc un entier p tel que $n \geqslant p$ entraîne $|\Pi_p - \Pi_n| \leqslant |\Pi_p|$. Donc tous les $|\Pi_n|$ sont bornés. Soit M leur borne supérieure. Pour ε donné, $0 \leqslant \varepsilon \leqslant \frac{M}{2}$, on peut alors déterminer q tel que $m \geqslant q$, $n \geqslant q$ entraîne $|\Pi_m - \Pi_n| \leqslant \frac{\varepsilon}{M}|\Pi_m| \leqslant \varepsilon$, donc les Π_n forment une suite de CAUCHY dans le corps complexe $\mathbb{C}$ .comme celui-ci est complet, les Π_n ont une limite Π . Alors on a,en faisant tendre n vers $+\infty$, dans la majoration ci-dessus, $|\Pi_m - \Pi| \leqslant \frac{\varepsilon}{M}|\Pi_m| \leqslant \frac{1}{2}|\Pi_m|$, donc $\Pi \neq 0$, et le produit infini est convergent.

Produit infini et série des logarithmes Il est évidemment tentant, pou étudier la convergence ou la divergence d'un produit infini, de prendre les logarithmes des termes, ce qui remplace le produit infini par une série. Si les termes sont réels, on peut prendre le logarithme dès que $u_n > 0$.

Or, s'il n'en est pas ainsi à partir d'une certaine valeur de n, cela signifie que le terme général u_n ne tend pas vers 1, pour n tendant vers $+\infty$, alors on sait tout de suite que le produit est divergent et l'étude est terminée.

Si RU contraire u_n tend vers 1 pour n tendant vers $+\infty$, alors, à partir d'un certain terme, tous les u_n sont strictement positifs, et il est possible de prendre leur logarithme. Supposons maintenant que les u_n soient complexes. On sait qu'il est assez délicat de prendre le logarithme d'un nombre complexe, car tout nombre complexe a une infinité de logarithmes. Si l'on pose $z = r e^{i\theta}$, on a la formule générale : $\log z = \log r + i\theta$; θ n'est défini qu'à un multiple près de 2π . Cependant, supposons que z varie dans le demi-plan $x = \mathcal{R} z > 0$. Alors il est possible de choisir son argument entre $-\frac{\pi}{2}$ et $+\frac{\pi}{2}$, et de définir son logarithme par la formule correspondante; le logarithme ainsi défini est une fonction continue. On dira que c est la détermination principale du logarithme et on le notera $\mathcal{L}og$ *

* On peut même définir une détermination principale du logarithme dans le complémentaire dans le plan complexe, de la demi-droite réelle $\leqslant 0$. Alors $-\pi < \mathcal{A}rg\, z < +\pi$, et $-i\pi < \mathcal{I}m(\mathcal{L}og\, z) < i\pi$.

En particulier si v est un nombre tel que $|v| < 1$, **alors** $1 + v$ est dans le demi-plan précédent, et le logarithme précédemment **défini** peut se représenter par le développement de **TAYLOR** :

$$(\mathrm{II},16;1) \qquad \mathcal{L}og(1+v) = \frac{v}{1} - \frac{v^2}{2} + \ldots \quad + (-1)^{n-1}\frac{v^n}{n} + \ldots$$

Soit alors $\prod_{n=0}^{\infty} u_n$ un produit infini ue nombres complexes. **Si** u_n **ne** tend pas vers 1 pour n tendant vers $+\infty$, **le** produit est divergent, et **l'étude** est terminée.

SI u_n tend vers 1 pour n tendant vers $+\infty$, alors, à partir d'une certaine valeur de n , u_n se trouve dans le demi-plan $\mathcal{R}\, u_n > 0$, et il est possible, après suppression d'un nombre **fini** de termes, d'utiliser les logarithmes. On voit alors Immédiatement :

Théorème 69 - <u>Pour que le produit infini $\prod_{n=0}^{\infty} u_n$, dont tous les termes vérifient $\mathcal{R}\, u_n > 0$, soit convergent, il faut et il suffit que la série $\sum_{n=0}^{\infty} \mathcal{L}og\, u_n$ soit convergente</u>.

Supposons d'abord le produit infini convergent. **Appelons** Π_n les produits partiels et S_n les sommes partielles de la série des logarithmes. On n'a pas nécessairement $\mathcal{L}og\, \Pi_n = S_n$, d'ailleurs on ne peut **même** pas affirmer que $\mathcal{R}\, \Pi_n$ 7 0 . Mals, **d'après** le critère de CAUCHY, il existe un entier p tel que $n \geqslant p$ **entraîne** $\left|\frac{\Pi_n}{\Pi_p} -1\right| <\cdot 1$, donc $\mathcal{R}\, \frac{\Pi_n}{\Pi_p} > 0$, et aussi $\mathcal{R}\, \frac{\text{l-r}}{\Pi_p} > 0$. Supposons un tel entier p choisi une fois pour toutes. Alors $\mathcal{L}og\, \frac{\Pi_n}{\Pi_p} = S_n - S_p + 2k_n i\pi$; **Comme** $\frac{\Pi_n}{\Pi_p}$ converge vers $\frac{\Pi}{\Pi_p}$, la **continuité** de la détermination **principale** du logarlthme montre que $S_n - S_p + 2k_n i\pi$ a une limite pour n Infini, donc la **série** de terme **général** $\mathcal{L}og\, u_n + 2(k_n - k_{n-1})ix$ est convergente; alors son terme **général** tend vers 0 , **mais, comme** u_n tend vers 1 **à** cause de la convergence du produit, $\mathcal{L}og\, u_n$ tend aussi vers 0 ; donc $k_{n-1} = k_n$ **à** partir d'une certaine valeur de n , mals alors la série de terme **général** $\mathcal{L}og\, u_n$ est aussi convergente.

Inversement, si la série est convergente, les S_n convergent vers une limite S, en vertu de la continuité de l'exponentielle, on en déduit bien que les $\Pi_n = e^{S_n}$ convergent vers $\Pi = e^{S} \neq 0$ .On voit pourquoi dans la démonstration, il était essentiel de supposer la valeur $\Pi \neq 0$. D'ailleurs, si les u_n sont réels >0 et si les Π_n tendent vers 0 ; $\mathcal{L}og\, u_n$ est le terme général d'une série divergente, de somme $-\infty$. On dit qu'un produit infini, déjà supposé convergent, est absolument convergent (resp semi-convergent) si la série des logarithmes (qui sont définis à partir d'un certain rang) est absolument convergente (resp semi-convergente) . Naturellement, il s agit là d une locution dont le sens pourrait être très ambigu; -dire que le produit est absolument convergent, ne signifie absolument pas que le produit $\Pi |u_n|$ soit convergent (celui-ci l'est toujours si le produit est convergent, et $\prod_{n=0}^{\infty} |u_n| = |\prod_{n=0}^{\infty} u_n|$).Si tous les u_n sont réels $\geqslant 1$, ou tous compris entre 0 et 1,convergence est synonyme de convergence absolue.

Théorème 70 - Pour que le produit Infini $\prod_{n=0}^{\infty}(1+v_n)$, $v_n \neq -1$ * soit absolument convergent, il faut et il suffit que la série $\sum_{n=0}^{\infty} |v_n|$ soit convergente.

Démonstration - Si le produit est absolument convergent, il est d'abord convergent, donc $1+v_n$ tend vers 1 et on peut prendre les $\mathcal{L}og$, pour $n \geq p$ assez grand. Ensuite la série

$\sum_{n \geq p} |\mathcal{L}og(1+v_n)|$ est supposée convergente; or pour la convergence des séries de nombres $\geqslant 0$ on a le droit de remplacer le terme général par un infiniment petit équivalent; pour n tendant vers $+\infty$, v_n tend vers 0 , alors $\mathcal{L}og\,(1+v_n) \sim v_n$, et par suite la série $\sum_{n \geq p} |v_n|$ est aussi convergente, et aussi la série $\sum_{n=0}^{+\infty} |v_n|$.

* Si l'on ne fait pas cette restriction $v_n \neq -1$, il pourra arriver que $\sum |v_n| < +\infty$,et que le produit infini ait un terme nul, donc ne soit pas convergent !

Réciproquement, supposons la série $\sum_{n=0}^{\infty} |v_n|$ convergente. Alors v_n tend vers 0, donc, pour $n \geqslant p$ assez grand, $\mathcal{R}(1+v_n) > 0$, et on peut prendre les logarithmes. Alors, pour n tendant vers $+\infty$, $|\mathcal{L}og\,(1+v_n)| \sim |v_n|$, donc la série $\sum_{n \geqslant p} |\mathcal{L}og\,(1+v_n)|$ est convergente; donc la série $\sum_{n \geqslant p} \mathcal{L}og\,(1+v_n)$ est aussi convergente, et par suite le produit $\prod_{n \geqslant p} (1+v_n)$ est convergent, d'après le théorème 69; et comme tous les t', sont $\neq -1$, le produit $\prod_{n=0}^{\infty} (1+v_n)$ n'a *aucun* terme nul et est aussi convergent; alors $\prod_{n=0}^{\infty} (1+v_n)$ est convergent et $\sum_{n \geqslant p} |\mathcal{L}og\,(1+v_n)|$ est convergente, donc le produit est absolument convergent.

<u>Exemple</u> Le produit infini :

(II,16;2) $$\prod_{n \geqslant 1} \left(1 + \frac{1}{n^{\alpha}}\right) \qquad \text{ou} \qquad \prod_{n \geqslant 2} \left(1 - \frac{1}{n^{\alpha}}\right)$$

est convergent si $\alpha > 1$, et divergent si $\alpha \leqslant 1$.

Il est en particulier divergent pour $\alpha = 1$ Remarquons que ce dernier cas se voit de façon Immédiate: car on connait les produits partiels :

(II,16;3) $$\left(1+\frac{1}{1}\right)\left(1+\frac{1}{2}\right)\left(1+\frac{1}{3}\right)\cdots\left(1+\frac{1}{n}\right) = n+1\,; \quad \left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\cdots\left(1-\frac{1}{n}\right) = \frac{1}{n}\,.$$

On peut même dire que la divergence du produit $\prod_{n=1}^{\infty}\left(1+\frac{1}{n}\right)$ ou $\prod_{n=2}^{\infty}\left(1-\frac{1}{n}\right)$ est plus immédiate que la divergence de la série $\sum_{n=1}^{\infty} \frac{1}{n}$, et peut servir à la démontrer en vertu du théorème 70 !

Produits infinis de fonctions réelles ou complexes

Soit $u_0, u_1, \cdots u_n, \cdots$ une suite de fonctions définies sur un ensemble E, à valeurs réelles ou complexes. Le produit infini $\prod_{n=0}^{\infty} u_n$ sera dit simplement convergent, si pour tout x de E, le produit infini $\prod_{n=0}^{\infty} u_n(x)$ de nombres réels ou complexes est convergent. Cela signifie que la suite de fonctions $\Pi_n = \prod_{0 \leqslant m \leqslant n} u_m$ est simplement convergente vers une fonction limite qui ne s'annule jamais.

Le produit est dit simplement absolument-convergent si, pour tout x de E , le produit de nombres $\prod_{n=0} u_n(x)$ est absolument convergent.

L'expression "convergence uniforme", pour un produit de fonctions, n'est pas claire. Dire que le produit de fonctions complexes $\prod_{n=0}^{\infty} u_n$ sur E converge uniformément vers la fonction Π peut signifier que les fonctions u_n et Π ne s'annulent jamais et que les Π_n convergent uniformement vers Π ; ou que les u_n et Π ne s'annulent jamais et que les $\frac{\Pi_n}{\Pi}$ convergent uniformément vers 1. Ces deux notions, comme on le voit facilement, ne coïncident pas nécessairement. Toutefois elles coïncident si la limite Π admet, sur E , une majoration et une minoration uniforme du type $0 < a \leq \Pi(x) \leq b < +\infty$. En effet, dans ce cas, l'inégalité $|\Pi_n - \Pi| \leq \varepsilon$ entraîne l'inégalité $\left|\frac{\Pi_n}{\Pi} - 1\right| \leq \frac{\varepsilon}{a}$, et l'inégalité $\left|\frac{\Pi_n}{\Pi} - 1\right| \leq \varepsilon$ entraîne l'inégalité $|\Pi_n - \Pi| \leq \varepsilon b$. C'est seulement dans ce cas que nous nous permettrons de parler de convergence uniforme d'un produit infini de fonctions. Mais on peut toujours parler de convergence uniforme <u>locale</u> d'un produit infini de fonctions continues sur un espace topologique E (et la limite Π est alors aussi continue). Si en effet les Π_n convergent localement uniformément vers Π , Π est continue d'après le théorème 65; comme elle est partout différente de 0, tout point a a un voisinage $\mathcal{V}'_a$ dans lequel $|\Pi|$ est bornée supérieurement et Inférieurement par des nombres > 0 fixes; alors, si $\mathcal{V}_a \subset \mathcal{V}'_a$ est un voisinage de a sur lequel les Π_n convergent uniformément vers Π, les $\frac{\Pi_n}{\Pi}$ convergent uniformément vers 1 sur $\mathcal{V}_a$.
Inversement, supposons que les $\frac{\Pi_n}{\Pi}$ convergent localement uniformément vers 1 . Pour tout a de E , il existe un voisinage $\mathcal{V}'_a$ et un n tel que $\left|\frac{\Pi_n(x)}{\Pi(x)} - 1\right| \leq \frac{1}{2}$ pour $x \in \mathcal{V}'_a$; cela entraîne $\frac{1}{2} \leq \left|\frac{\Pi_n(x)}{\Pi(x)}\right| \leq \frac{3}{2}$, donc $\frac{2}{3}\,|\Pi_n(x)| \leq |\Pi(x)| \leq 2\,|\Pi_n(x)|$. Comme Π_n est continue et que $\Pi_n(a) \neq 0$, il existe un voisinage $\mathcal{V}''_a \subset \mathcal{V}'_a$ de a sur lequel $|\Pi_n|$ est borné supérieurement et inférieurement par des constantes > 0, donc aussi $|\Pi|$; si alors $\mathcal{V}_a \subset \mathcal{V}''_a$ est un voisinage sur lequel les $\frac{\Pi_n}{\Pi}$ convergent uniformément vers 1 , les Π_n y convergent uniformément vers Π.

Application à la fonction ζ de Riemann

La fonction ζ de RIEMANN est définie par la formule :

(II,16 ;4) $$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Si nous posons $s = \sigma + i\tau$, et si $\delta > 0$, on voit que, considérée comme série de fonctions dans la région $\sigma \geq 1 + \delta$ du plan complexe, cette série est normalement convergente.

On a en effet :

(II.16;5) $$\left|\frac{1}{n^{\sigma+i\tau}}\right| = \frac{1}{n^{\sigma}} \leq \frac{1}{n^{1+\delta}} .$$

Comme, pour tout n , la fonction $s \longrightarrow \frac{1}{n^{s}}$ est continue dans le demi-plan $\sigma \geqslant 1 + \delta$, on voit-que la somme, c'est-à-dire la fonction ζ , est continue dans ce même demi-plan; et comme ceci est vrai pour tout $\delta > 0$, la fonction ζ est continue dans tout le demi-plan $\sigma > 1$.

Considérons maintenant le produit Infini, où p parcourt l'ensemble de tous les nombres premiers

(II.16;6) $$G(s) = \prod_{p} \left(\frac{1}{1 - \frac{1}{p^{s}}} \right) .$$

Un terme quelconque de ce produit est toujours $\neq 0$. Par ailleurs, le dénominateur $1 - \frac{1}{p^{s}}$ est toujours $\neq 0$, pour $\sigma > 0$. En outre dans ce cas le module de $\frac{1}{p^{s}}$ est $\frac{1}{p^{\sigma}} < 1$ pour $\sigma > 0$, et par conséquent le théorème 70 est applicable.

Le produit infini est absolument convergent si et seulement si son inverse l'est, c'est-à-dire si la série $\sum_{p} \frac{1}{p^{\sigma}}$ est convergente. Ce la se produit donc certainement pour $\sigma > 1$, puisque cette série a une somme majorée par celle de la série $\sum_{n=1}^{\infty} \frac{1}{n^{\sigma}}$, qui est elle-même convergente.

Nous nous bornerons donc à considérer toujours $\sigma > 1$.

Théorème 71 - Pour $\sigma > 1$, on a l'égalité $G(s) = \zeta(s)$.

Démonstration - Pour démontrer ce théorème, il est naturellement possible de supposer s fixé une fois pour toutes. Alors, si $\varepsilon > 0$ est donné, on peut, puisque la série $\sum_{n=1} \frac{1}{n^{\sigma}}$ et le produit G sont convergents, trouver un entier m , ayant les propriétés suivantes :

a) le reste $\sum_{n>m} \frac{1}{n^{\sigma}}$ est majoré par $\frac{\varepsilon}{2}$;

b) si nous appelons $G_m(s)$ le produit partiel formé des m premiers facteurs du produit infini, on a $|G_m(s) - G(s)| \leq \frac{\varepsilon}{2}$.

On a alors pour tout nombre premier p , le développement en série géométrique absolument convergente :

$$\frac{1}{1-\frac{1}{p^s}} = 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \cdots = \sum_{k=0}^{\infty} \frac{1}{p^{ks}} . \tag{II,16;7}$$

En vertu de la règle relative au produit de plusieurs séries absolument convergentes (théorème 61), on peut écrire :

$$G_m(s) = \sum_{k_1,k_2,\ldots k_m} \left(\frac{1}{p_1}\right)^{k_1 s} \cdots \left(\frac{1}{p_m}\right)^{k_m s} , \tag{II,16;8}$$

où $p_1 = 2$, $p_2 = 3$, $p_3 = 5, \ldots p_m$ sont les m premiers nombres premiers, et $k_1 , k_2 , \ldots k_m$ des entiers ≥ 0 ; ce qui prouve que l'on a :

$$G_m(s) = \sum_{\nu} \frac{1}{\nu^s} , \tag{II,16;9}$$

oùl'entier ν parcourt la suite de tous les entiers qui, dans leur décomposition en facteurs premiers, ne contiennent que les m premiers nombres premiers p_1 , $p_2 , \cdots p_m$. Si alors nous considérons la différence $G_m(s) - \zeta(s)$, elle se compose d'une partie des termes de la série $\sum_n \frac{1}{n^s}$, qui tous correspondent à des indices $n > m$; on a donc les inégalités :

$$\left| G_m(s) - \zeta(s) \right| \leq \sum_{n>m} \frac{1}{n^\sigma} \leq \frac{\varepsilon}{2} , \tag{II,16;10}$$

d'où l'on déduit l'inégalité :

$$\left| G(s) - \zeta(s) \right| \leq \left| G(s) - G_m(s)\right| + \left| G_m(s) - \zeta(s)\right| \leq \varepsilon . \tag{II,16;11}$$

Comme alors ε est arbitraire, on en déduit bien que l'on a $G(s) = \zeta(s)$.

Corollaire - La fonction ζ ne s'annule jamais, pour $\sigma > 1$.

Elle est en effet égale à la valeur d'un produit infini convergent.

Les résultats précédents ne sont évidemment plus valables pour $\sigma = 1$. En particulier, on sait que la série $\sum_{n=1} \frac{1}{n}$ est divergente. Nous allons de même démontrer :

Théorème 72 - Le produit infini $\prod_p \left(\frac{1}{1-\frac{1}{p}}\right)$ est divergent.

Pour cela remarquons que, si $A > 0$ est un nombre quelconque, on peut trouver un entier m , tel que :

$$1 + \frac{1}{2} + \cdots + \frac{1}{m} \geq A . \tag{II,16;12}$$

Si alors nous considérons le produit partiel G_m, les développements en série géométrique utilisés précédemment sont encore valables, et par conséquent G_m est une somme $\sum \frac{1}{\nu}$, dans laquelle ν parcourt tous les entiers dont les seuls facteurs premiers sont $p_1, p_2, \ldots p_m$. Il en résulte en particulier que l'on a l'inégalité :

$$G_m(1) = \sum \frac{1}{\nu} \geqslant 1 + \frac{1}{2} + \cdots + \frac{1}{} \geqslant A . \tag{II,16;13}$$

Comme A est arbitraire, cela prouve bien que le produit infini (dont tous les facteurs sont.> 1) : est divergent : $G(1) = +\infty$.

Corollaire - Il existe une infinité de nombres premiers, et même la série $\sum_p \frac{1}{p}$ est divergente.

En effet la divergence de cette série est exactement equivalente à celle du produit infini $\prod_p \left(1-\frac{1}{p}\right) = \frac{1}{G(1)}$.

Remarque : Considérons maintenant la série alternée

$$\zeta_a(s) = \frac{1}{1^s} - \frac{1}{2^s} + \frac{1}{3^s} + \cdots + (-1)^{n-1} \frac{1}{n^s} + \cdots \tag{II,16;14}$$

Cette série est convergente, comme nous allons le voir, pour $\sigma > 0$. Nous allons même montrer qu'elle est uniformément convergente sur tout compact du demi-plan ouvert $\sigma > 0$ du plan complexe. Soit donc K un tel compact. Nous remarquerons d'abord que, sur K, $|s|$ est borne supérieurement par un nombre S, puisque c'est une fonction continue. Ce la même manière, σ est bornée inférieurement par un nombre $\delta > 0$, puisque c'est une fonction continue partout > 0 sur K.

Appliquons alors le théorème d'ABEL (théorème 63). On a $\frac{(-1)^{n-1}}{n^s} = u_n v_n$, où $v_n = (-1)^{n-1}$, $u_n = \frac{1}{n^s}$. Les $|\sigma_{m,n}|$ sont majorés par 1. Montrons que la suite des u_n est à variation bornée. On a

$$\frac{1}{n^s} - \frac{1}{(n+1)^s} = s \int_n^{n+1} \frac{dx}{x^{s+1}} \,{}^{*} \tag{II,16;15}$$

d'où

* C'est un procédé général pour majorer une différence; on écrit : $f(n+1) - f(n) = \int_n^{n+1} f'(t)\, dt$

$$(\text{II},16;16) \quad \left|\frac{1}{n^s} - \frac{1}{(n+1)^s}\right| \leqslant |s| \int_n^{n+1} \frac{dx}{x^{\sigma+1}} \qquad \text{et}$$

$$(\text{II},16;17) \quad \left|\frac{1}{n^s} - \frac{1}{(n+1)^s}\right| + \left|\frac{1}{(n+1)^s} - \frac{1}{(n+2)^s}\right| + \cdots$$

$$\leqslant |s| \int_n^{+\infty} \frac{dx}{x^{\sigma+1}} = \frac{|s|}{\sigma n^\sigma} .$$

La série est donc bien convergente, et la formule (II,14;31) donne, pour le reste, la majoration :

$$(\text{II},16;18) \quad \left| \sum_{n \geqslant m+1} \frac{(-1)^{n-1}}{n^s} \right| \leqslant \frac{|s|}{\sigma (m+1)^\sigma} .$$

La convergence est uniforme sur K , puisqu'alors

$|R_m| \leqslant \frac{S}{\delta (m+1)^\delta}$, qui ne dépend pas de s et tend vers 0

pour m tendant vers $+\infty$.

Si on suppose $\sigma > 1$, il existe une relation simple entre les fonctions ζ et ζ_a . On a en effet la formule :

$\zeta_a(s) = \zeta(s) - 2 \sum_{n=1}^{\infty} \left(\frac{1}{2n}\right)^s$, d'où l'on déduit:

$$(\text{II},16;19) \qquad \zeta_a(s) = \zeta(s) \left(1 - \frac{1}{2^{s-1}}\right) \quad \text{ou} \quad \zeta(s) = \frac{\zeta_a(s)}{1 - \frac{1}{2^{s-1}}} .$$

La propriété de convergence uniforme démontrée pour ζ_a montre que cette fonction est continue sur tout compact K du demi-plan ouvert $\sigma > 0$. Elle est donc partout continue dans ce demi-plan; en particulier, lorsque s tend vers **1**, $\zeta_a(s)$ tend vers $\zeta_a(1) = \mathcal{L}og\, 2$. Alors la formule (II,16;19) montre que, lorsque s tend vers **1**, $\zeta(s)$ est équivalente à $\frac{\mathcal{L}og\, 2}{1 - e^{(1-s)\mathcal{L}og\, 2}} \sim \frac{\mathcal{L}og\, 2}{(s-1)\mathcal{L}og\, 2}$. En fait la formule (II,16;19) permet de prolonger la fonction ζ dans le demi-plan $\sigma > 0$ * .

D'autres méthodes de prolongement permettent plus généralement de définir la fonction ζ dans tout le plan complexe, et de montrer que c'est une fonction holomorphe de la variable complexe s , c'est-à-dire une fonction continue et à **dérivée** première **continue** par rapport à cette variable complexe, dans le complémentaire du point $s = 1$ du plan complexe. Ce point

* Naturellement ce prolongement n'a plus rien à voir avec la somme de la série $\sum_{n=1} \frac{1}{n^s}$, qui n'a pas de sens pour $\sigma \leqslant 1$.

$s = 1$ est un pôle, $\zeta(1) = \infty$. Cette fonction prolongée s'annule aux points $s = -2, -4, -6, \dots$. L'dtude de cette fonction donne des renseignements sur la répartition des nombres premiers, comme nous avons déjà pu le voir par un exemple simple; RIEMANN a émis l'hypothèse, qui jusqu'à présent n'a encore jamais été démontrée, que la fonction ζ prolongée a tous ses zéros en dehors des précédents, sur la demi-droite verticale $\sigma = \frac{1}{2}$ La démonstration de cette hypothèse donnerait des renseignements extrêmement précis sur la répartition de la suite des nombres premiers.

De toute façon, les propriétés déjà actuellement *connues* de la fonction ζ permettent de montrer que le $n^{ième}$ nombre premier est équivalent, pour n tendant vers l'infini, a $n \operatorname{Log} n$, ou encore que le nombre des nombres premiers compris entre 1 et N est dquivalent, pour N tendant vers $+\infty$, à $\dfrac{N}{\operatorname{Log} N}$.

La théorie des nombres premiers est une des théorie les plus intéressantes, et les plus difficiles des mathématiques.

III

CALCULDIFFÉRENTIEL

§ 1 ESPACES AFFINES

Jusqu'à la classe de mathématiques spéciales on a considéré comme plus ou moins intuitive, la notion d'espace de la géométrie élémentaire, dont les éldments étaient appelds points, On peut alors introduire la notion de vecteurs libres à l'aide d'une relation d'équivalence comme il est indiqué page 14. L'espace des vecteurs libres est un espace vectoriel possédant les propriétés habituelles.

Il n'est plus possible, à partir àe maintenant, d'utiliser ces notions insuffisantes apprises antérieurement, car l'espace de la géométrie élémentaire n'a jamais été défini d'une façon rigoureuse On doit considérer que les mathématiques sont fondées de la façon suivante :

Après la théorie générale des ensembles, on définit. correctement l'ensemble $\mathbb{N}$ des nombres entiers $\geqslant 0$, puis l'ensemble $\mathbb{Z}$ des nombres entiers de signes quelconques. On introduit ensuite le corps $\mathbb{Q}$ des nombres ration--nels, à partir d'une relation d'équivalence, comme il est dit page 14. En mathématiques spéciales, à partir de la notion de coupure, on a défini en toute rigueur le corps $\mathbb{R}$ des nombres réels, à partir du corps $\mathbb{Q}$ des nombres rationnels; et enfin le corps $\mathbb{C}$ des nombres complexes. On peut alors introduire la notion générale abstraite d'espace vectoriel sur un corps (qui sera, en général, le corps des réels ou le corps des complexes).

La notion d'espace vectoriel et les propriétés de ces espaces ont été étudiées antérieurement. C'est à partir des espaces vectoriels que nous introduirons ici en toute rigueur l'espace de la géométrie élémentaire, qui est un espace euclidien aff'ine. Ii est évidemment très voisin d'un espace vectoriel, mais ne possède pas d'origine privilégiée. Dans la suite nous noterons toujours les éléments d'un espace vectoriel par une lettre surmontée d'une flèche, $\vec{X}$, et nous appellerons ces éléments des vecteurs.

Définition - On appelle espace affine E , sur le corps $\mathbb{K}$ desreels ou des complexes, un ensemble non vide, dont les éléments sont appelés points, auquel sont associés, d'une part un autre ensemble, noté $\vec{E}$ muni d'une structure d'espace vectoriel sur le corps des réels ou des complexes, appelé espace vectoriel associé à E , et dont les éléments sont appelés vecteurs, et, d'autre part, une application de $E \times E$ dans $\vec{E}$, ayant des propriétés que nous détaillerons plus loin.

Si a et b sont des points de E , l'élément associé au couple (a,b) par l'application précédente est donc un vecteur de $\vec{E}$, qu'on note $\overrightarrow{ab}$, et qu'on appelle le vecteur d'origine a et d'extrémité b. Les propriétés que doit posséder l'application de $E\times E$ dans $\vec{E}$, sont les suivantes :

1°) La relation de Chasles : Quels que soient a., a , c , dans E :

$$\overrightarrow{ab} + \overrightarrow{bc} + \overrightarrow{ca} = \vec{0}, \text{ vecteur nul de } \vec{E} \qquad (\text{III},1;1)$$

De oette relation on déduit en particulier (en prenant les trois points a, b, c, confondus en a,) que, quel que soit a , le vecteur $\overrightarrow{aa}$ est le vecteur $\vec{0}$, origine de l'espace vectoriel. On voit aussi, en prenant simplement a d'une part, et d'autre part b = c , que $\overrightarrow{ab}$ et $\overrightarrow{ba}$ sont deux vecteurs opposés.

2°) Quel que soit le point a fixé, l'application $x \longrightarrow \overrightarrow{ax}$ doit être une bijection de E sur $\vec{E}$. Nous adopterons généralement les notations suivantes :

Tout d'abord, au lieu de $\overrightarrow{ab}$ on peut noter $\overrightarrow{b-a}$, le vecteur d'origine a et d'extrémité b ; la relation de CHASLES nous montre en effet que l'on a $\overrightarrow{b-a} + \overrightarrow{c-b} + \overrightarrow{a-c} = 0$, cette nouvelle notation est donc compatible avec les propriétés usuelles de la soustraction. D'autre part, si a est un point de E et $\vec{h}$ un vecteur de $\vec{E}$, la propriété de bijection 2°) nous affirme qu'il existe un point 8 de E , et un seul, tel que $\overrightarrow{b-a} = \vec{h}$; il est commode de noter ce point par $a + \vec{h}$; on a alors, en vertu de la relation de CHASLES, la relatinn suivante : $a + (\vec{h} + \vec{k}) = (a + \vec{h}) + \vec{k}$, donc cette nouvelle notation est compatible avec les propriétés habituelles de l'addition. La dimension de l'espace vectoriel associé $\vec{E}$ s'appelle aussi dimension de l'espace affine E. On est amené à considérer que l'ensemble vide est aussi un espace affine, mais sans espace vec-toriel associé.

Naturellement un espace vectoriel est un espace affine particulier; il suffit de considérer ici l'espace affine et l'espace vectoriel comme confondus, et de faire correspondre à deux éléments quelconques $\vec{a}$, $\vec{b}$, de l'espace vectoriel, levecteur $\overrightarrow{ab} = \vec{b} - \vec{a}$. En particulier le corps des scalaires $\mathbb{K}$ lui-même est un espace affine de dimension 1.

On appelle référentiel ou système de référence d'un espace affine de dimension finie, le système formé d'une origine 0 de E et d'une base $(\vec{e}_i)_{i\in I}$ * de l'espace vectoriel associé $\vec{E}$. Si alors nous considérons un point quelconque x de E, le vecteur $\overrightarrow{x-0}$ possède des coordonnées $(x_i)_{i\in I}$ par rapport à la base choisie dans $\vec{E}$; ces coordonnées sont appelées les coordonnées de x par rapport au système de référence considéré, et on a la formule

(III, 1; 2) $$x = 0 + \sum_{i\in I} x_i \vec{e}_i .$$

Variétés affines Soit F un sous-ensemble non vide d'un espace affine E. Supposons qu'il ait la propriété suivante :

Il existe un sous-espace vectoriel $\vec{F}$ de $\vec{E}$, tel que, pour tout couple (a,b) de $F \times F$, le vecteur $\overrightarrow{b-a}$ appartienne à $\vec{F}$, et, pour tout couple $(a,\vec{h})$ de $F \times \vec{F}$, le point $a+\vec{h}$ appartienne à F. Si un tel sous-espace vectoriel $\vec{F}$ existe, il est évidemment unique, puisqu'il est exactement l'ensemble de tous les vecteurs $\overrightarrow{b-a}$ pour tous les couples (a, b) de $F \times F$. Dans ces conditions,

on dit que F est un sous-espace affine ou une variété affine, souvent même une variété linéaire, de E, et que $\vec{F}$ est son sous-espace vectoriel associé.

* Si $\vec{E}$ est de dimension n, I est un ensemble quelconque d'"indices" à n éléments, et $(\vec{e}_i)_{i\in I}$ est une "famille" de n vecteurs de $\vec{E}$. Assez souvent $I = \{1, 2, \dots n\}$, ensemble des n premiers entiers > 0, et la base est une "suite" de n vecteurs de $\vec{E}$.

F possède en effet une structure d'espace affine avec $\vec{F}$ comme espace vectoriel associé, et, comme application de $F \times F$ dans $\vec{F}$, la restriction de l'application donnée de $E \times E$ dans $\vec{E}$. On convient aussi que la partie vide de E est une variété affine de E, sans espace vectoriel associé. E lui-même est une variété affine; un point est une variété affine de dimension 0. On appelle droite une variété affine de dimension 1, on appelle plan une variété affine de dimension 2. Un sous-espace vectoriel $\vec{F}$ de $\vec{E}$ est dit un hyperplan si ses sous-espaces vectoriels supplémentaires ont la dimension 1; une variété affine F de E est appelée un hyperplan si son sous-espace vectoriel associé est un hyperplan. Si E est de dimension finie n, un hyperplan est simplement une variété affine de dimension $n-1$.

Deux variétés affines de E de même dimension sont dites parallèles si elles ont le même sous-espace vectoriel associé.* En particulier deux variétés confondues sont parallèles. On voit que, dans cette manière d'introduire la théorie des espaces affines, le 'postulat' d'Euclide est un théorème, d'ailleurs évident : pour tout point de l'espace on peut mener une variété-affine parallèle à une variété donnée, et une seule. Cela revient à dire que, si $\vec{F}$ est un sous-espace vectoriel de $\vec{E}$ et a un point de E, il existe une variété affine et une seule, contenant a et d'espace vectoriel associé $\vec{F}$: c'est l'ensemble des $a + \vec{h}$, $\vec{h} \in \vec{F}$.

L'intersection d'une famille finie ou infinie de variétés affines d'un espace affine, est une variété affine : il en résulte que, si A est une partie quelconque d'un espace affine, il existe une variété affine plus petite que toutes les autres, qui contienne A à savoir l'intersection de toutes les variétés affines contenant A. On l'appelle la variété affine engendrée par A. Si deux variétés affines ont les dimensions p et q, et si on appelle i la dimension de leur intersection, et s la dimension de la variété affine engendrée par leur réunion, on démontre aisément que, si l'intersection n'est pas vide, on a la formule : $p + q = i + s$.

Le produit $E_1 \times E_2$ de deux espaces affines a évidemment une structure d'espace affine, d'espace vectoriel associé $\vec{E}_1 \times \vec{E}_2$, en posant : $\overrightarrow{(b_1, b_2) - (a_1, a_2)} = (\overrightarrow{b_1 - a_1}, \overrightarrow{b_2 - a_2})$.

Applications linéaires, applications affines

Soit E et F deux espaces affines. On dit qu'une application u de E dans F est une application affine, s'il

* Alors, si E est un plan (espace affine à 2 dimensions), 2 droites distinctes sont parallèles, si et seulement si elles ne se coupent pas.

existe une application linéaire, notée $\vec{u}$, de $\vec{E}$ dans $\vec{F}$, telle que

$$(\text{III},1;3) \qquad \overrightarrow{u(b)-u(a)} = \vec{u}(\overrightarrow{b-a}) .$$

Dans ce cas, l'application $\vec{u}$ est manifestement unique, puisque sa valeur sur n'importe quel vecteur de $\vec{E}$ est connue * . Si E et F sont de dimension finie, et si l'on choisit dans chacun des deux espaces affines un système de référence, à savoir une origine a de E et une base $(\vec{e}_j)_{j\in J}$ de $\vec{E}$, puis une origine b de F et une base $(\vec{f}_i)_{i\in I}$ de $\vec{F}$, une application affine u est entièrement connue, si l'on connait les coordonnées $(c_i)_{i\in I}$ du point $u(a)$ dans l'espace affine F, et les coordonnées $u_{i,j}$, $i \in I$, de chaque vecteur $\vec{u}(\vec{e}_j)$, $j\in J$, dans l'espace vectoriel $\vec{F}$. L'application u est alors définie par la formule

$$(\text{III},1;4) \qquad \begin{cases} u(a) = b + \sum_{i\in I} c_i \vec{f}_i \\ \vec{u}(\vec{e}_j) = \sum_{i\in I} u_{i,j} \vec{f}_i \end{cases}$$

On peut encore dire qu'elle fait correspondre, au point

$x = a + \sum_{j\in J} x_j \vec{e}_j$ de E, de coordonnées $(x_j)_{j\in J}$
le point $y = b + \sum_{i\in I} y_i \vec{f}_i$ de F, de coordonnées $(y_i)_{i\in I}$
suivant la formule :

$$(\text{III},1;5) \qquad y_i = c_i + \sum_{j\in J} u_{i,j}\, x_j \ ; \quad i \in I .$$

Si en particulier F est le corps des réels, ou le corps des complexes, muni de son système de référence canonique constitué par son origine et le vecteur unité, ont voit qu'on pourra parler d'une fonction réelle ou complexe affine . C'est une fonction qui, avec le système de référence choisi dans E, fera correspondre au point x, de coordonnées $(x_j)_{j\in J}$, le nombre réel ou complexe

$$(\text{III},1;6) \qquad u(x) = c + \sum_{j\in J} u_j\, x_j \ , \quad c = u(a),\ u_j = \vec{u}(\vec{e}_j) .$$

* Très fréquemment, on ne mettra pas de flèche sur l'application linéaire associée, et on notera u à la fois l'application affine et son application linéaire associée; la notation $\vec{u}$ a en effet le défaut de faire croire que $\vec{u}$ est un vecteur de $\vec{E}$ ou $\vec{F}$.

On voit que ce qui est appelé souvent une fonction linéaire, doit désormais s'appeler une fonction affine, et que ce qui est appelé souvent une fonction linéaire homogène ($\vec{E}$ espace vectoriel, a = $\vec{0}$, c = $\vec{0}$) doit s'appeler désormais une fonction linéaire. $y = ax + b$ est une fonction affine, $y = ax$ sa fonction linéaire associée.

Si $\vec{h}$ est un vecteur de l'espace vectoriel associé à un espace affine E , la bijection $x \to x + \vec{h}$ de E dans lui-même est appelée la translation de vecteur $\vec{h}$. C'est manifestement une application affine, dont l'application linéaire associée est l'application identiaue. Récioroauement, toute application affine dont l'application linéaire associée est l'application identiaue est une translation. Car $\overrightarrow{u(b)-u(a)} = \overrightarrow{b-a}$, donc $\overrightarrow{u(b)-b} = \overrightarrow{u(a)-a}$, donc $\overrightarrow{u(x)-x}$ est indépendant de x ; si $\vec{h}$ est sa valeur on a $u(x) = x + \vec{h}$

Espaces affines normés

On dit qu'un espace affine est normé, si l'espace vectoriel associé est normé. On notera donc que la norme est une fonction définie sur l'espace vectoriel associé et non sur l'espace affine lui-même. On peut parler de la norme d'un vecteur, mais non de la norme d'un point.

Un espace affine normé possède une métrique définie par la fonction distance : $d(x,y) = \| \overrightarrow{x-y} \|$.

Cette fonction distance est invariante par les translations, en ce sens que $d(x+\vec{h}, y+\vec{h}) = d(x,y)$; d'autre part, dans une homothétie de centre 0 quelconque et de rapport λ (définie par l'application $x \to x' = 0 + \lambda(\overrightarrow{x-0})$ de E dans lui-même), elle est multipliéepar $|\lambda|$ (en ce sens que $d(x',y') = |\lambda|\, d(x,y)$), . Si E est un espace affine normé, l'application $(x,y) \to \overrightarrow{y-x}$ de $E \times E$ dans $\vec{E}$, et l'application $(x, \vec{h}) \to x + \vec{h}$ de E x $\vec{E}$ dans E sont continues.

Si E et F sont des espaces affines normés, E de dimension finie, toute application affine de E dans F est nécessairement continue. Il suffit en effet de répéter le raisonnement de la page 104. Par contre, si E et F sont de dimension infinie, il n'en est plus du tout ainsi On a d'ailleurs vu page 105 qu'une application linéaire d'un espace vectoriel normé dans un autre pouvait n'être pas continue.

Théorème 1 - 1°) Un espace affine normé est complet si et seulement si l'espace vectoriel associé est complet.

2°) Un sous-espace affine F de E est fermé, si et seulement si son sous-espace vectoriel associé $\vec{F}$ est fermé dans $\vec{E}$ (en particulier, un sous-espace affine de dimension finie est toujours fermé).

3°) <u>Pour qu'une application affine u d'un espace affine normé E dans un espace affine normé F soit continue, il faut et il suffit que l'application linéaire associée soit continue * ; dans ce cas, u est uniformément continue.</u>

<u>Démonstration</u> - 1°/ Soit a une origine choisie dans E.

L'application $\vec{X} \to a + \vec{X}$ est une bijection de $\vec{E}$ sur E, conservant les distances, c'est-à-dire conservant la structure d'espace métrique; E est donc complet en même temps que $\vec{E}$.

2°/ Soit $a \in F$. Alors $\vec{X} \to a + \vec{X}$ est un homéomorphisme de $\vec{E}$ sur E, et l'image de $\vec{F}$ est F, donc F est fermé dans E si et seulement si $\vec{F}$ est fermé dans $\vec{E}$.

3°/ Si u est une application affine continue de E dans F, l'application linéaire associée $\vec{u}$ est définie par

4 (III,1;7) $$\vec{u}(\vec{X}) = \overrightarrow{u(a+\vec{X}) - u(a)},$$

a quelconque fixé; elle est par conséquent manifestement continue. Réciproquement, si $\vec{u}$ est continue, l'inégalité

5 (III,1;7 bis) $$\|\overrightarrow{u(x) - u(y)}\| = \|\vec{u}(\vec{x}-\vec{y})\| \leq \|\vec{u}\| \, \|\vec{x}-\vec{y}\|$$

montre que u est lipschitzienne et par suite uniformément continue.

<u>Théorème 2</u> - <u>Soit E un espace affine normé, f une fonction scalaire affine sur E, non constante. Alors l'équation $f(x) = 0$ définit un hyperplan affine H.</u>

<u>Inversement tout hyperplan affine a une infinité d'équations de cette forme; toutes les fonctions affines correspondantes sont proportionnelles à l'une d'entre elles.</u>

<u>Tout hyperplan affine H est fermé ou dense; il est fermé si et seulement si les fonctions affines f définissant son équation sont continues.</u>

<u>Démonstration</u> - 1°/ Soit f une fonction scalaire affine non constante, f_0 la forme linéaire sur $\vec{E}$ associée ** . L'espace E contient alors au moins deux points a et b, tels que

* Rappelons qu'une application linéaire est continue, si et seulement si elle est continue à l'origine (théorème 47 chapitre II).

** Voir * page 171; si l'on mettait $\vec{f}$, on écrirait que $f(\vec{X})$ est un vecteur, alors que c'est un scalaire.

$f(a) \neq f(b)$. La formule

$$f(a+t(\overrightarrow{b-a})) = f(a) + t f_0(\overrightarrow{b-a}) , \qquad f_0(\overrightarrow{b-a}) \neq 0 ,$$

montre qu'on peut choisir t pour que les deux quantités égales de cette formule soient nulles; autrement dit f s'annule en au moins un point c de E . En choisissant c comme origine, on établit une bijection $\vec{X} \to c + \vec{X}$ de $\vec{E}$ sur E , ce qui va nous permettre de raisonner sur $\vec{E}$ à la place de E , soit $\vec{H}$ l'ensemble défini par l'équation $f_0(\vec{X}) = 0$, ou encore l'ensemble&' $(\{\vec{0}\})$ ou encore le noyau de la forme linéaire f_0 ; soit d'autre part H l'ensemble de E défini par l'équation $f(x) = 0$. Puisque $f(c) = 0$, on a $x \in H$ si et seulement si $\overrightarrow{x - c} \in \vec{H}$; autrement dit on a $H = c + \vec{H}$.

Si nous démontrons que $\vec{H}$ est un sous-espace vectoriel hyperplan, on aura bien démontré que H est un sous-espace affine hyperplan. Mais, puisque f n'est pas constante, f_0 n'est pas identiquement nulle; il existe donc un élément $\vec{e}$ de $\vec{E}$ tel que $f_0(\vec{e}) \neq 0$, et, en remplaçant au besoin $\vec{e}$ par un multiple de $\vec{e}$, on peut toujours supposer que $f_0(\vec{e}) = 1$. Tout élément $\vec{X}$ de $\vec{E}$ s'écrit alors, d'une manière et d'une seule, sous la forme

(III,1;7 ter) $$\vec{X} = \vec{Y} + \lambda \vec{e} , \qquad \lambda \in \mathbb{K} , \qquad \vec{Y} \in \vec{H} \qquad (\text{ou } f_0(\vec{Y}) = 0).$$

Cette formule équivaut en effet à

$$\lambda = f_0(\vec{X}) , \qquad \vec{Y} = \vec{X} - f_0(\vec{X}) \vec{e}$$

Cela prouve exactement que le sous-espace vectoriel; et le sous-espace vectoriel à une dimension engendré par $\vec{e}$ sont supplémentaire dans $\vec{E}$, c'est-a-dire que $\vec{H}$ est un hyperplan.

2°/ Réciproquement, soit H un hyperplan de E . Son sous-espace vectoriel associé $\vec{H}$ est alors par définition un hyperplan de $\vec{E}$.

Soit alors $\vec{e}$ un vecteur supplémentaire de $\vec{H}$ dans $\vec{E}$ ($\vec{e} \notin \vec{H}$). Tout vecteur $\vec{X}$ de $\vec{E}$ admet alors une décomposition unique de la forme (III,1;7 ter). Le scalaire λ dépend de $\vec{X}$, et nous pourrons l'appeler $\lambda = f_0(\vec{X})$. La fonction $f_0 : \vec{X} \to f_0(\vec{X})$ est une forme linéaire sur $\vec{E}$; elle n'est pas identiquement nulle et l'ensemble de ses zéros est exactement $\vec{H}$.

Toute forme linéaire g_0 telle que l'équation $g_0(\vec{X}) = 0$ définisse $\vec{H}$ est proportionnelle à f . Si en effet $K = g_0(\vec{e})$, on a $g_0(\vec{X}) = K f_0(\vec{X})$ pour $\vec{X} = \vec{e}$ par définition, et pour $\vec{X} \in \vec{H}$, puisqu'alors les deux membres sont nuls, donc pour $\vec{X} \in \vec{E}$ quelconque. Si maintenant C est un point quelconque de H , l'ensemble H est défini par l'une quelconque des équations $g(x) = 0$, où $g(x) = g_0(\overrightarrow{x - c})$: toutes les fonctions affines correspondantes g sont proportionnelles à l'une d'entre elles.

3°/ Si f est continue, l'ensemble H, image réciproque par f l'ensemble fermé $\{0\}$ du corps des scalaires, est ferme dans E. Supposons au contraire f, donc f_0 discontinue. D'après le théorème 47 du Chapitre II, pour tout entier n on peut trouver un vecteur $\vec{a}_n$ de $\vec{E}$ tel que $|f_0(\vec{a}_n)| \geq n \|\vec{a}_n\|$; en multipliant $\vec{a}_n$ par un facteur scalaire, on peut supposer que $\|\vec{a}_n\| \leq \frac{1}{n}$, et que $f_0(\vec{a}_n) = 1$. Soit alors x un point quelconque de E . Considérons la suite des points $x_n = x - f(x)\vec{a}_n$.
On aura $f(x_n) = 0$, autrement dit ces points x_n appartiendront à H.

Mais les $f(x)\vec{a}_n$ convergent vers $\vec{0}$ et par suite les x_n convergent vers x pour n tendant vers l'infini, on voit donc bien que tout point x de E est adhérent à H, et que H est bien dense, ce qui achève la démonstration du théorème.

Ce théorème met en évidence un fait quelque peu surprenant, auquel on n'est pas habitué par la considération des espaces de dimension finie : il peut arriver qu'un hyperplan soit dense. Nous aurons d'ailleurs des occasions, ultérieurement, de trouver, dans les espaces vectoriels normés, des sous-espaces vectoriels denses (un célèbre théorème de Weierstrass dit que, dans l'espace $\mathcal{C}([0,1])$ des fonctions complexes continues sur $[0,1]$, le sous-espace des polynômes est dense).

Ensembles convexes dans les espaces affines

Soit a et b deux points d'un espace affine E sur le corps $\mathbb{K}$ des réels ou des complexes. On appelle segment d'extrémités a et b, l'ensemble des points qui peuvent s'écrire sous la forme $a + t\,\overrightarrow{b-a}$, t réel, $0 \leq t \leq 1$. On note par $[a, b]$ ce segment, qu'on appelle aussi un segment fermé. On note par $[a,b[$ (resp. $]a, b]$, resp. $]a,b[$) le même segment privé du point b (resp. du point a, resp. des deux points a, b); on emploie ici les expressions "segment semi-ouvert" ou "segment ouvert" bien qu'ils ne soient pas ouverts si l'espace affine est normé .

On appelle partie convexe de E toute partie qui, toutes les fois qu'elle contient deux points, contient tout entier le segment qui les admet comme extrémités.

Dans un espace affine, on définit sans peine la notion de barycentre. Une partie A de E est alors convexe, si et seulement si, toutes les fois qu'elle contient un nombre fini de points, elle contient aussi leur barycentre pour n'importe quel système de masses ≥ 0, on le voit aisément, en remarquant qu'un barycentre de plusieurs points se construit en effectuant plusieurs fois la construction d'un barycentre de 2 points, et qu'un barycentre de 2 points, pour des masses ≥ 0, est sur le segment qui les joint.

La partie vide, une partie réduite à un point, E lui-même, et plus généralement toute variété affine de E est un ensemble convexe. Si E est normé, toute boule ouverte ou fermée est convexe. Toute intersection d'une famille finie ou infinie de parties convexes est une partie convexe. On peut montrer que l'intérieur et l'adhérence d'une partie convexe dans un espace affine normé sont encore convexes.

Espaces vectoriels et affines euclidiens

Soit $\vec{E}$ un espace vectoriel sur le corps des réels $\mathbb{R}$. On appelle produit scalaire euclidien sur $\vec{E}$ une forme bilinéaire $(\vec{X}, \vec{Y}) \longrightarrow B(\vec{X}, \vec{Y}) \in \mathbb{R}$, symétrique, c'est-à-dire telle que $B(\vec{X}, \vec{Y}) = B(\vec{Y}, \vec{X})$, et définie positive, c'est-à-dire telle que.

(III,1;8) $$B(\vec{X}, \vec{X}) > 0 \quad \text{pour} \quad \vec{X} \neq \vec{0}$$

Deux vecteurs $\vec{X}$, $\vec{Y}$, de $\vec{E}$, sont dits orthogonaux pour le produit scalaire B, si $B(\vec{X}, \vec{Y}) = 0$.

Un espace vectoriel muni d'un produit scalaire euclidien est appelé espace vectoriel euclidien. Un espace affine dont l'espace vectoriel associé est euclidien est appele espace affine euclidien. Ce qu'on appelle l'espace à 3 dimensions de la geometrie élémentaire est simplement un espace affine euclidien à 3 dimensions *.

Théorème 2 bis. Soit $\vec{E}$ un espace vectoriel sur $\mathbb{R}$, B une forme bilinéaire symétrique sur $\vec{E} \times \vec{E}$, ~~telle que l'on ait~~ $B(\vec{X},\vec{X}) \geq 0$ pour tout $\vec{X}$ de $\vec{E}$. On a alors l'inégalité de CAUCHY-SCHWARZ

(III,1;9) $$|B(\vec{X}, \vec{Y})| \leq \sqrt{B(\vec{X}, \vec{X})}\ \sqrt{B(\vec{Y}, \vec{Y})}$$

et l'inégalité de MINKOWSKI :

(III,1;10) $$\sqrt{B(\vec{X}+\vec{Y}, \vec{X}+\vec{Y})} \leq \sqrt{B(\vec{X}, \vec{X})} + \sqrt{B(\vec{Y}, \vec{Y})} \;;$$

en outre, si B est définie positive, c'est-à-dire si $B(\vec{X},\vec{X}) > 0$ pour $\vec{X} \neq \vec{0}$, il s'agit d'inégalités strictes $<$, sauf si $\vec{X}$ et $\vec{Y}$ sont proportionnels dans le cas de (III,1;9), s'ils sont proportionnels avec un coefficient de proportionnalité ≥ 0 ** dans le cas de (III, 1;10)

* Ce n'est pas tout à fait exact : en géométrie élémentaire, l'unité de longueur n'est pas nécessairement fixée. Ce n'est qu'après le choix d'une unité de longueur aue l'espace devient un espace affine euclidien

** Si $\vec{X}$ ou $\vec{Y}$ est nul, on considère qu'ils sont proportionnels, avec coefficient de proportionnalité ≥ 0. Voir Cours d'Algèbre, Chapitre III, n°7, théorème 7.1 et corollaire, Démonstration - On a toujours

$$B(\vec{X}+\lambda\vec{Y}\,,\,\vec{X}+\lambda\vec{Y}) \geqslant 0\,, \tag{III,1;11}$$

ou

$$B(\vec{X},\vec{X}) + 2\lambda\, B(\vec{X},\vec{Y}) + \lambda^2\, B(\vec{Y},\vec{Y}) \geqslant 0 \tag{III,1;12}$$

Un trinôme réel ne peut être toujours $\geqslant 0$ que s'il n'a pas 2 racines réelles, donc son discriminant est $\leqslant 0$, d'où (III,1;9). Soit B définie positive. Si $\vec{Y} \neq \vec{0}$, $B(\vec{Y},\vec{Y})$ est >0 le trinôme n'est pas dégénéré; si alors $\vec{X}$ n'est pas proportionnel à $\vec{Y}$ le vecteur $\vec{X}+\lambda\vec{Y}$ est $\neq 0$ pour tout λ réel, donc le Premier membre de (III,1;11) est toujours >0; alors le trinôme (III,1;12) est >0 quel que soit λ, donc il n'a aucune racine réelle, son discriminant est <0, et on a (III,1;9) avec $<$. On a donc toujours $<$, sauf si $\vec{Y} = \vec{0}$ ou si $\vec{Y} \neq \vec{0}$, $\vec{X} = \lambda_0\vec{Y}$, c'est-à-dire si $\vec{X}$ et $\vec{Y}$ sont proportionnels. Quant à l'inégalité de MINKOWSKI, elle équivaut à

$$B(\vec{X}+\vec{Y}\,,\,\vec{X}+\vec{Y}) \leqslant B(\vec{X},\vec{X}) + B(\vec{Y},\vec{Y}) + 2\sqrt{B(\vec{X},\vec{X})\,B(\vec{Y},\vec{Y})} \tag{III,1;13}$$

ou à

$$B(\vec{X},\vec{X}) + 2B(\vec{X},\vec{Y}) + B(\vec{Y},\vec{Y}) \tag{III,1;14}$$

$$\leqslant B(\vec{X},\vec{X}) + B(\vec{Y},\vec{Y}) + 2\sqrt{B(\vec{X},\vec{X})\,B(\vec{Y},\vec{Y})}\,,$$

ce qui résulte de (III,1;9); si B est définie positive, l'égalité ne peut avoir lieu que si $B(\vec{X},\vec{Y}) \geqslant 0$ et si on a (III,1;9) avec le signe $=$, c'est-à-dire si $\vec{X}$ et $\vec{Y}$ sont proportionnels, avec cette fois un coefficient de proportionnalité $\geqslant 0$.

Il en résulte que, si E est euclidien, la fonction $\vec{X} \to \sqrt{B(\vec{X},\vec{X})}$ est une <u>norme</u> sur $\vec{E}$; c'est elle qui sert à définir la distance en géométrie élémentaire; <u>un espace euclidien, vectoriel ou affine, est normé.</u> On a l'habitude de noter par $(\vec{X} \mid \vec{Y})$ le produit scalaire dans un espace vectoriel euclidien, et par $\|\vec{X}\|$ la norme $\sqrt{(\vec{X} \mid \vec{X})}$

Espaces vectoriels et affines hermitiens

Soient maintenant $\vec{E}$ et F des espaces vectoriels sur le corps des complexes $\mathbb{C}$. On appelle application semi-linéaire u de $\vec{E}$ dans $\vec{F}$ une application vérifiant

$$(\text{III},1;15)\quad \begin{cases} u(\vec{X}+\vec{Y}) = u(\vec{X})+u(\vec{Y}) & \text{pour } \vec{X}\in\vec{E},\ \vec{Y}\in\vec{E} \\ u(\lambda\vec{X}) = \bar{\lambda}\,u(\vec{X}) & \text{pour } \vec{X}\in\vec{E},\ \lambda\in\mathbb{C}. \end{cases}$$

Si $\vec{F} = \mathbb{C}$, u est une <u>forme semi-linéaire</u>.

On appelle <u>forme sesquilinéaire</u> * sur $\vec{E}\times\vec{E}$ une fonction $(\vec{X},\vec{Y}) \longrightarrow B(\vec{X},\vec{Y}) \in \mathbb{C}$,linéaire en $\vec{X}$ pour $\vec{Y}$ fixé, semi-linéaire en $\vec{Y}$ pour $\vec{X}$ fixé. Autrement dit :

$$(\text{III},1;16)\quad \begin{cases} B(\vec{X}_1+\vec{X}_2,\vec{Y}) = B(\vec{X}_1,\vec{Y}) + B(\vec{X}_2,\vec{Y}) \\ B(\vec{X},\vec{Y}_1+\vec{Y}_2) = B(\vec{X},\vec{Y}_1) + B(\vec{X},\vec{Y}_2) \\ B(\lambda\vec{X},\vec{Y}) = \lambda B(\vec{X},\vec{Y}), \quad \lambda\in\mathbb{C} \\ B(\vec{X},\mu\vec{Y}) = \bar{\mu}B(\vec{X},\vec{Y}), \quad \mu\in\mathbb{C}. \end{cases}$$

On appelle <u>produit scalaire</u> hermitien sur un espace vectoriel $\vec{E}$ sur le corps des complexes, une <u>forme sesquilinéaire</u> $(\vec{X},\vec{Y}) \longrightarrow B(\vec{X},\vec{Y})$, <u>hermitienne</u>, c'est-à-dire telle que $B(\vec{Y},\vec{X}) = \overline{B(\vec{X},\vec{Y})}$, et <u>définie positive</u>, c'est-à-dire vérifiant (III,1;8). Un espace vectoriel sur $\mathbb{C}$, muni d'un produit scalaire hermitien, s'appelle aussi <u>espace vectoriel hermitien</u>. Un espace affine sur $\mathbb{C}$, dont l'espace vectoriel associé est hermitien, s'appelle <u>espace affine hermitien</u>.

<u>Théorème 2 ter</u>** - <u>On a un énoncé analogue à 2 bis, en remplaçant</u> $\mathbb{R}$ par $\mathbb{C}$, et "<u>bilinéaire symétrique</u>" par "<u>sesquilinéaire hermitienne</u>".

<u>Démonstration</u> - En effet, au lieu de (III,1;12),on a cette fois, pour tout λ complexe :

$$(\text{III},1;17)\quad B(\vec{X},\vec{X}) + \lambda B(\vec{Y},\vec{X}) + \bar{\lambda} B(\vec{X},\vec{Y}) + \lambda\bar{\lambda} B(\vec{Y},\vec{Y}) \geqslant 0 .$$

Soit alors $B(\vec{X},\vec{Y}) = \rho e^{i\theta}$, $\rho = |B(\vec{X},\vec{Y})|$. Posons $\lambda = t e^{i\theta}$, t réel (non nécessairement $\geqslant 0$). (III,1;17) dit que, pour tout t réel, on a :

$$(\text{III},1,18)\quad B(\vec{X},\vec{X}) + 2|B(\vec{X},\vec{Y})|t + B(\vec{Y},\vec{Y})t^2 \geqslant 0 ,$$

ce qui redonne bien (III,1;9).

* sesqui = une fois et demi ; sesquilinéaire = linéaire + semi-linéaire.

** Voir cours d'Algèbre, Chapitre III, n°7, théorème 7.1 et corollaire.

Ici (III,1;10) est équivalent à :

(III,1;18 bis) $$B(\vec{X},\vec{Y}) + B(\vec{Y},\vec{X}) \leq 2\sqrt{B(\vec{X},\vec{X})\,B(\vec{Y},\vec{Y})}$$ ou

(III,1;18 ter) $$\mathcal{R}\big(B(\vec{X},\vec{Y})\big) \leq \sqrt{B(\vec{X},\vec{X})\,B(\vec{Y},\vec{Y})}$$ ($\mathcal{R}$ = partie réelle de)

ce qui résulte encore de (III,1;9).

Si B est définie positive, si $\vec{Y} \neq \vec{0}$ et $\vec{X}$ non proportionnel à $\vec{Y}$ on trouve encore que (III,1;18) est > o pour tout t réel, donc on a encore (III,1;9) avec <; ici encore, on ne peut donc avoir égalité que si $\vec{X}$ et $\vec{Y}$ sont proportionnels. Pour (III,1;10) on ne peut avoir égalité que si en outre $\mathcal{R}\, B(\vec{X},\vec{Y}) = |B(\vec{X},\vec{Y})|$ c'est-à-dire si $B(\vec{X},\vec{Y})$ est réel et ≥ 0, donc si le coefficient de proportionnalité est ≥ 0. Par suite un espace vectoriel ou affine hermitien est encore normé.

Les résultats que nous allons donner sont valables à la fois pour les espaces euclidiens sur $\mathbb{R}$ et les espaces hermitiens sur $\mathbb{C}$; nous les énoncerons pour les espaces hermitiens; il sera entendu que, dans le cas euclidien, les scalaires sont réels, et qu'alors semi-linéaire veut dire linéaire, sesquilindaire veut dire bilinéaire, $\bar{\lambda}$ veut dire λ .

Isomorphisme (ou semi-isomorphisme) d'un espace euclidien (ou hermitien) de dimension finie et de son dual. Soit $\overleftarrow{E}'$ le dual d un espace vectoriel quelconque $\vec{E}$ sur un corps $\mathbb{K}$ * . . Si8 est un élément de $\overleftarrow{E}'$ c'est une forme linéaire sur $\vec{E}$, soit $\vec{X} \longrightarrow \overleftarrow{\alpha}(\vec{X})$. Il est commode d'écrire $\overleftarrow{\alpha}.\vec{X}$ ou $<\overleftarrow{\alpha}, \vec{X}>$ au lieu de $\overleftarrow{\alpha}(\vec{X})$. On sait que $\overleftarrow{E}'$ est aussi un espace vectoriel, et alors $(\overleftarrow{\alpha}, \vec{X}) \longrightarrow <\overleftarrow{\alpha},\vec{X}> \in \mathbb{K}$ est une forme bilinéaire sur $\overleftarrow{E}' \times \vec{E}$, qu'on appelle la forme bilinéaire fondamentale; on l'appelle aussi oroduit scalaire de $\overleftarrow{\alpha} \in \overleftarrow{E}$, et de $\vec{X} \in \vec{E}$ mais ce produit scalaire n'a aucun rapport avec celui d'un'espace euclidien ou hermitien, car $\overleftarrow{\alpha}$ et $\vec{X}$ n'appartiennent pas au même espace vectoriel ** . Aussi faut-il distingue; soigneusement les notations $<\overleftarrow{\alpha}, \vec{X}>$ et $(\vec{X}\,|\,\vec{Y})$. Mais, si E est euclidien ou hermitien, les deux produits scalaires existent a la fois. Pour $\vec{Y}$ fixé, $\vec{X} \rightarrow (\vec{X}\,|\,\vec{Y})$, est une forme linéaire sur $\vec{E}$; il lui correspond donc un élément $\overleftarrow{Y}$ du dual, tel que

Y Il est commode de noter avec une flèche à l'envers le dual et ses éléments. Les vecteurs de $\overleftarrow{E}$ sont aussi appelés des covecteurs.

** En outre, le produit scalaire < , > est toujours bilinéaire, et non sesquilinéairs.

(III,1;19) $$\langle \overleftarrow{\gamma}, \vec{X} \rangle = (\vec{X} \mid \vec{Y}), \text{ pour tout } \vec{X} \text{ de } \vec{E}.$$

Cet élément $\overleftarrow{\gamma}$ dépend de $\vec{Y}$, nous le noterons donc $\overleftarrow{\gamma}_{\vec{Y}}$ ou $\overleftarrow{\gamma(\vec{Y})}$. Alors la semi-linéarité du second membre de (III,1;19) par rapport à $\vec{Y}$, pour $\vec{X}$ fixé, montre que $\overleftarrow{\gamma}_{\vec{Y}}$ dépend semi-linéairement de $\vec{Y}$:

(III,1;20) $$\begin{cases} \overleftarrow{\gamma}_{\vec{Y}_1+\vec{Y}_2} = \overleftarrow{\gamma}_{\vec{Y}_1} + \overleftarrow{\gamma}_{\vec{Y}_2} \\ \overleftarrow{\gamma}_{\lambda\vec{Y}} = \bar{\lambda}\, \overleftarrow{\gamma}_{\vec{Y}} \end{cases}$$

Donc $\vec{Y} \longrightarrow \overleftarrow{\gamma}_{\vec{Y}}$ est une application semi-linéaire γ de $\vec{E}$ dans son dual $\overleftarrow{E}'$. Cette application γ est injective; car, si $\overleftarrow{\gamma}_{\vec{Y}} = \overleftarrow{0}$, cela veut dire que $\langle \overleftarrow{\gamma}_{\vec{Y}} \mid \vec{X} \rangle = (\vec{X} \mid \vec{Y}) = 0$ quel que soit $\vec{X}$ de $\vec{E}$; en faisant $\vec{X} = \vec{Y}$, on voit que $(\vec{Y} \mid \vec{Y}) = 0$ donc $\vec{Y} = \vec{0}$, ce qui prouve bien que, γ est injective. Si alors **E** est de dimension finie, comme E et $\overleftarrow{E}'$ ont même dimension, cela signifie que γ est une bijection de $\vec{E}$ sur $\overleftarrow{E}'$. La donnée d'une structure euclidienne définit donc un isomorphisme entre l'espace et son dual; la donnée d'une structure hermitienne définit un semi-isomorphisme. Utilisons simplement le fait que γ est une surjection. Si $\overleftarrow{\alpha}$ est un élément du dual, c'est-a-dire une forme linéaire sur $\vec{E}$, il existe un élément et un seul $\vec{Y}$ de $\vec{E}$, tel que

$$(\vec{X} \mid \vec{Y}) = \langle \overleftarrow{\alpha} \mid \vec{X} \rangle \qquad \text{pour tout } \vec{X} \text{ de } \vec{E} ;$$

$\vec{Y}$ n'est autre que $\overrightarrow{\gamma^{-1}(\overleftarrow{\alpha})}$. Ainsi :

Théorème 2 quarto*- Si $\overleftarrow{\alpha}$ est une forme linéaire sur un espace euclidien ou hermitien $\vec{E}$ de dimension finie, il existe un vecteur $\vec{Y}$, déterminé d'une manière unique, tel que la forme $\overleftarrow{\alpha}$ soit le produit scalaire $\vec{X} \longrightarrow (\vec{X} \mid \vec{Y})$.

Bases orthonormales

On appelle base orthonormale d'un espace euclidien ou hermitien de dimension finie $\vec{E}$, une base $(\vec{e}_i)_{i\in I}$ dont les éléments sont deux à deux orthogonaux, et tous de longueur 1.

* Voir Cours d'Algèbre, Chapitre III, fin du n°4, page 107.

Théorème 2 quinto** - Tout espace vectoriel euclidien ou hermitien de dimension finie a des bases orthonormales.

Démonstration -

C'est évident si la dimension n est égale à 1 * ; car si, dans ce cas, $\{\vec{f}_1\}$ est une base quelconque, $\left\{\vec{e}_1 = \frac{\vec{f}_1}{\|\vec{f}_1\|}\right\}$ est une base orthonormale. Supposons alors démontrke l'existence d'une base orthonormale dans tout espace hermitien ou euclidien de dimension $\leqslant n-1$, et démontrons la dans un espace $\vec{E}$ de-dimension n. Soit $\vec{f}_1$ un vecteur $\neq \vec{0}$, et $\vec{e}_1 = \frac{\vec{f}_1}{\|\vec{f}_1\|}$; l'ensemble des vecteurs orthogonaux a $\vec{e}_1$ est un hyperplan $\vec{H}_1$ de $\vec{E}$ car c'est l'ensemble des vecteurs $\vec{X}$ vérifiant l'équation linéaire

$\overleftarrow{\gamma}\vec{e}_1(\vec{X}) = (\vec{X} \mid \vec{e}_1) = 0$, et la forme linéaire $\overleftarrow{\gamma}\vec{e}_1$ n'est pas $\equiv 0$ puisque $\vec{e}_1 \neq \vec{0}$ et que γ est injective. D'ailleurs cet hyperplan ne contient pas $\vec{e}_1$, car $(\vec{e}_1 \mid \vec{e}_1) > 0$. Donc $\vec{H}_1$ et la droite engendrée par $\vec{e}_1$ sont supplémentaires; $\vec{H}_1$ est appelé l'hyperplan orthogonal à $\vec{e}_1$. Alors $\vec{H}_1$ est un espace euclidien ou hermitien de dimension $n-1$; il contient donc, d'après l'hypothèse de récurrence, au moins une base orthonormale $\vec{e}_2, \ldots \vec{e}_n$, et alors $\vec{e}_1, \vec{e}_2, \ldots \vec{e}_n$, est une base orthonormale de $\vec{E}$, ce qui démontre le théorème.

Si $(\vec{e}_i)_{i \in I}$ est une base quelconque de $\vec{E}$, le produit scalaire des 2 vecteurs $\vec{X} = \sum_{i\in I} X_i \vec{e}_i$, $\vec{Y} = \sum_{i\in I} Y_i \vec{e}_i$ s'écrit :

$$(\vec{X} \mid \vec{Y}) = \sum_{i,j\in I} (\vec{e}_i \mid \vec{e}_j) X_i \overline{Y}_j = \sum_{i,j\in I} g_{i,j} X_i \overline{Y}_j \tag{III,1;21}$$

avec $g_{i,j} = (\vec{e}_i \mid \vec{e}_j)$. On a $g_{i,j} = \overline{g_{j,i}}$, et l'inégalité $(\vec{X} \mid \vec{X}) > 0$ pour $\vec{X} \neq \vec{0}$ s'écrit :

$$\sum_{i,j\in I} g_{i,j} X_i \overline{X}_j > 0 \tag{III,1;22}$$

, sauf si $X_i = 0$ pour tout i de I ; on dit que la matrice des $g_{i,j}$ est une matrice hermitienne définie positive.

* On peut même commencer la récurrence à $n = 0$. Dans un espace vectoriel de dimension 0 , une base est vide; il n'y en a qu'une, et elle est orthonormale !

** Voir Cours d'Algèbre, Chapitre III, n°5, théorème 5.3 page 113.

La base considérée est orthonormale si et seulement si la matrice des $g_{i,j}$ est la matrice identique :

$$(\vec{X} \mid \vec{Y}) = \sum_{i \in I} x_i \bar{y}_i \qquad (\text{III},1;23)$$

$$\|\vec{X}\|^2 = \sum_{i \in I} |x_i^2|$$

Espaces euclidiens ou hermitiens généralisés (**)

On a besoin, en théorie de la relativité restreinte, de la notion d'espace euclidien ou hermitien généralisé ou espace lorentzien ou espace de MINKOWSKI. C'est un espace vectoriel $\vec{E}$ de dimension finie (en physique, de dimension 4 sur $\mathbb{R}$), muni d'un produit scalaire généralisé, c'est-à-dire d'une forme sesquilinéaire hermitienne, non définie positive, Notons-le toujours $(\vec{X} \mid \vec{Y})$. On n'a pas nécessairement $(\vec{X} \mid \vec{X}) > 0$ pour $\vec{X} \neq \vec{0}$; on peut alors avoir $(\vec{X} \mid \vec{X}) = 0$ pour certains vecteurs $\vec{X}$, appelés isotropes, et dont la réunion est le cône isotrope * . On fait toutefois une hypothèse essentielle : la forme sesquilinéaire est non dégénérée, autrement dit il n'existe aucun

vecteur? $\neq \vec{0}$ totalement isotrope, c'est-à-dire orthogonal à tous les vecteurs de l'espace. Naturellement il n'y a pas d'inégalités de CAUCHY-SCHWARZ et de MINKOWSKI, et une telle structure ne définit pas de norme. Par contre l'application γ de $\vec{E}$ dans $\overleftarrow{E}'$ est toujours semi-linéaire et injective; car, si $\overleftarrow{\gamma_Y} = \overleftarrow{0}$ cela signifie que Y est orthogonal à tout $\vec{X}$ de $\vec{E}$ donc totalement isotrope donc nul; alors γ est encore bijective; si $\overleftarrow{\alpha}$ est une forme linéaire sur $\vec{E}$, il lui correspond encore un $\vec{Y}$ de $\vec{E}$ et un seul tel que $\langle \overleftarrow{\alpha}, \vec{X} \rangle = (\vec{X} \mid \vec{Y})$ pour tout $\vec{X}$ de $\vec{E}$, c'est toujours $\overrightarrow{\gamma^{-1}(\overleftarrow{\alpha})}$.

On appelle base orthonormale d'un espace euclidien ou hermitien généralisé une base $(\vec{e}_i)_{i \in I}$, dont les éléments sont deux à deux orthogonaux, et tous de carré scalaire ± 1.

Il existe toujours de telles bases . Le même procédé de recurrence que pour la demonstration du théorème 2 quinto permet d'en construire , mais avec quelques précautions.

* Il y a une différence importante avec les isotropes de Mathématiques Spéciales : il s'agit en physique d'un espace vectoriel sur $\mathbb{R}$, le produit scalaire est réel.

** Ces espaces sont appelés, dans le cours d'Algèbre : pseudo-euclidiens et pseudo-hermitiens, Chapitre III, n°8, définition 8.2, page 130. Ce que nous appelons ici bases orthonormales est appelé là bases pseudo-orthonormales.

Pour $n=1$, on part toujours d'un $\vec{f}_1 \neq \vec{0}$ quelconque; dire que $\vec{f}_1$ n'est pas totalement isotrope équivaut, pour $n=1$, à dire qu'il n'est pas isotrope, puisque tous les vecteurs lui sont proportionnels ; on peut alors prendre $\vec{e}_1 = \frac{\vec{f}_1}{\sqrt{|(\vec{f}_1 \mid \vec{f}_1)|}}$. et on a bien $(\vec{e}_1 \mid \vec{e}_1) = \pm 1$.

Mais, dans le passage de $n-1$ à n les choses sont plus délicates. Montrons d'abord qu'il existe au moins un vecteur $\vec{f}_1$ qui n'est pas isotrope. Si tout vecteur était isotrope, c'est-à-dire orthogonal à lui-même, l'égalite

(III,1;24) $$(\vec{X}+\vec{Y} \mid \vec{X}+\vec{Y}) = (\vec{X} \mid \vec{X}) + (\vec{Y} \mid \vec{Y}) + 2\mathcal{R}(\vec{X} \mid \vec{Y})$$

montre que $\mathcal{R}(\vec{X} \mid \vec{Y})$ serait nul quels que soient $\vec{X}$ et $\vec{Y}$; mais $\mathcal{R}(\vec{X} \mid i\vec{Y}) = \mathcal{J}m(\vec{X} \mid \vec{Y})$ si $\mathbb{K} = \mathbb{C}$, donc on aurait toujours $(\vec{X} \mid \vec{Y}) = 0$; alors tout vecteur serait totalement isotrope, ce qui est contraire à l'hypothèse; on peut donc bien trouver un vecteur $\vec{f}_1$ non isotrope, et poser $\vec{e}_1 = \frac{\vec{f}_1}{\sqrt{|(\vec{f}_1 \mid \vec{f}_1)|}}$. L'hyperplan $\vec{H}_1$ orthogonal à $\vec{e}_1$ ne contient pas $\vec{e}_1$, puisque $\vec{e}_1$ n'est pas isotrope; $\vec{H}_1$ est donc encore supplémentaire de la droite engendrée par $\vec{e}_1$.

D'autre part, sur $\vec{H}_1$ le produit scalaire est encore non dégénéré: car, s'il existait un vecteur $\neq \vec{0}$ de $\vec{H}_1$ orthogonal à tous les vecteurs de $\vec{H}_1$, comme il serait aussi orthogonal à $\vec{e}_1$, il serait orthogonal à tous les vecteurs de $\vec{E}$ donc totalement isotrope, ce qui est impossible. Donc $\vec{H}_1$ est un espace euclidien ou hermitien généralisé de dimension $n-1$, donc il a une base orthonormale $\vec{e}_2, \ldots \vec{e}_n$, et comme $\vec{e}_1 \notin \vec{H}_1$, $\vec{e}_1, \vec{e}_2, \ldots \vec{e}_n$ est bien une base orthonormale de $\vec{E}$.

Si alors $(\vec{e}_i)_{i \in I}$, est une base orthonormale de $\vec{E}$, le produit scalaire $(\vec{X} \mid \vec{Y})$, avec $\vec{X} = \sum_{i \in I} X_i \vec{e}_i$, $\vec{Y} = \sum_{i \in I} Y_i \vec{e}_i$, s'écrit, si l'on pose $(\vec{e}_i \mid \vec{e}_i) = \varepsilon_i = \pm 1$,

(III,1;25) $$\begin{cases} (\vec{X} \mid \vec{Y}) = \sum_{i \in I} \varepsilon_i X_i \overline{Y}_i, & \text{et} \\ (\vec{X} \mid \vec{X}) = \sum_{i \in I} \varepsilon_i |X_i^2|, \end{cases}$$

On démontre en outre un théorème important, appelé <u>loi d'inertie : le nombre p des ε_i qui sont > 0 et le nombre q de ceux qui sont < 0 est indépendant de la base orthonormale choisie</u>. Considérons en effet une base particu-

lière $(\vec{e}_i)_{i\in I}$, et soit J (resp K) le sous-ensemble de I formé des i pour lesquels $\varepsilon_i = +1$ (resp. -1). Soit p (resp. q) le nombre des éléments de J (resp. K)
Il existe au moins un sous-espace vectoriel $\vec{F}$ de $\vec{E}$, de dimensions p, sur lequel le produit scalaire est défini positif : à savoir celui qui est engendré par les $\vec{e}$, $i \in 1$. Mais il n'existe pas de sous-espace vectoriel $\vec{G}$ de dimension $> p$ ayant la même propriété. Si en effet $\vec{G}$ est n'importe quel sous-espace vectoriel de dimension $> p$ il coupe nécessairement le-sous-espace vectoriel de dimension $n-p$ engendré par les e_i, $i \in K$ suivant un sous-espace vectoriel non réduit à $\vec{0}$, puisque la somme des dimension2 de ces sous-espaces est $> n$; si $\vec{X}$ est un vecteur $\neq 0$ de cette intersection, on a nécessairement $(\vec{X}|\vec{X}) < 0$, donc il est impossible que, sur $\vec{G}$, le produit scalaire soit défini positif. Ainsi les nombres p et $q = n-p$ on une définition intrinsèque, indépendante de la base initialement choisie; p (resp. q) est la dimension maxima des sous-espaces vectoriels de $\vec{E}$ sur lesquels le produit scalaire est défini positif (resp. défini négatif).

En résumé. nous pouvons énoncer :

Théorème 2 sexto* - Tout espace euclidien ou hermitien généralisé a des bases orthonormales. Le nombre des vecteurs d'une telle base, dont le carré scalaire est $+1$ (resp -1), est independant de la base choisie; c'est la dimension maxima des sous-espaces vectoriels sur lesquels le produit scalaire est défini positif (resp. défini négatif).

En physique, dans la théorie de la relativité restreinte, l'univers physique d'espace-temps est un espace affine E_4 à quatre dimensions sur le corps des réels; son espace vectoriel associé $\vec{E}_4$ est muni d'un produit scalaire pour lequel $p = 3$, $q = 1$. Un référentiel galiléen de **E** est un référentiel formé d'une origine et de 4 vecteurs $\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_0$, formant une base orthonormale de $\vec{E}_4$, avec $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 1$, $\varepsilon_0 = -1$.
Un vecteur d'espace-temps $\vec{X} \in \vec{E}$ a alors 4 coordonnées $X_1, X_2, X_3, X_0 = cT$, où T est la coordonnée temps pour le référentiel galiléen considéré et c la vitesse de la lumière; son carré scalaire est $X_1^2 + X_2^2 + X_3^2 - X_0^2 = L^2 - c^2T^2$, L étant la longueur spatiale du vecteur pour le référentiel galiléen considéré.

* Voir Cours d'Algèbre, Chapitre **III**, n°7, théorème 7.4, page 117.

§ 2 FONCTIONS RÉELLES D'UNE VARIABLE RÉELLE. CONTINUITÉ A DROITE, A GAUCHE

Soit Ω une partie de la droite réelle $\mathbb{R}$, et F un espace topologique quelconque. On dit qu'une application f de Ω dans F est continue à droite en un point a de Ω, si $f(x)$ tend vers $f(a)$ lorsque x tend vers a. dansa, par valeurs $\geqslant a$. Si a- est isolé à droite dans Ω, c'est-à-dire s'il existe un nombre $\eta > 0$ tel que l'intervalle $]a, a+\eta[$ ne contienne aucun point de Ω , alors toute application de Ω dans F est continue à droite au point a . L'application f est continue à droite en a , si et seulement si la restriction de f à la partie $x \geqslant a$ de Ω est continue au point a., .

Même définition pour la continuité à gauche, Alors l'application f est continue au point a, si et seulement si elle est à la fois continue à droite et à gauche. Si cl. est isolé à gauche, la continuité en a est équivalente à la continuité à droite, Dans la suite nous donnerons tous les énoncés en supposant que Ω est un ouvert de $\mathbb{R}$. Cela laisse de côté un cas tout de même important dans la pratique : celui où Ω est un intervalle semi-ouvert ou fermé. La plupart des théorèmes seront encore valables dans ce cas, avec éventuellement de petites modifications que le lecteur fera de lui-même.

Discontinuité de première espèce. Fonctions réglées

On dit que a est une discontinuité de première espèce, de la fonction f , si $f(x)$ tend-vers une limite, notée $f(a+0)$, lorsque x tend vers a par valeurs strictement supérieures et si $f(x)$ tend vers une limite, notée $f(a-0)$, lorsque x tend vers a par valeurs strictement inférieures, ces deux limites n'étant pas toutes deux égales à la valeur de f au point a. Si f est continue ou a une discontinuité de première espèce au point a , et si l'espace F est métrique, la quantité $\operatorname{Sup}[d(f(a), f(a-0)), d(f(a),f(a+0)), d(f(a-0), f(a+0))]$ s'appellera l'oscillation de f au pointa . Elle est nulle, si et seulement si f est continue au pointa . Si F est un espace affine, on peut calculer la différence $f(a+0) - f(a-0)$, qui est un élément de l'espace vectoriel associé $\vec{F}$, et qu'on appellera le saut de f au pointa. . Ce saut ne fait pas intervenir la valeur $f(a)$ de f au point a lui-même. Il peut donc être nul même si f est discontinue : $f(a+0)$ et $f(a-0)$ peuvent être égaux sans être égaux à $f(a)$.

Naturellement une fonction f qui est discontinue en un point a n'a pas en général une discontinuité de première espèce, en ce sens que les deux limites $\lim_{x>a, x\to a} f(x)$, $\lim_{x<a, x\to a} f(x)$ n'existent pas nécessairement.

Par exemple, la fonction égale à $\sin\frac{1}{x}$ pour $x \neq 0$

et à 0 pour $x = 0$ est discontinue à l'origine, mais n'y a pas une discontinuité de première espèce.

Théorème 3 - Si une application f d'un ouvert Ω de la droite réelle $\mathbb{R}$ dans un espace métrique F n'a que des points de continuité et des discontinuités de première espèce, elle est partout continue, sauf au plus en une infinité dénombrable de points de Ω.

Démonstration - Bornons-nous à le montrer pour $\Omega = \mathbb{R}$. Soit c un point quelconque de $\mathbb{R}$; l'oscillation $\omega(c)$ au point c peut être quelconque, mais nécessairement l'oscillation $\omega(x)$ en x tend vers 0 , lorsque x tend vers c par valeurs distinctes de c. . En effet, d'après la définition même de $f(c+0)$, quel que soit $\varepsilon > 0$, il existe $\eta > 0$ tel que l'inégalité $c < x < c+\eta$ entraîne l'inégalité $d(f(x), f(c+0)) \leqslant \frac{\varepsilon}{2}$. Alors, pour tout point y et tout point z de l'intervalle $]c, c+\eta[$, on a $d(f(y), f(z)) \leqslant d(f(y), f(c+0)) + d(f(c+0), f(z)) \leqslant \varepsilon$. En prenant $y = x$ et en faisant tendre z vers 5 par valeurs strictement supérieures puis strictement inférieures, on aura $d(f(x+0), f(x)) \leqslant \varepsilon$, $d(f(x-0), f(x)) \leqslant \varepsilon$ .En faisant tendre $y < x$ et $z > x$ vers x , on aura $d(f(x-0), f(x+0)) \leqslant \varepsilon$; d'où finalement $\omega(x) \leqslant \varepsilon$ pour $c < x < c+\eta$.
En agissant de même à gauche du point c , on voit que notre assertion est vraie. Considérons alors l'intervalle $[-n, +n]$. Dans cet intervalle, l'ensemble des points ou l'oscillation est $\geqslant \frac{1}{k}$ est nécessairement fini. Si en effet il n'en était pas ainsi, on pourrait trouver une suite infinie de points distincts de cet intervalle, en chacun desquels l'oscillation serait $\geqslant \frac{1}{k}$. Comme cet intervalle est compact, on pourrait extraire de cette suite une suite partielle qui serait convergente vers un point C de l'intervalle, et formée d'éléments tous distincts de c (théorème 25 du chapitre II). Or, d'après ce que nous venons de voir, l'oscillation en ces points, qui est toujours $\geqslant \frac{1}{k}$, devrait tendre vers 0 , ce qui serait contradictoire; ainsi l'ensemble des points de l'intervalle $[-n, +n]$ où l'oscillation $\geqslant \frac{1}{k}$ est bien nécessairement fini. En prenant la réunion de ces ensembles finis de points exceptionnels pour $k = 1, 2$, etc, on voit que, dans l'intervalle $[-n, +n]$, l'ensemble des points x où l'oscillation $\omega(x)$

est > 0 est nécessairement au plus dénombrable; et par suite, sur la droite entière $\mathbb{R}$, l'ensemble des points où l'oscillation est >0, est une réunion dénombrable d'ensembles au plus dénombrables et par suite est lui-même au plus dénombrable. En tout point x n'appartenant pas à cet ensemble exceptionnel au plus dénombrnhle, $\omega(x) = 0$ et f est continue.

On appelle <u>fonction réglée,</u> définie sur une partie de la droite $\mathbb{R}$, à valeurs clans un espace topo10 - gique F , une fonction qui n'a que des points de continuité ou des discontinuités de première espèce. D'après ce que nous venons de voir, elle est en fait partout continue, sauf au plus en une infinité dénombrable de points exceptionnels, si **F** est métrisable.

<u>Remarque</u> . Cette infinité dénombrable de pointsexceptionnels peut effectivement se présenter.

Considérons par exemple la fonction réelle f d'une variable réelle, définie comme suit :

$$\begin{cases} f(x) = 0 , \text{ si } x \text{ est irrationnel} \\ f\left(\frac{p}{q}\right) = \frac{1}{q} , \text{ si } \frac{p}{q} \text{ est une fraction ratinnnelle} \end{cases}$$

irréductible, $q > 0$.

On voit facilement que, -si x tend vers a par valeurs distinctes de a , $f(x)$ tend vers 0 . Soit en effet $\varepsilon > 0$ donné. soit q_0 un entier $\geqslant \frac{1}{\varepsilon}$ Les nombres rationnels de dénominateur $< q_0$, situés dans l'intervalle$[a-1, a+1]$, sont en nombre fini; donc il existe $\eta > 0$ tel que tout nombre rationnel de l'intervalle $[a-\eta, a+\eta]$ ait un dénominateur $\geqslant q_0$, sauf peut être a. lui-même s'il est rationnel. Alors $|x-a| \leqslant \eta$, $x \neq a$, entraîne $f(x) \leqslant \varepsilon$, ce qui prouve notre affirmatinn. Alors f est continue en a si et seulement si $f(a) = 0$ c'est-à-dire si et seulement si a est irrationnel. Si a est rationnel, f admet en a une discontinuité de première espèce, avec $f(a+0) = f(a-0) = 0$. La fonction f est donc réglée, et admet bien une infinité dénombrable (et dense) de points de discontinuité. Son saut est partout nul (mais pas son oscillation !).

Si F est un espace topologique non métrisable une fonction f définie sur $\mathbb{R}$ à valeurs dans F peut être réglée et partout discontinue.

Dérivée d'une fonction réelle de variable réelle

Soit f une application d'un ouvert Ω de $\mathbb{R}$ dans $F = \mathbb{R}$. On appelle <u>dérivée de f en un point a de Ω</u>, la limite, si elle existe, de $\frac{f(a+h)-f(a)}{h}$, lorsque h tend vers 0 par valeurs $\neq 0$ de manière que $a+h$ appartienne à Ω. Comme Ω est ouvert, $a+h$ est dans Ω pour $|h|$ assez petit. Si cette limite existe seulement lorsque h tend vers 0 par valeurs > 0 on l'appellera la dérivée à droite de f. On peut parler de même de la dérivée à gauche de f au point a. La dérivée de f existe au point a, si et seulement si la dérivée à gauche et la dérivée à droite existent toutes deux et sont égales *.

Une fonction qui admet une dérivée en a est continue en a ; si elle admet une dérivée à droite en a, elle est continue à droite en a. L'existence et la valeur de la dérivée de f en $a \in \Omega$ ne dépendent que des valeurs de f au voisinage de a. La dérivée de f en a se note généralement $f'(a)$ ou $\frac{df}{dx}(a)$ ou $Df(a)$. Si la dérivée existe partout dans Ω, la fonction $x \longrightarrow f'(x)$ s'appelle fonction dérivée de ou simplement dérivée. On la note f', ou $\frac{df}{dx}$, ou Df. On peut alors chercher si, à son tour, elle admet une dérivée ; d'où les notions de dérivée seconde, troisième, etc .... La dérivée d'ordre m en a se note $f^{(m)}(a)$, ou $\frac{d^m f}{dx^m}(a)$, ou $\left(\frac{d}{dx}\right)^m f(a)$, ou $D^m f(a)$; la fonction dérivée d'ordre m se note $f^{(m)}$, ou $\frac{d^m f}{dx^m}$, ou $\left(\frac{d}{dx}\right)^m f$, ou $D^m f$.

Bien noter qu'on ne parle de la dérivée d'ordre m en $a \in \Omega$ que si toutes les dérivées d'ordre $\leq m-1$ existent, sinon dans Ω tout entier, au moins dans un même voisinage de a dans Ω.

On dit que f est m fois dérivable dans Ω si elle admet une dérivée d'ordre m en tout point de Ω ; elle est alors à fortiori k fois dérivable, pour $k \leq m$, et ses dérivées d'ordre $\leq m-1$ sont continues dans Ω. On dit que f est m fois continuement dérivable ou de classe C^m, si elle admet des derivees continues jusqu'à l'ordre m inclusivement. Elle est alors à fortiori de classe C^k, pour $k \leq m$. Une fonction m fois dérivable est sûrement de classe C^{m-1}. On dit que f est indéfiniement dérivable ou de classe C^∞, si elle admet des dérivées successives de tous les ordres; elles sont alors toutes continues.

* Si, au lieu d'un ouvert, on prend pour Ω un intervalle fermé $[a,b]$, on dira encore que f a une dérivée au point a (resp. b) si elle admet une dérivée à droite (resp. à gauche) en ce point; d'ailleurs cela correspond bien à la défintion générale de la dérivée, puisqu'on considère les valeurs de h pour lesquelles $a+h$ (ou $b+h$) est dans Ω.

Il existe un certain nombre de questions dans lesquelles une notation confuse risque d'aboutir a des erreurs. Que signifie par exemple la notation $f'(2x)$?

Signifie-t-elle la dérivée de la fonction $x \longrightarrow f(2x)$, ou la valeur de la dérivée f' de f, au point d'abscisse $2x$? <u>Nous considérerons toujours que c'est cette dernière</u> **signification** <u>qui est valable</u>. Quand nous voudrons écrire la dérivde de la fonction $x \longrightarrow f(2x)$ au point a, nous l'écrirons $\left(f(2x)\right)'_{x=a}$, et la fonction dérivée de la fonction $x \longrightarrow f(2x)$ s'écrira $\left(f(2x)\right)'$; on a donc

$\left(f(2x)\right)' = 2f'(2x)$, $\left(f(2x)\right)^{(m)} = 2^m f^{(m)}(2x)$. On distinguera de même $D^m f(2x)$ ou $(D^m f)(2x) = f^{(m)}(2x)$, et $D^m\left(f(2x)\right) = \left(f(2x)\right)^{(m)}$

On peut naturellement parler d'une dérivée ayant la valeur $+\infty$ ou la valeur $-\infty$, et de même pour une dérivée à gauche et une dérivée à droite; cependant, <u>sauf mention expresse du contraire, quand nous parlerons d'une dérivée, il s'agira toujours d'une dérivée de valeur finie.</u>

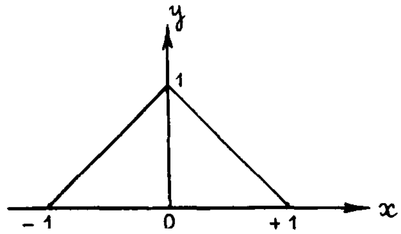
soit f une fonction réelle définie sur un intervalle $[a,b]$ de $\mathbb{R}$. On dit qu'elle est de classe C^m par morceaux si elle est réglée, et s'il existe des points $a_0 = a$, $a_1, \cdots a_{n-1}$, $a_n = b$, $a_i < a_{i+1}$, tels que, si f_i est la fonction égale 3 f dans $]a_i, a_{i+1}[$, à $f(a_i+0)$ en a_i, à $f(a_{i+1}-0)$ en a_{i+1}, est de classe C^m dans $[a_i, a_{i+1}]$ Par exemple, si le graphe de la fonction $y = f(x)$ est une ligne polygonale, f est continue, non dérivable, mais C^∞ par morceaux.

<u>Théorème 4 (Rolle)- Soit f une fonction réelle continue sur un intervalle fermé borné $[a,b]$</u>, <u>ayant en tout point de l'intervalle ouvert $]a,b[$ une dérivée, finie, ou égale à $+\infty$,ou égale à $-\infty$.</u> Si alors $f(a) = f(b) = 0$, i<u>l existe au moins un</u> <u>point c de $]a,b[$ tel que</u> $f'(c) = 0$.

<u>Démonstration</u> Si f est identiquement nulle, le résultat est évident. Sinon elle, elle prend au moins une valeur > 0 ou au moins une valeur < 0 . Plaçons-nous par exemple dans le premier cas : f étant une fonction continue sur le compact $[a,b]$, il existe au moins un point C de $]a,b[$ où elle est maxima; alors, en un tel point, la dérivée à droite est nécessairement $\leqslant 0$, tandis que la dérivée à gauche est nécessairement $\geqslant 0$. Comme la fonction est supposée, en un tel point, avoir une dérivée, ces dérivées à gauche et a droite sont égales, et la dérivée $f'(c)$ est nulle.

2

Remarque - On voit pourquoi il n'est pas nécessaire de supposer f dérivable en a ou b, ni de supposer la dérivée finie. Par contre il est essentiel de supposer l'existence d'une dérivée et non seulement d'une dérivée à droite et d'une dérivée à gauche. La fonction $f(x) = |1-x|$, dans l'intervalle $[-1, +1]$, est continue, nulle aux extrémités, et a en tout point une dérivée à droite et une dérivée à gauche; au point $x = 0$, où f est maxima, la dérivée à droite vaut -1 et la dérivée à gauche $+1$; en aucun point la dérivée à droite ou la dérivée à gauche n'est nulle. Si on suppose seulement que f a dans $]a,b[$, une dérivée à droite f'_d, on peut seulement conclure à l'existence d'un point c_1 de $]a,b[$ tel que $f'_d(c_1) \leq 0$, et d'un point c_2 de $]a,b[$ tel que $f'_d(c_2) \geq 0$. *



$y = |1-x|$, $|x| \leq 1$

Théorème 5 (Formule des accroissements finis)

Si f est une fonction réelle continue sur un intervalle $[a,b]$, admettant une dérivée, finie ou égale à $+\infty$ ou à $-\infty$, en tout point de $]a,b[$, il existe au moins un point c de $]a,b[$ tel que

$$(\mathrm{III},2;1) \qquad f'(c) = \frac{f(b)-f(a)}{b-a}.$$

La formule des accroissements finis s'écrit souvent de la manière suivante :

$$(\mathrm{III},2;2) \qquad \frac{f(x+h)-f(x)}{h} = f'(x+\theta h),$$

où θ est un nombre de l'intervalle $]0,1[$.

Remarque - Si l'on suppose seulement l'existence d'une dérivée à droite f'_d, on pourra affirmer l'existence d'un point c_1 tel que $f'_d(c_1) \leq \frac{f(b)-f(a)}{b-a}$, et d'un point c_2 tel que $f'_d(c_2) \geq \frac{f(b)-f(a)}{b-a}$.

*) La démonstration est plus délicate que la précédente. Voir la remarque qui suit le lemme du théorème 13.

Théorème 6 - (Formule de TAYLOR)

Si f est une fonction m fois continuement dérivable dans l'intervalle $[a,b]$, admettant en tout point de $]a,b[$ une dérivée d'ordre $m+1$, finie ou égale à $+\infty$ ou à $-\infty$, il existe au moins un point c de $]a,b[$ tel que

$$(\mathrm{III},2;3)\quad f(b)-f(a)-(b-a)f'(a)-\frac{(b-a)^2}{2!}f''(a)-\dots-\frac{(b-a)^m}{m!}f^{(m)}(a) = \frac{(b-a)^{m+1}}{(m+1)!}f^{(m+1)}(c).$$

On écrit souvent cette formule sous la forme suivante :

$$(\mathrm{III},2;4)\quad f(x+h)=f(x)+h f'(x)+\frac{h^2}{2!}f''(x)+\dots+\frac{h^m}{m!}f^{(m)}(x)+\frac{h^{m+1}}{(m+1)!}f^{(m+1)}(x+\theta h),$$

où θ est un nombre de $]0,1[$.

La formule des accroissements finis est évidemment un cas particulier de la formule de TAYLOR. Celle-ci se démontre de la façon suivante : on considère la fonction

$$(\mathrm{III},2;5)\quad g(x)=f(b)-f(x)-(b-x)f'(x)-\dots-\frac{(b-x)^m}{m!}f^{(m)}(x)-\frac{(b-x)^{m+1}}{(m+1)!}\lambda,$$

où le nombre λ est déterminé de façon que $g(a)=0$.

Comme $g(b)=0$, on applique à cette fonction le théorème de ROLLE; cela donne un point c de $]a,b[$ tel que $g'(c)=0$, ou $g^{(m+1)}(c)=\lambda$, d'où le résultat.

Nous n'insistons pas sur ces théorèmes. Ce sont des théorèmes fondamentaux du calcul différentiel, mais ils ont déjà été étudiée antérieurement.

Fonctions monotones On dit qu'une fonction réelle définie sur un ouvert $\Omega\subset\mathbb{R}$ * est monotone, si elle est croissante ou si elle est décroissante. Une fonction monotone n'est pas nécessairement continue; mais, d'après un théorème connu de mathématiques spéciales, elle possède nécessairement une limite à droite $f(a+0)$ et une limite à gauche $f(a-0)$ en tout point a de Ω ** C'es deux limites ne sont pas nécessairement égales, ni égales à $f(a)$.

* Comme il a été dit page 184, Ω peut être aussi un intervalle non ouvert.

** Si par exemple f est croissante, $f(a+0)=\operatorname{Inf.}_{x>a}(f(x))$, et $f(a-0)=\operatorname{Sup.}_{x<a}(f(x))$. Alors $f(a-0)\leqslant f(a)\leqslant f(a+0)$; le saut de f en a est $\geqslant 0$.

Une fonction monotone est par conséquent réglée. On peut alors lui appliquer le theorème 3; la fonction est partout continue, sauf au plus en une infinité dénombrable de points.

Il n'est pas inutile de donner un exemple simple d'une fonction strictement croissante, ayant une infinité dénombrable dense de discontinuités. Soit h une fonction > 0 définie sur l'ensemble $\mathbf{Q}$ des nombres rationnels, et telle que la somme $\sum_{r \in \mathbf{Q}} h(r)$ soit finie (il s'agit de la somme d'une infinité dénombrable de nombres positifs, qui ne sont pas donnés dans un ordre particulier, donc au sens du chapitre II page 124). Considérons alors la fonction f définie comme suit

$$f(x) = \sum_{r \in \mathbf{Q},\, r < x} h(r) \; ; \tag{III,2;6}$$

Elle est partout définie , > 0 , et croissante; elle est même trivialement strictement croissante, puisque l'on a la formule :

$$f(x) - f(y) = \sum_{r \in \mathbf{Q},\, y \leq r < x} h(r), \qquad \text{pour } y < x . \tag{III,2;7}$$

Soit a un nombre réel quelconque. Soit $\varepsilon > 0$. Puisque la somme $\sum_{r \in \mathbf{Q}} h(r)$ est convergente, il existe un nombre fini de nombres rationnels, $r_0 , r_1 , \dots , r_n$, tels que

$$\sum_{r \in \mathbf{Q},\, r \neq r_0, r_1, \dots, r_n} h(r) \leq \varepsilon . \tag{III,2;8}$$

Il existe alors un nombre $\eta > 0$ tel que l'on ait

$$x \neq r_0, r_1, \dots, r_n , \text{ pour } a-q \leq x \leq a+\eta , \quad x \neq a . \tag{III,2;9}$$

On a alors les deux inégalités suivantes, en vertu de (III,2;7) :

$$f(a) \geq f(x) \geq f(a)-\varepsilon , \text{ pour } a - \eta \leq x < a . \tag{III,2;10}$$

$$f(a) + h(a) \leq f(x) \leq f(a) + h(a) + \varepsilon , \text{ pour } a \leq x \leq a+\eta ,$$

en convenant que $h(a) = 0$ si a est irrationnel.
Ceci montre que l'on a #(a-0)= $f(a)$, et $f(a+0) = f(a) + h(a)$.

La fonction8 est donc partout continue a gauche; elle est continue à droite, et par conséquent continue, en tout point a irrationnel; elle est discontinue en tout pointa rationnel, et son saut est alors égal à $h(a)$; elle a bien une infinité dénombrable et dense de points de discontinuité de première espèce, a savoir tous les points rationnels.

Une fonction croissante n'est donc pas nécessairement dérivable. Mais, pour les fonctions dérivables, il existe un critère très important et déjà connu de croissance.

Théorème 7 - Soit une fonction réelle, définie sur un intervalle (ouvert, semi-ouvert ou fermé) de $\mathbb{R}$, et ayant partout une dérivée, finie ou égale à $+\infty$ ou à $-\infty$. Pour qu'elle soit croissante, il faut et il suffit que sa dérivée soit partout $\geqslant 0$.

Démonstration - Si la fonction est croissante, il est bien évident que la quantité $\frac{f(x+h)-f(x)}{h}$ est $\geqslant 0$, et que par conséquent sa limite lorsque h tend vers 0, limite qui est supposée exister, est aussi $\geqslant 0$.

Réciproquement, supposons4 continue et dérivable, et de dérivée (finie ou non) partout $\geqslant 0$. Alors on a, d'après la formule des accroissements finis, $\frac{f(x+h)-f(x)}{h} = f'(x+\theta h) \geqslant 0$, ce qui prouve bien que f est croissante.

Remarques - 1°/ La même démonstration prouve que, si la dérivée f' est >0, alors f est strictement croissante. Mais il peut arriver que f soit strictement croissante, sans que sa dérivée soit partout 7 0 ; c'est ce que montre l'exemple de la fonction $f(x) = x^3$, dont la dérivée, partout $\geqslant 0$, est nulle à l'origine.

2°/ En utilisant les remarques qui suivent les théorèmes 4 et 5, on démontre le même théorème avec dérivée à droite (ou dérivée a gauche) au lieu de dérivee.

Les fonctions dérivées et le théorème des valeurs intermédiaires

La dérivée d'une fonction n'est naturellement pas nécessairemer continue; c'est ce que montre l'exemple de la fonction f définie]

$f(x) = x^2 \sin\frac{1}{x}$ pour $x \neq 0$, $f(0) = 0$; f est partout continue; elle admet en tout point $x \neq 0$ la dérivée $f'(x) = 2x\sin\frac{1}{x} - \cos\frac{1}{x}$; à l'origine elle admet un? dérivée nulle; sa dérivée est donc partout définie sur $\mathbb{R}$, mais discontinue à l'origine, car $\cos\frac{1}{x}$ ne tend pas vers0 quand $x \neq 0$ tend vers 0. Cependant, si une fonction définie sur un intervalle est dérivable, sa fonction dérivée possède, comme les fonctions continues, la propriété de ne pas pouvoir prendre deux valeurs sans prendre aussi toutes les valeurs intermédiaires. Pour simplifier supposons la fonction définie sur $\mathbb{R}$. Soit $f'(a)=\alpha$, $f'(b)=\beta$, et soit γ un nombre strictement compris entre α et β. Alors, si l'on choisit un nombre h assez petit, on a nécessairement

$$\frac{f(a+h)-f(a)}{h} = \alpha' < \gamma \quad , \text{ et } \quad \frac{f(b+h)-f(b)}{h} = \beta' > \gamma .$$

Le nombre h étant ainsi choisi et fixé, la fonction : $x \longrightarrow \frac{f(x+h)-f(x)}{h}$ est continue. Il en résulte que, prenant les deux valeurs α' et β', elle prend aussi toutes les valeurs intermédiaires, et en particulier γ. Il existe donc un point x tel que $\frac{f(x+h)-f(x)}{h} = \gamma$. Si alors nous appliquons à l'intervalle $[x, x+h]$ la formule des accroissements finis, nous trouvons bien l'existence d'un point c tel que $f'(c) = \gamma$, ce qui prouve notre affirmation.

Fonctions convexes

soit f une fonction réelle définie sur un intervalle I de $\mathbb{R}$. On dit que f est convexe si elle possède la propriété suivante :

Quels que soient les points $A_1 = (a_1, f(a_1))$,

et $A_2 = (a_2, f(a_2))$ au graphe de f dans $\mathbb{R} \times \mathbb{R}$, la corde $A_1 A_2$ est au dessus de l'arc joignant A_1 à A_2 sur le graphe de f. Cela se traduit par l'inégalité

$$(\mathrm{III},2;6) \qquad f(ta_1+(1-t)a_2) \leqslant t\,f(a_1)+(1-t)\,f(a_2), \qquad 0 \leqslant t \leqslant 1 .$$

Cela revient exactement à dire que l'ensemble des points (x,y) de $\mathbb{R}^2$ vérifiant l'inégalité $y \geqslant f(x)$ est un ensemble convexe de $\mathbb{R}^2$ (voir définition des ensembles convexes page 174).

Compte tenu de ce que nous avons dit sur les propriétés barycentriques des ensembles convexes, la relation (III,2;6) entraîne alors la relation plus générale

$$(\mathrm{III},2;7) \qquad f\left(\frac{\alpha_1 x_1+\cdots+\alpha_n x_n}{\alpha_1+\cdots+\alpha_n}\right) \leqslant \frac{\alpha_1 f(x_1)+\cdots+\alpha_n f(x_n)}{\alpha_1+\cdots+\alpha_n},$$

pour $x_1, x_2, \ldots, x_n \in I$, $\alpha_1, \alpha_2, \ldots, \alpha_n \geqslant 0$.

Théorème 7 bis - Pour qu'une fonction réelle f définie sur l'intervalle I de $\mathbb{R}$ soit convexe, il faut et il suffit qu'elle ait les propriétés suivantes :

1°) La fonction f est continue dans l'intérieur $\mathring{I}$ de I, et en outre, si I est semi-ouvert ou fermé, elle vérifie, en l'une quelconque des extrémités ℓ de I, la relation

$$(\mathrm{III},2;8) \qquad \lim_{x\in I,\, x\to \ell} f(x) \leqslant f(\ell) \;{}^{*}$$

* ou encore elle est semi-continue supérieurement aux extrémités de I.

2°) En tout point x de $\mathring{I}$, la fonction f possède une dérivée à gauche f'_g et une dérivée a droite f'_d . ces dérivées coincident partout, sauf au plus en une infinité dénombrable de points de $\mathring{I}$; et l'on a les inégalités

$$f'_g \leqslant f'_d \ ; \quad f'_d(x_1) \leqslant \frac{f(x_2)-f(x_1)}{x_2-x_1} \leqslant f'_g(x_2) \quad \text{pour } x_1 < x_2 . \tag{III,2;9}$$

Il en résulte en particulier que, si f est une fonction dérivable, elle est convexe si et seulementsi sa dérivée Premiere est une fonction croissante; et que, si f est une fonction 2 fois dérivables, elle est convexe si et seulement si sa dé- seconde f'' est $\geqslant 0$.

Démonstration - 1°/ Supposons d'abord f convexe. Si $0 < h < k$, et $[x, x+k] \subset I$, le point $(x+h, f(x+h))$ est au-dessous du segment de droite $[(x, f(x)), x+k, f(x+k)]$. On en déduit l'inégalité suivante :

$$\frac{f(x+h)-f(x)}{h} \leqslant \frac{f(x+k)-f(x)}{k} \tag{III,2;10}$$

Cela prouve que la fonction $h \longrightarrow \dfrac{f(x+h)-f(x)}{h}$ est croissante pour $h > 0$. Elle a donc nécessairement une limite quand h tend vers 0 . Autrement dit f possède en tout point de $\mathring{I}$ une dérivée a droite finie, ou égale a $-\infty$.

En outre, on aura

$$f'_d(x) \leqslant \frac{f(x+k)-f(x)}{k} , \tag{III,2;11}$$

ce qui est exactement la première moitié du 2ème système d'inégalité (III,2;9), en posant $x = x_1$, $x+k = x_2$.

Mais on pourrait faire un raisonnement analogue avec h et $k < 0$, donc f a une dérivée, a gauche f'_g , finie ou égale à $+\infty$ en tout point de I , et l'on a la 2ème moitié du 2ème système d'inégalités (III,2;9).

(III,2;10) devient, en posant $x+h = y$, $x+k = y+l$.

2 $$\frac{f(y)-f(y-h)}{h} \leqslant \frac{f(y+l)-f(y-h)}{h+l} , \tag{III,2;12}$$

et le raisonnement analogue avec des accroissements négatifs donne

$$\frac{f(y+l)-f(y-h)}{h+l} \leqslant \frac{f(y+l)-f(y)}{l} \tag{III,2;13}$$

Retenons seulement

$$\frac{f(y-h)-h(y)}{-h} \leqslant \frac{f(y+l)-f(y)}{l} \tag{III,2;14}$$

En faisant tendre h et l vers 0 , on a $f'_g(y) \leq f'_d(y)$, ce qui achève de démontrer (III,2;9); en outre ces 2 dérivées sont finies.

La fonction f'_d est alors une fonction croissante dans $\mathring{I}$; elle a donc au plus une infinité dénombrable de points de discontinuité. Si x est un point de continuité de f'_d , on a nécessairement, pour $h > 0$, $f'_d(x-h) \leq f'_g(x) \leq f'_d(x)$, d'où l'on déduit, en faisant tendre h vers 0 , $f'_g(x) = f'_d(x)$, ainsi f admet bien une dérivée, sauf au plus en une infinité dénombrable de points.

Alors f étant en tout point dérivable a gauche et dérivable a droite, est continue à gauche et continue à droite, donc partout continue dans l'intérieur $\mathring{I}$ de l'intervalle I. Puisque f'_d est croissante, elle est ou bien toujours ≤ 0, dans $\mathring{I}$ ou bien toujours ≥ 0 , ou bien ≤ 0 strictement a gauche d'un point c et ≥ 0 strictement a droite de c . La fonction f est donc ou bien décroissante dans $\mathring{I}$, ou bien croissante dans $\mathring{I}$, ou bien elle est décroissante a gauche de c , croissante à droite de c ; elle admet alors un minimum au point c , puisqu'elle y est continue. De toute façon elle a donc nécessairement une limite a droite a l'extrémité gauche de I , et une limite à gauche a l'extrémité droite de I .

Considérons par exemple le cas de l'extrémité droite b, et supposons $b \in I$. Si $b' < x < b$, le point $(x, f(x))$ doit être au-dessous de la corde $[(b', f(b')), (b, f(b))]$. En passant à la limite lorsque x tend vers b , on voit qu'il en est encore de même du point $(b, f(b-0))$; ce qui signifie bien que l'on a l'inégalité $f(b-0) \leq f(b)$. Ceci achève de démontrer toutes les propriétés de f données dans l'énoncé du théorème.

2°/ Inversement supposons que f vérifie toutes ces propriétés; toutefois, en ce qui concerne le 2ème système d'inégalités (III,2;9), nous supposerons seulement vérifié $f'_d(x_1) \leq f'_g(x_2)$. Nous allons démontrer qu'elle est convexe dans l'intérieur $\mathring{I}$ de I , et alors l'inégalité (III,2;8) montrera qu'elle l'est dans I si celui-ci est semi ouvert ou fermé. Soient donc a, b deux points de $\mathring{I}$, $a < b$.

La fonction g, définie par $g(x) = f(x) - f(a) - (x-a)\dfrac{f(b)-f(a)}{b-a}$ vérifie encore toutes ces propriétés dans $\mathring{I}$ Mais en outre $g(a) = g(b) = 0$. Pour démontrer la convexité de f , nous devons simplement montrer que, pour $a \leq x \leq b$, on a $g(x) \leq 0$. S'il n'en était pas ainsi, la fonction g possèderait dans l'intervalle $]a, b[$ un maximum > 0 en un point c.

Le raisonnement donné au théorème 4 de Rolle montrerait que $g'(c) \geqslant 0$ et $g'_d(c) \leqslant 0$. L'inégalité (III,2;9) (1er système) montrerait alors que ces deux quantités sont nulles, Mais comme g'_d est croissante par (III,2;9), elle serait nécessairement $\geqslant 0$ dans l'intervalle $[c, b]$. Le maximum $g(c) > 0$ devrait donc être $\leqslant g(b) = 0$, ce qui serait contradictoire. Il est donc oien démontré que f est convexe.

si f est dérivable dans I , les conditions précédentes reviennent bien à dire que f' est croissante; si elle est 2 fois dérivables, que f'' est $\geqslant 0$.

§ 3 DÉRIVÉE D'UNE APPLICATION D'UN ESPACE AFFINE DANS UN AUTRE. VECTEUR DÉRIVÉ D'UNE FONCTION D'UNE VARIABLE SCALAIRE

Considérons une application4 d'un ouvert Ω du corps des scalaires $\mathbb{K}$ dans un espace affine normé F * .On peut alors donner'un sens, pour $a \in \Omega$, à la formule

$$(\mathrm{III},3;1) \qquad \overrightarrow{f'(a)} = \lim_{h \neq 0,\, h \to 0,\, a+h \in \Omega} \frac{\overrightarrow{f(a+h) - f(a)}}{h} \in \vec{F} .$$

Dans le deuxième membre, nous avons d'abord la différence $\overrightarrow{f(a+h) - (a)}$ de deux points de F , qui est un vecteur de l'espace vectoriel associé $\vec{F}$. On peut diviser ce vecteur par le scalaire $h \neq 0$, et l'on peut chercher la limite de ce vecteur dans $\vec{F}$ lorsque h tend vers 0 , puisque l'espace vectoriel $\vec{F}$ est supposé normé. Si $\overrightarrow{f'(a)}$ existe, on 1 appelle le vecteur dérivé ou la dérivée de f en a . L 'existence de la dérivée et sa valeur ne dépendent pas de la norme,mais seulement de la topologie de F , puisqu'il en est ainsi de la notion de limite. On peut de même parler de dérivée à gauche et de dérivée à droite, si $\mathbb{K} = \mathbb{R}$. On peut ensuite considérer la fonction dérivée $\overrightarrow{f'} : x \to \overrightarrow{f'(x)}$, si la dérivée existe partout dans Ω , c'est une application de Ω dans l'espace vectoriel normé $\vec{F}$. On peut ensuite prendre les dérivées ultérieures, dans les mêmes conditions qu'au § 2; elles se noteront de la même manière que pour les fonctions réelles (à savoir : $\overrightarrow{f''}, \ldots \overrightarrow{f^{(m)}}, \ldots$ etc...); ce sont toutes, si elles existent, des applications de Ω dans $\vec{F}$. Notons que f prend ses valeurs dans l'espace

* $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$; F est supposé affine sur $\mathbb{K}$. Si $\mathbb{K} = \mathbb{R}$, et si F est donné comme affine sur $\mathbb{C}$, on se bornera à considérer F comme affine sur $\mathbb{R}$.

affine F, et que ses dérivées $\vec{f}', \vec{f}'', \dots$ prennent leurs valeurs dans l'espace vectoriel associé $\vec{F}$. Si $E = F = \mathbb{R}$, on retombe sur la dérivée usuelle d'une fonction réelle d'une variable réelle.

On pourra alors parler des classes $C^1, C^2, \dots C^m, \dots, C^\infty$, de fonctions 1 fois, 2 fois,.... m fois continuement dérivables ou indéfiniment dérivables, à valeurs dans F. Ici comme au § 2, une fonction dérivable est continue: et on ne parlera de dérivée seconde en $a \in \Omega$ que si la dérivée première est définie, sinon dans Ω tout entier, au moins dans tout un voisinage de a. Un exemple particulièrement important nous est fourni par la mécanique. Dans ce cas, $\mathbb{K} = \mathbb{R}$ la variable x est la variable temps t, et l'espace F est l'espace affine ordinaire à trois dimensions. Le mouvement d'une particule est alors une fonction $t \longrightarrow M(t)$, fonction de la variable réelle t à valeurs dans F. Sa dérivée première $\overrightarrow{dM}/dt$ est le vecteur vitesse; sa dérivée seconde $\overrightarrow{\frac{d^2M}{dt^2}}$ le vecteur accélération. Ils appartiennent à $\vec{F}$.

Si F est de dimension finie, et si l'on y a choisi un référentiel, formé d'une origine b et d'une base $(\vec{f}_i)_{i\in I}$ de $\vec{F}$, alors la position de tout point de F se représente par ses coordonnées $(y_i)_{i\in I}$, et la donnée de la fonction f définie sur $\Omega \subset \mathbb{K}$ à valeurs dans F devient équivalente à la donnée des fonctions scalaires $(F_i)_{i\in I}$, avec la formule :

(III,3;2) $$f(x) = b + \sum_{i\in I} F_i(x)\vec{f}_i \; ; \quad \text{ou} \quad y_i = F_i(x) .$$

Dans ces conditions, la fonction dérivée est donnée par la formule :

(III,3;3) $$\overrightarrow{f'(x)} = \sum_{i\in I} F_i'(x)\vec{f}_i .$$

<u>Pour qu'une fonction à valeurs dans un espace affine normé de dimension finie soit dérivable, il faut et il suffit que ses composantes, sur un réferentiel quelconque, soient des fonctions scalaires dérivables, et les composantes de la dérivée sont les dérivées des composantes.</u>

Cas général: dérivée partielle suivant un vecteur.

Soit maintenant f une application d'un ouvert Ω d'un espace affine normé E dans un espace affine normé F; il n'est évidemment plus possible de parler de dérivée au sens précédent. Nous introduirons d'abord la notion de <u>dérivée ou dérivée partielle suivant un vecteur $\vec{X}$ de $\vec{E}$</u>. Soit a un point de Ω. On appelle dérivée de f en a suivant le vecteur $\vec{X}$

la dérivée, si elle existe, pour $t = 0$, de la fonction $t \to f(a + t\vec{X})$. Cette dérivée se note $\overrightarrow{D_{\vec{X}} f}(a) \in \vec{F}$.

On a donc

$$(\mathrm{III},3;4) \qquad \overrightarrow{D_{\vec{X}} f}(a) = \left(\frac{d}{dt}\left(f(a + t\vec{X})\right)\right)_{t=0} = \lim_{\substack{t \neq 0,\, t \to 0 \\ a + t\vec{X} \in \Omega}} \frac{\overrightarrow{f(a+t\vec{X}) - f(a)}}{t}$$

Ici t est un scalaire; il parcourt l'ensemble $\mathbb{K}_{a,\vec{X}}$ des éléments de $\mathbb{K}$ pour lesquels $a + t\vec{X} \in \Omega$. La fonc - tion $t \longrightarrow f(a + t\vec{X})$ est donc une application de $\mathbb{K}_{a,\vec{X}}$ dans F.. $\mathbb{K}_{a,\vec{X}}$ est l'image réciproque de l'ouvert Ω par l'application continue $t \to a + t\vec{X}$ de $\mathbb{K}$ dans E c'est donc un ouvert de $\mathbb{K}$, contenant l'origine, ce qui permet de chercher une dérivée au point $t = 0$.

L'existence et la valeur de la dérivée suivant $\vec{X}$ en a dépendent seulement de la topologie de $\vec{F}$ et non de sa norme, puisqu'il en est ainsi de la notion de limite. Si $\vec{X} = \vec{0}$, la dérivée existe en tout point de Ω, et elle est nulle. Si $\mathbb{K} = \mathbb{R}$, et si on se donne une direction de demi-droite, il existe un vecteur X et un seul, de norme 1, ayant cette direction; la dérivée suivant ce vecteur $\vec{X}$ s'appelle dérivée suivant la direction considérée; elle dépend évidemment de la norme de $\vec{E}$. Par exemple, si E est euclidien de dimension finie, si S est une hypersurface régulière contenue dans Ω, a un point de S, 3 une normale orientée à S en a, la dérivée normale $\frac{d}{d\vec{\nu}}$ en a est la dérivée suivant le vecteur unitaire de $\vec{\nu}$.

<u>Si E est le corps des scalaires, et si $\vec{X}$ est l'élément 1 de ce corps, la dérivée suivant $\vec{X}$ est ce que nous avons appelé simplement la dérivée dans (III,3;1)</u> : $\overrightarrow{D_1 f}(a) = \vec{f'}(a)$. Si $\overrightarrow{D_{\vec{X}} f}(x)$ existe pour tout x, on appelle <u>fonction dérivée suivant $\vec{X}$ de f</u>, ou simplement dérivée suivant $\vec{X}$ de f, la fonction $\overrightarrow{D_{\vec{X}} f} : x \to \overrightarrow{D_{\vec{X}} f}(x)$, C'est donc, pour? fixé, une application de Ω dans $\vec{F}$. On pourra ensuite chercher si $\overrightarrow{D_{\vec{X}} f}$ admet à son tour une dérivée en a, suivant un vecteur $\vec{Y}$ (distinct ou non de $\vec{X}$); s'il en est ainsi, on la notera $D_{\vec{Y}} D_{\vec{X}} f(a)$; ce sera une dérivée partielle du second ordre.

Et ainsi de suite .

Matrice dérivée, déterminant jacobien

Si F est de dimension finie, et si b, $(\vec{f}_i)_{i\in I}$, est un référentiel de F, alors on a la formule : *

(III,3;5) $$f(x) = b + \sum_{i\in I} F_i(x)\vec{f}_i \;;\quad \overrightarrow{D_{\vec{X}} f}(x) = \sum_{i\in I} D_{\vec{X}} F_i(x)\vec{f}_i\,.$$

Supposons maintenant que E soit de dimension finie, et soit a, $(\vec{e}_j)_{j\in J}$, un référentiel de E ; alors les dérivées suivant les vecteurs $\vec{e}_j$ de la base de E sont aussi ce qu'on appelle usuellement les dérivées partielles de f ; autrement dit, on a la définition :

(III,3;6) $$\overrightarrow{\partial_j f}(x) = \frac{\overrightarrow{\partial f}}{\partial x_j}(x) = \overrightarrow{D_{\vec{e}_j} f}(x) = \lim_{t\to 0,\, t\neq 0} \overrightarrow{\frac{f(x+t\vec{e}_j)-f(x)}{t}}\,.$$

Si enfin E et F sont tous deux de dimension finie, et si on a choisi à la fois un référentiel de E et un référentiel de F, alors la dérivée $\overrightarrow{\partial_j f} = \frac{\overrightarrow{\partial f}}{\partial x_j}$ s'exprime par :

(III,3;7) $$\frac{\overrightarrow{\partial f}}{\partial x_j} = \sum_{i\in I} \frac{\partial F_i}{\partial x_j}\vec{f}_i\,.$$

La matrice des $\frac{\partial F_i}{\partial x_j}(x)$, c'est-a-dire dans le cas particulier où $I = \{1,2,\dots m\}$, $J = \{1,2,\dots n\}$:

(III,3;8) $$\begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x) & \frac{\partial F_1}{\partial x_2}(x) & \cdots & \frac{\partial F_1}{\partial x_n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial F_m}{\partial x_1}(x) & \frac{\partial F_m}{\partial x_2}(x) & \cdots & \frac{\partial F_m}{\partial x_n}(x) \end{pmatrix}$$

s'appelle matrice dérivée de f au point x de Ω. Si $m = n$, don déterminant s'appelle le déterminant jacobien de f au point x, par rapport aux référentiels considérés. On note souvent par $\frac{D(y_1,y_2,\dots y_n)}{D(x_1,x_2,\dots x_n)}$ le déterminant jacobien de la fonction $y = f(x)$ définie, dans les systèmes référentiels considérés, par les fonctions scalaires $y_i = F_i(x) = F_i(x_1,x_2,\dots x_n)$, $i = 1,2,\dots n$; on dit aussi que c'est le déterminant jacobien des n fonctions $y_i = F_i$ des n variables $x_1, x_2, \dots x_n$. Pour $n = 1$, le déterminant jacobien se réduit à la dérivée usuelle.

* Les F_i étant des fonctions scalaires, $D_{\vec{X}} F_i(x)$ est un scalaire.

Insuffisance de la dérivée suivant un vecteur

La notion de dérivée suivant un vecteur est manifestement insuffisante. En effet :

1°/ Une fonction peut avoir en tout point une dérivée partielle suivant tout vecteur, sans être pour cela nécessairement continue. Considérons en effet la fonction f scalaire définie sur $\mathbb{R}^2$ par :

$$(\mathrm{III},3;9) \qquad f(x,y) = \begin{cases} 0 \text{ pour } (x,y) = (0,0) \\ \dfrac{x^5}{(y-x^2)^2 + x^8} \text{ pour } (x,y) \neq (0,0). \end{cases}$$

Cette fonction est le quotient de 2 polynomes. Pour la dérivée suivant un vecteur, les règles de dérivation d'une somme, d'un produit, d'un quotient, sont les mêmes que pour la dérivée de fonctions d'une variable scalaire, puisque l'on se ramène à la dérivée pour $t = 0$ d'une fonction de t (formule (III,3;4)). Donc f est dérivable en tout point $\neq (0,0)$, puisqu'alors son dénominateur n'est pas nul. Cherchons sa dérivée à l'origine, suivant un vecteur (X,Y). Si $Y \neq 0$, on a, pour $t \neq 0$:

$$(\mathrm{III},3;10) \qquad f(tX,tY) = \frac{t^5X^5}{t^2Y^2 + \cdots} = \frac{X^5}{Y^2}t^3 + \cdots .$$

Comme f est nulle à l'origine, donc pour $t = 0$ sa dérivée à l'origine suivant le vecteur considéré est'nulle.

Si $Y = 0$, $X \neq 0$, on a, pour $t \neq 0$:

$$(\mathrm{III},3;11) \qquad f(tX,tY) = \frac{t^5X^5}{t^4X^4 + \cdots} = [illegible] + \cdots .$$

Comme f est nulle à l'origine, donc pour $t = 0$, sa dérivée suivant le vecteur $(X,0)$ est X . Suivant le vecteur $(0,0)$, sa dérivée est toujours nulle. Ainsi f est dérivable suivant tout vecteur. Or elle est discontinue à l'origine : sur la parabole $y - x^2 = 0$, on a

$f(x,y) = \dfrac{x^5}{x^8} = \dfrac{1}{x^3}$ pour $x \neq 0$, expression qui tend vers ∞ quand $x \neq 0$ tend vers 0 .

2°/ Si l'on ne fait aucune hypothèse de continuité sur les dérivées partielles, il peut n'exister aucune liaison entre les dérivées suivant les divers vecteurs de $\vec{E}$ en un même point de Ω . Bien entendu, si $D_{\vec{X}}f(a)$ existe, il en est de même de $\overrightarrow{D_{\lambda\vec{X}}f(a)}$, pour> scalaire, et on a

(III,3;12) $$\overrightarrow{D_{\lambda\vec{X}} f}\,(a) = \lambda\, \overrightarrow{D_{\vec{X}} f}\,(a)\ .$$

En effet, c'est évident si $\lambda = 0$; et sinon, $\overrightarrow{\dfrac{f(a+t\lambda\vec{X}) - f(a)}{t}}$, pour $t \in \mathbb{K}$, $a+t\lambda\vec{X} \in \Omega$, est de la forme $\lambda \dfrac{f(a+s\vec{X}) - f(a)}{s}$, avec $s = t\lambda \in \mathbb{K}$, $a+s\vec{X} \in \Omega$, et $s \neq 0$ tend vers 0 si $t \neq 0$ tend vers 0, ce qui donne bien (III,3;12). Mais on pourrait souhaiter que, pour a fixé, $\overrightarrow{D_{\vec{X}} f}\,(a)$ dépende linéairement de $\vec{X}$. Il n'en est évidemment rien, comme le montre le même exemple (III,3;9), dans lequel, à l'origine, la dérivée suivant le vecteur $(X,0)$ est X, alors que sa dérivée suivant (X,Y), $Y \neq 0$, est nulle.

Le défaut de la dérivée partielle suivant un vecteur est de ne faire intervenir, autour de chaque point a, que le comportement de la fonction sur les droites issues de a dans l'espace affine E, alors qu'il est bon, en réalité, de connaître un comportement global de f dans tout un voisinage de a. C'est pourquoi nous introduirons la notion d'application dérivée.

Dérivée totale ou application dérivée

Soit f une application d'un ouvert Ω d'un espace affine normé E dans un espace affine normé F ; on dit que <u>f admet, au point a de Ω, une application dérivée ou dérivée totale ou différentielle ou différentielle totale L, si L est une application linéaire continue de $\vec{E}$ dans $\vec{F}$, et si l'on a, pour $a+\vec{h} \in \Omega$</u> :

(III,3;13) $$f(a+\vec{h}) = f(a) + L.\vec{h} + \varphi(\vec{h})\,\|\vec{h}\|\ ,$$

<u>où $\varphi(\vec{h})$ tend vers $\vec{0}$ lorsque $\vec{h} \neq \vec{0}$ tend vers $\vec{0}$</u>. Cela revient encore à dire que l'accroissement :

$\overrightarrow{\Delta f} = \overrightarrow{f(a+\vec{h}) - f(a)}$ peut se représenter par un accroissement linéaire continu $L.\vec{h}$, avec une erreur qui est infiniment petite devant $\|\vec{h}\|$ lorsque $\vec{h}$ tend vers $\vec{0}$ dans $\vec{E}$. Notons que $\varphi(\vec{0})$ peut être choisie n'importe comment et n'a pas d'intérêt, mais $\varphi(\vec{h})$ est déterminée de manière unique pour $\vec{h} \neq \vec{0}$, par $\varphi(\vec{h}) = \dfrac{\overrightarrow{f(a+\vec{h}) - f(a)} - L.\vec{h}}{\|\vec{h}\|}$. Dire que f a pour dérivée L en a, revient à dire que cette quantité, bien définie pour $\vec{h} \neq \vec{0}$ tel que $a+\vec{h} \in \Omega$, tend vers $\vec{0}$ quand $\vec{h}$ tend vers $\vec{0}$. On conviendra toujours de choisir $\overrightarrow{\varphi(0)} = \vec{0}$, de manière que $\vec{\varphi}$ devienne continue à l'origine, et que la restriction $\vec{h} \neq 0$ puisse être levée.

L'existence et la valeur de l'application dérivée ne dépendent que des topologies de $\vec{E}$ et de $\vec{F}$, et non de leurs normes, à cause du théorème 12 du chapitre II.

Théorème 8 - Si l'application f admet une dérivée au point a cette dérivée L est unique. Dans ce cas, f est continue au point a ; en outre, f a une dérivée en a suivant tout vecteur $\vec{X}$ de $\vec{E}$, et l'application $\vec{X} \to \overrightarrow{D_{\vec{X}} f}(a)$ est une application linéaire continue de $\vec{E}$ dans $\vec{F}$, qui n'est autre que L elle-même, c'est-à-dire :

(III,3;14) $$\overrightarrow{D_{\vec{X}} f}\,(a) = L\cdot\vec{X}\,.$$

Ainsi l'existence et la connaissance de l'application dérivée L en a entrainent l'existence et la connaissance de la dérivée en a suivant tout vecteur $\vec{X}$ de $\vec{E}$.

Démonstration : Si $\vec{h}$ tend vers $\vec{0}$, $L\cdot\vec{h}$ tend vers $\vec{0}$ puisque l'application L est supposée continue, et il en est de même de $\overrightarrow{\varphi(\vec{h})}\|\vec{h}\|$, donc f est bien continue au point a. Soit $\vec{X} \in \vec{E}$. On a, en faisant $\vec{h} = t\vec{X}$ dans (III,3;13), et en remarquant que, Ω étant ouvert, $a + t\vec{X}$ est dans Ω pour $|t|$ assez petit, la formule :

(III,3;15) $$\frac{\overrightarrow{f(a+t\vec{X}) - f(a)}}{t} = L\cdot\vec{X} + \frac{|t|}{t}\varphi(t\vec{X})\|\vec{X}\|\;;$$

cela prouve l'égalité (III,3;14) en faisant tendre $t \neq 0$ vers 0, et par là-même l'unicité de la dérivée, puisque $L\cdot\vec{X}$ est connue pour tout vecteur $\vec{X}$ de $\vec{E}$.

Notation f étant une application de $\Omega \subset E$ dans F, on pourra noter par $f'(a)$ ou $\frac{df}{dx}$ (a) l'application dérivée de f au point a ; on a donc $f'(a) \in \mathcal{L}(\vec{E}\,;\vec{F})$.
Si alors $\vec{X}$ est un vecteur de $\vec{E}$, on pourra noter par $f'(a)\cdot\vec{X}$, ou $\frac{df}{dx}(a)\cdot\vec{X} \in \vec{F}$, la valeur de cette application dérivée sur le vecteur $\vec{X}$.

On a donc la formule :

(III,3;14bis) $$\overrightarrow{D_{\vec{X}} f}\,(a) = f'(a)\cdot\vec{X} \in \vec{F}\,.$$

Remarques 1°/ Soient E et F des espaces affines sur le corps $\mathbb{C}$ des complexes. Ils sont à fortiori affines sur le corps $\mathbb{R}$ des réels. Une appliuation L de $\vec{E}$ dans? , linéaire quand on considère $\vec{E}$ et $\vec{F}$ comme vectoriels sur $\mathbb{C}$, l'est à fortiori quand on les considère comme vectoriels sur $\mathbb{R}$.

Donc une application f de $\Omega \subset E$ dans F , ayant une dérivée $L \in \mathcal{L}(\vec{E};\vec{F})$ en $a \in \Omega$, quand on considère E et F comme affines sur $\mathbb{C}$,a à fortiori L comme dérivée quand on les considère comme affines sur $\mathbb{R}$. L'inverse n'est pas exact ;c'est ce que nous verrons plus en détail dans la théorie des fonctions anaiytiques de variables complexes.

2°/ Il résulte de la définition que la dérivée $f'(a)$ n'est plus un vecteur de $\vec{F}$, mais une application linéaire continue de $\vec{E}$ dans $\vec{F}$; c'est $f'(a).\vec{X}$ qui-est un vecteur de F, pour $\vec{X} \in \vec{E}$. Mais supposons que E soit le corps des scalaires $\mathbb{K}$. Alors nous avons défini un vecteur dérivé $\overrightarrow{f'(a)} \in \vec{F}$ par (III,3;1), et une application dérivée $f'(a) \in \mathcal{L}(\mathbb{K},\vec{F})$ par (III,3;13). La liaison entre les deux notions est bien simple. Si l'un existe, il en est de même de l'autre, et

$$\overrightarrow{f'(a)} = f'(a)\cdot\vec{1}\ , \quad \vec{1} \in \mathbb{K}\ . \qquad (III,3;15^{bis})$$

En effet, si [illegible] $\in \mathcal{L}(\mathbb{K};\vec{F})$ existe, le théorème 8 dit bien que $\overrightarrow{f'(a)} = D_{\vec{1}}f\,(a) \in \vec{F}$ existe et vaut $f'(a).\vec{1}$. Inversement, si le vecteur dérivé $\overrightarrow{f'(a)}$ existe, on a, pour tout $h \in$ K :

$$f(a+h) = f(a) + h\,\overrightarrow{f'(a)} + \vec{\alpha}\,|h|\ ,$$

où $\vec{\alpha}$ tend vers 0 quand-k tend vers 0 . Cela veut bien dire que $f'(a) \in \mathcal{L}(\mathbb{K};\vec{F})$ existe, et c'est l'application $X \rightarrow X\ \overrightarrow{f'(a)}$. On peut donc écrire Indifféremment $f'(a).X$ ou $\overrightarrow{f'(a)}\,X$ pour $X \in \mathbb{K}$.

Le mot dérivée est donc ambigu, puisqu'il peut désigner Indifféremment l'application dérivée $f'(a)$ ou le vecteur dérivé $\overrightarrow{f'(a)} = f'(a).1$. Cette ambiguité n'est guère gênante dans la pratique, D'ailleurs, dans le calcul matriciel, on représente bien un vecteur $\vec{Y}$ de $\mathbb{K}^n$, par une matrice colonne, c'est-à-dire la matrice de l'application linéaire $L : X \longrightarrow \vec{Y}\,X$ de $\mathbb{K}$ dans $\mathbb{K}^n$, et on a bien aussi $\vec{Y} = L\cdot 1$.

Z

si $E = F = \mathbb{K}$, le vecteur dérivé est la dérivée usuelle $f'(a) \in \mathbb{K}$, et l'application dérivée est l'homothétle $x \rightarrow f'(a)\,x$ de $\mathbb{K}$ dans $\mathbb{K}$.

3°/ Si E est de dimension finie, et si on y a choisi un référentiel $a, (\vec{e}_j)_{j\in J}$, alors la dérivée $f'(x) \in \mathcal{L}(\vec{E};\vec{F})$ est reliée aux dérivées partielles $f'(x)\cdot\vec{e}_j = \frac{\partial \vec{f}}{\partial x_j}(x)$ par

$$(\mathrm{III},3;15\,ter)\quad f'(x)\cdot\vec{X} = f'(x)\cdot\Big(\sum_{j\in J} X_j\,\vec{e}_j\Big) = \sum_{j\in J} X_j\, f'(x)\cdot\vec{e}_j = \sum_{j\in J} X_j\,\frac{\partial \vec{f}}{\partial x_j}(x)\,.$$

Si maintenant E et F sont de dimension finie, et si, dans chacun d'eux, on a choisi un référentiel, à savoir $a, (\vec{e}_j)_{j\in J}$, dans E, et $b, (\vec{f}_i)_{i\in I}$, dans F, alors chaque point de E et de F est entiérement déterminé par ses coordonnées, et l'application f de Ω dans F peut être définie par un système de m fonctions de n variables scalaires, à savoir $y_i = F_i\big((x_j)_{j\in J}\big)$, $i\in I$ (si $I=\{1,2,\dots m\}$, et $J=\{1,2,\dots n\}$, alors ce sont des fonctions $y_i = F_i(x_1,x_2,\dots x_n)$, pour $i=1,2,\dots m$). Dans ce cas, l'application dérivée au point x, si elle existe, est définie comme suit; on pose $\vec{X} = \sum_{j\in J} X_j\,\vec{e}_j$, et on appelle $\vec{Y} = \sum_{i\in I} Y_i\,\vec{e}_i$ son image par l'application dérivée: alors :

$$(\mathrm{III},3;16)\quad \begin{cases} \displaystyle\sum_{i\in I} Y_i\,\vec{f}_i = f'(x)\cdot\vec{X} = \sum_{j\in J}\frac{\partial \vec{f}}{\partial x_j}(x)\,X_j = \sum_{i,j}\frac{\partial F_i}{\partial x_j}(x)\,X_j\,\vec{f}_i \\ \text{ou} \\ \displaystyle Y_i = \sum_{j\in J}\frac{\partial F_i}{\partial x_j}(x)\,X_j\,,\ i\in I. \end{cases}$$

Cela montre que la matrice de l'application dérivée $f'(x)$, par rapport aux référentiels considérés, est ce que nous avons appelé la matrice dérivée (III,3;8). Comme toujours pour la matrice d'une application linéaire, les colonnes de la matrice représentent des vecteurs, qui sont les images, par l'application linéaire, des vecteurs de la base de $\vec{E}$, autrement dit les $f'(x)\cdot\vec{e}_j = \frac{\partial \vec{f}}{\partial x_j}(x)$.
L'existence de l'application dérivée $f'(x)$ entraîne celle des $\frac{\partial \vec{f}}{\partial x_j}$ donc des $\frac{\partial F_i}{\partial x_j}$, et de la matrice dérivée, mais la réciproque n'est pas vraie, comme le montre l'exemple (III,3;9). Le déterminant jacobien en x, si E et F ont même dimension, est le déterminant de $f'(x) \in \mathcal{L}(\vec{E};\vec{F})$ par rapport aux référentiels considérés.

Rappelons qu'on peut aussi parler du déterminant d'une application linéaire d'un espace de dimension finie $\vec{E}$ dans lui-même, sans spécifier aucune base (parce que le déterminant, calculé pour une base, est indépendant de cette base). On peut donc aussi parler du déterminant jacobien de f au point x, ou déterminant de $f'(x)$, si f est une application d'un ouvert de E dans E lui-même, sans spécifier aucun referentiel de E.

Notation différentielle

Au lieu de représenter par x, y_0, etc les points de E, F, et par $\vec{X}, \vec{Y}$ etc..., les points de $\vec{E}$, $\vec{F}$, il est souvent commode de noter par x, y, les premiers et par $\vec{dx}, \vec{dy}$, en lettres minuscules, les autres; alors l'expression de la dérivée de l'application f de Ω dans F, s'écrit :

(III,3;17) $$\vec{dx} \to \vec{dy} = f'(x) \cdot \vec{dx} .$$

Si en particulier E est de dimension finie et muni d'un référentiel, la formule s'écrit alors :

(III,3;18) $$\vec{dx} = \sum_{j\in J} dx_j \, \vec{e}_j \longrightarrow \vec{dy} = f'(x)\cdot \vec{dx} = \sum_{j\in J} \frac{\partial \vec{f}}{\partial x_j} dx_j$$

Si f est une fonction sur $\mathbb{K}^2$, donc de deux variables x, y, scalaires, on la note fréquemment z, on appelle $\vec{p}$ et $\vec{q}$ ses dérivées partielles en x, y, et sa dérivée totale s'écrit, en notation différentielle :

(III,3;19) $$(dx, dy) \longrightarrow \vec{dz} = \vec{p}\,dx + \vec{q}\,dy .$$

La dernière formule (III,3;16) s'écrit, en remplaçant X_j, Y_i par dx_j, dy_i :

(III,3;19 A) $$dy_i = \sum_{j\in J} \frac{\partial F_i}{\partial x_j}(x)\, dx_j \quad ; \quad i \in I .$$

Interprétation géométrique de l'application dérivée: variété différentiable et variété linéaire tangente

Soit f une application d'un ouvert Ω de E dans F. Son graphe ou graphique est l'ensemble $\mathscr{C}$ des points $(x, f(x))$, $x \in \Omega$, de $E \times F$: Si f est dérivable en tout point de Ω, on dit que $\mathscr{C}$ est une variété différentiable de $E \times F$, d'équation $y = f(x)$. Par exemple, si $E = \mathbb{R}^2$, $F = \mathbb{R}$, $\mathscr{C}$ est une surface de $\mathbb{R}^3$ d'équation $z = f(x, y)$;

si $E = \mathbb{R}$, $F = \mathbb{R}^2$, $\mathcal{C}$ est une courbe de $\mathbb{R}^3$ d'équations $y = g(x), z = h(x)$ *.

Qu'appellera-t-on **variété** linéaire tangente à $\mathcal{C}$ au point **A=** $(a, f(a))$ **?**

Considérons un ensemble quelconque% d'un espace affine normé **G** . Soit A un point de $\mathcal{C}$. On dit qu'un vecteur? de $\vec{G}$ est tangent en **A** à $\mathcal{C}$, **s'il** existe une suite de points $M_0, M_1, M_2, \dots M_n$, de'% , tendent vers A pourn tendant vers l'infini, et une suite de scalaires $\lambda_0, \lambda_1, \lambda_2, \dots \lambda_n, \dots$, réels $\geqslant 0$, tels que les $\lambda_n \overrightarrow{AM_n}$ tendent vers $\vec{X}$ pour n tendant vers l'infini. Le vecteur $\vec{0}$ est toujours tangent; si $\vec{X}$ est tangent, $\lambda \vec{X}$ l'est aussi pour λ réel $\geqslant 0$. Si $\vec{X} \neq \vec{0}$ est tangent, nécessairement les λ_n **tendent** vers $+\infty$. Comme la notion de limite elle-même, les vecteurs tangents en A à $\mathcal{C}$ ne dépendent que de la topologie et non de la norme de $\vec{G}$. L'ensemble des vecteurs tangents en A à $\mathcal{C}$ est le contingent vectoriel de $\mathcal{C}$ au point A ; l'ensemble des points $A + \vec{X}$, où $\vec{X}$ parcourt le contingent vectoriel, est le **contingent affine de $\mathcal{C}$ en A.**

Theorème **8** A - Soit f une application d'un **ouvert** Ω de **E** dans F ,dérivable au point a de Ω . Le contingent vectoriel (resp. affine) au point $A = (a, f(a) = b)$ de l'ensemble% d'équation $y = f(x)$ dans $E \times F$, est le sous-espace vectoriel de $\vec{E} \times \vec{F}$ d'**équation**:

$$\vec{Y} = f'(a) \cdot \vec{X} \qquad \text{(III,3;19 bis)}$$

(resp. le sous-espace affine de **E × F** d'équation :

$$\overrightarrow{y-b} = 3'(a) \cdot x-a) \quad . \qquad \text{(III,3;19 ter)}$$

* Au lieu de x et y, nous prenons x, y, z parce que $E \times F = \mathbb{R}^3$. Bien entendu, la phrase : "$\mathcal{C}$ a pour équation $y = f(x)$ " ne-veut rien dire; elle est une abréviation évidente de la phrase : "$\mathcal{C}$ est l'ensemble des couples (x, y) de $E \times F$ vérifiant $y = f(x)$"ou"$\mathcal{C} = \{(x, y) ; x \in \Omega, \ y = f(x)\}$"

<u>Démonstration</u> - La deuxième affirmation (correspondant aux resp.) est identique à la première; démontrons la première. soit $x_0, x_1, \ldots, x_n, \ldots$ une suite de points de Ω, tendant vers a, et soit $\lambda_0, \lambda_1, \ldots \lambda_n, \ldots$ une suite de scalaires réels ≥ 0, tels que les $\lambda_n(\overrightarrow{x_n - a})$ tendent vers une limite $\vec{X}$. D'après la définition (III,3;13) de l'application dérivée, on a :

(III,3;19 quarto) $$\overrightarrow{f(x_n) - b} = f'(a) \cdot \overrightarrow{x_n - a} + \vec{\alpha}_n \|\overrightarrow{x_n - a}\|,$$

où $\vec{\alpha}_n$ tend vers $\vec{0}$ pour n tendant vers l'infini.

Alors

(III,3;19 quinto) $$\lambda_n \overrightarrow{f(x_n) - b} = f'(a) \cdot \lambda_n (\overrightarrow{x_n - a}) + \vec{\alpha}_n \lambda_n \|\overrightarrow{x_n - a}\|$$

Au second membre, le premier terme converge pour n infini vers $f'(a) \cdot \vec{X}$ puisque $\lambda_n \overrightarrow{x_n - a}$ converge vers $\vec{X}$ et que $f'(a)$ est supposée continue; le second terme converge vers $\vec{0}$, puisque $\|\lambda_n (\overrightarrow{x_n - a})\|$ converge vers $\|\vec{X}\|$ et $\|\vec{\alpha}_n\|$ vers 0. Donc le premier membre converge vers $\vec{Y} = f'(a) \cdot \vec{X}$. De cela découle bien le théorème. En effet : 1°) Soit $(\vec{X}, \vec{Y})$ un vecteur tangent en A à $\mathscr{C}$. si $M_1, M, \ldots M_n, \ldots$ est une suite de points $(x_n, f(x_n))$ de $\mathscr{C}$ tendant vers A, $\lambda_0, \lambda_1, \ldots \lambda_n, \ldots$ une suite de scalaires ≥ 0, tels que les $\lambda_n \overrightarrow{AM_n}$ aient la limite $(\vec{X}, \vec{Y})$ dans $\vec{E} \times \vec{F}$, cela veut dire que $\lambda_n (\overrightarrow{x_n - a})$ tend vers $\vec{X}$, et que $\lambda_n (\overrightarrow{f(x_n) - b})$ tend vers $\vec{Y}$; nous venons de voir qu'on a bien (III,3;19 bis).

2°) Inversement, soit $(\vec{X}, \vec{Y})$ un vecteur de $\vec{E} \times \vec{F}$ vérifiant (III,3;19 bis). Considérons la suite des x_n de Ω définie par $x_n = a + t_n \vec{X}$, t_n réels > 0, tendant vers 0 pour n infini, et la suite λ_n définie par $\lambda_n = \frac{1}{t_n}$. Alors $\lambda_n (\overrightarrow{x_n - a}) = \vec{X}$. Donc, d'après ce que nous avons vu plus haut, $\lambda_n (\overrightarrow{f(x_n) - b})$ tend vers $f'(a) \cdot \vec{X}$, c'est-à-dire $\vec{Y}$; la suite des $M_n = (x_n, f(x_n))$ appartient bien à $\mathscr{C}$ et tend vers A, et $\lambda_n \overrightarrow{AM_n}$ converge bien vers $(\vec{X}, \vec{Y})$, qui est donc bien un vecteur tangent en A à $\mathscr{C}$.

Le sous-espace vectoriel de $\vec{E} \times \vec{F}$ d'équation (III,3;19 bis) (resp. le sous-espace affine de $E \times F$ d'équation (III,3;19 ter)) s'appelle <u>le sous-espace vectoriel tangent en A à la variété $\mathscr{E}$ (resp. la variété linéaire tangente en A à la variété $\mathscr{E}$)</u>; la fonction affine (application affine de E dans F) : $x \to b + f'(a) \cdot \overrightarrow{x-a}$ s'appelle <u>la fonction affine tangente en a à la fonction f.</u>
<u>Ainsi l'équation de la variété linéaire tangente en $A = (a, f(a) = b)$ à une variété différentiable d'équation $y = f(x)$ dans $E \times F$ s'obtient en remplaçant dans la différentielle $\overrightarrow{dy} = f'(a) \cdot \overrightarrow{dx}$, $\overrightarrow{dx}$ par $\overrightarrow{x-a}$ et $\overrightarrow{dy}$ par $\overrightarrow{y-b}$.</u>

Ainsi, si $E = \mathbb{R}^2$, $F = \mathbb{R}$, le plan tangent au point (a, b, c) à la surface d'équation $z = f(x,y)$ a l'équation

(III,3;19 sexto)
$$z - c = \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b)$$
$$= p(x-a) + q(y-b).$$

Si $E = \mathbb{R}$, $F = \mathbb{R}^2$, la tangente en (a,b,c) à la courbe d'équations $y = g(x)$, $z = h(x)$, a pour équations :

(III,3;19 septimo)
$$y - b = g'(a)(x-a)$$
$$z - c = h'(a)(x-a)$$

Gradient d'une fonction réelle sur un espace euclidien

Soit E un espace affine euclidien de dimension finie sur le corps des réels. Soit f une fonction réelle définie sur un ouvert Ω de E. Si en un point x de E f admet une dérivée $f'(x)$, celle-ci est une application linéaire de $\vec{E}$ dans $\mathbb{R}$, c'est-à-dire une forme linéaire ou un élément de $\vec{E}'$. D'après ce que nous avons vu au théorème 2 quarto, il existe un vecteur $\vec{Y}$ et un seul de $\vec{E}$, tel que l'on ait

(III,3;20)
$$f'(x) \cdot \vec{X} = (\vec{X} \mid \vec{Y}), \text{ pour tout } \vec{X} \text{ de } \vec{E}.$$

Ce vecteur s'appelle le gradient de f en x, et on le note $\overrightarrow{grad} f(x)$. En remplaçant $\vec{X}$ par $\overrightarrow{dx}$, et $f'(x).\overrightarrow{dx}$ par $df \in \mathbb{R}$, suivant la notation différentielle, on a donc :

$$(\mathrm{III},3;21) \quad \begin{cases} \overrightarrow{D_{\vec{X}}} f(x) = (\overrightarrow{grad} f(x) | \vec{X}) \\ df = (\overrightarrow{grad} f(x) | \overrightarrow{dx}) \end{cases}$$

soit $(\vec{e}_i)_{i \in I}$ une base orthonormale de $\vec{E}$. Alors la formule (III,3;21) donne, compte tenu de la formule du produit scalaire par rapport à une base orthonormale,

$$(\mathrm{III},3;22) \quad df = \sum_{i \in I} Y_i \, dx_i \;,$$

si les Y_i sont les composantes du gradient; mais on a aussi (III,3;18), donc

$$(\mathrm{III},3;23) \quad Y_i = (\overrightarrow{grad} f(x))_i = \frac{\partial f}{\partial x_i} .$$

Les composantes du gradient de f par rapport à une base orthonormale, sont les dérivées partielles $\frac{\partial f}{\partial x_i}$.

Ceci subsiste dans le cas d'un espace euclidien généralisé; mais alors on doit utiliser la formule (III,1;25), donc (III,3;22) est à remplacer par

$$(\mathrm{III},3;24) \quad df = \sum_{i \in I} \varepsilon_i \, Y_i \, dx_i \;,$$

donc les composantes du gradient sont $\varepsilon_i \frac{\partial f}{\partial x_i}(x)$. En particulier, en relativité restreinte, ce sont

$$\frac{\partial f}{\partial x_1}, \quad \frac{\partial f}{\partial x_2}, \quad \frac{\partial f}{\partial x_3}, \quad - \frac{\partial f}{\partial x_0} = - \frac{1}{c} \frac{\partial f}{\partial t} .$$

La fonction $x \to \overrightarrow{grad} f(x)$, si le gradient est défini sur Ω tout entier, est ce qu'on appelle un champ de vecteurs sur Ω, ou application de Ω dans $\vec{E}$. On note $\overrightarrow{grad}$ cette application.

Théorème 8 bis (dérivée d'une constante, d'une fonction affine)

Une application constante est dérivable et de dérivée nulle.

Une application affine continue f de E dans F est dérivable en tout point a de E, et sa dérivée est l'application linéaire associée : $f'(a) = \vec{f} \in \mathcal{L}(\vec{E};\vec{F})$.

La fonction dérivée (que nous définirons plus loin, page 44) est donc une application constante de E dans $\mathcal{L}(\vec{E};\vec{F})$: $x \longrightarrow f'(x) = \vec{f}$. Ce résultat généralise le fait que la dérivée de la fonction affine $y = \alpha x + \beta$ est la constante α.

Evident.

Théorème 8 ter (La dérivation est une opération linéaire)*

Si f est une application de Ω dans F, et $\vec{g}$ une application de Ω dans $\vec{F}$, et si f et $\vec{g}$ ont des ***dérivées*** en $a \in \Omega$, alors la fonction $f+\vec{g} : x \longrightarrow f(x) + \overrightarrow{g(x)}$, a une dérivée en a, égale à la **somme** des dérivées :

(III,3;25) $$(f+\vec{g})'(a) = f'(a) + g'(a).$$

Si λ est une constante scalaire, la fonction $\lambda\vec{g} : x \rightarrow \lambda\overrightarrow{g(x)}$ a une dérivée en a, qui est $\lambda\, g'(a)$.

Evident.

Cas où F est un produit d'espaces affines

Supposons que f soit une application de $\Omega \subset$ E dans un produit $F_1 \times F_2 \times \cdots \times F_m$ d'espaces affines normés. Elle est définie par des applications f_i de E dans F_i , $i = 1,2,\ldots m$ (voir chapitre 1, page 7). On sait que f est Continue si et seulement si les f_i sont continues (théorème 17 du chapitre II et suite). On voit aussitôt que f est affine si et seulement si les f_i sont affines.

Théorème 8 quarto - Soient E et F_i , $i = 1,2,\ldots m$, des espaces affines normés. Soit f une application d'un ouvert Ω de E dans $F = F_1 \times F_2 \times \ldots \times F_m$, définie par des applications f_i de Ω dans les F_i . Pour que f soit dérivable au point a de Ω , il faut et il suffit que les f_i soient dérivables en a , et alors $f'(a)$ est l'application linéaire continue de $\vec{E}$ dans $\vec{F} = \vec{F}_1 \times \vec{F}_2 \times \ldots \times \vec{F}_m$ définie par les f'_i (a) , applications linéaires continues de $\vec{E}$ dans les $\vec{F}_i$.

Autrement dit, les **composantes** de la dérivée sont les dérivées des composantes, et on a :

(III,3;26) $$f'(a) = \left(f'_1(a),\ f'_2(a), \ldots f'_m(a)\right) ;\ \overrightarrow{df} = \left(\overrightarrow{df_1},\ \overrightarrow{df_2}, \ldots, \overrightarrow{df_m}\right).$$

Démonstration - Supposons les f_i dérivables en a.

Pour un accroissement $\overrightarrow{dx}$ de x, on a des accroissements $\overrightarrow{\Delta y_i}$ de $y_i = f_i(x)$, avec

$$\overrightarrow{\Delta y_i} = f'_i(a) \cdot \overrightarrow{dx} + \vec{\alpha}_i \, \| \overrightarrow{dx} \| , \tag{III,4;27}$$

où $\vec{\alpha}_i$ tend vers $\vec{0}$ avec $\overrightarrow{dx}$ Alors, d'après la définition de la structure affine Produit :

$$(\text{III } 4;28) \quad \left\{ \begin{aligned} \overrightarrow{\Delta(y_1, y_2, \dots, y_m)} &= \left(\overrightarrow{\Delta y_1}, \overrightarrow{\Delta y_2}, \dots, \overrightarrow{\Delta y_m}\right) \\ &= \left(f'_1(a) \cdot \overrightarrow{dx} , \quad f'_2(a) \cdot \overrightarrow{dx} \quad , \dots , f'_m(a) \cdot \overrightarrow{dx}\right) \\ &+ \left(\vec{\alpha}_1 \, \| \overrightarrow{dx} \| \quad , \quad \vec{\alpha}_2 \, \| \overrightarrow{dx} \| \quad , \dots , \quad \vec{\alpha}_m \, \| \overrightarrow{dx} \| \right) \\ &= \left(f'_1(a) , \; f'_2(a) , \dots , \; f'_m(a)\right) \cdot \overrightarrow{dx} \\ &+ \left(\vec{\alpha}_1 , \; \vec{\alpha}_2 , \dots , \vec{\alpha}_m \right) \| \overrightarrow{dx} \| , \end{aligned} \right.$$

où $(\vec{\alpha}_1 , \vec{\alpha}_2 , \dots , \vec{\alpha}_m)$ tend vers $\vec{0}$ dans $\vec{F}_1 \times \vec{F}_2 \times \dots \times \vec{F}_m$ **quand** $\overrightarrow{dx}$ tend vers $\vec{0}$, **d'après** la définition de la topologie produit. Cela prouve bien que f est dérivable en a , et que $f'(a) = \left(f'_1(a) , f'_2(a) , \dots , f'_m(a)\right)$.

Démonstration identique en sens inverse pour la réciproque.

Cas <u>où E est un produit d'espaces affines</u>. <u>Applications dérivées partielles</u>.

Si maintenant E est un produit $E_1 \times E_2$ alors l'application f de $\Omega \subset E_1 \times E_2$ **dans** F devient'une fonction de deux variables. Nous l'écrirons sous la forme $y = f(x_1, x_2)$. Pour x_1 fixé en a_1, on peut alors considérer l'application partielle $f_{a_1} : x_2 \longrightarrow f(x_1, x_2)$, et chercher si celle-ci admet une application dérivée au pointa, . S'il en est bien ainsi, cette application dérivée est une application de E_2 dans $\vec{F}$, on l'appelle <u>application **dérivée partielle ou différentielle** partielle</u> en x_2 **de f au point (a_1 , a_2).** On la notera par **l'une quelcon**que des notations : $\partial_2 f \; (a_1, a_2)$ ou $\dfrac{\partial f}{\partial x_2} (a_1 , a_2)$.

On pourra de même considérer l'application **partielle** f_{a_2} **et son** application **dérivée** partielle correspondante $\frac{\partial f}{\partial x_1}(a_1 \cdot, a_2)$ **, au** même point (a_1, a_2) . De la même manière que précédemment (exemple (III,3;9)) le fait que f possède des applications dérivées partielles en un point, n'entraîne pas nécessairement qu'elle possède une application dérivée totale en ce point (ni même qu'elle y soit continue) * C'est la réciproque qui est vraie. Plus précisément :

Théorème 9 - Si E est un produit de 2 espaces affines, $E = E_1 \times E_2$, et si f possède une application dérivée $f'(a_1, a_2)$ en un point $a = (a_1, a_2)$ de Ω, partie ouverte de $E_1 \times E_2$, alors elle possède en ce point des applications dérivées Partielles et son application dérivée totale est donnée par la formule :

(III,3;29) $$\frac{df}{dx}(a_1, a_2) \cdot (\vec{X}_1, \vec{X}_2) = \frac{\partial f}{\partial x_1}(a_1, a_2) \cdot \vec{X}_1 + \frac{\partial f}{\partial x_2}(a_1, a_2) \cdot \vec{X}_2 .$$

Démonstration Nous avons vu au théorème **51** du chapitre II **que** $\frac{df}{dx}(a_1, a_2)$, **application** linéaire continue de $\vec{E}_1 \times \vec{E}_2$ dans $\vec{F}$, définit des applications linéaires continues L_1 **et** L_2 de $\vec{E}_1$ et $\vec{E}_2$ dans $\vec{F}$, et qu'elle est donnée par la formule :

(III,3;30) $$\frac{df}{dx}(a_1, a_2) \cdot (\vec{X}_1, \vec{X}_2) = L_1 \cdot \vec{X}_1 + L_2 \vec{X}_2 .$$

Si alors on donne à la variable un **accroissement** $\vec{h} = (\vec{0}, \vec{h}_2)$, d'après la définition même de l'application dérivée, l'accroissement **de** f est donné par la formule :

(III,3;31) $$\overrightarrow{f(a_1, a_2 + \vec{h}_2) - f(a_1, a_2)} = L_2 \cdot \vec{h}_2 + \overrightarrow{\varphi(\vec{0}, \vec{h}_2)} \|\vec{h}_2\|$$

où $\overrightarrow{\varphi(0, \vec{h}_2)}$ tend vers $\vec{0}$ lorsque $\vec{h}_2$ tend vers $\vec{0}$ **; mais** cela signifie exactement que L_2 est l'application dérivée partielle de f par rapport à la deuxième variable, au

* Mais nous verrons au théorème 15 que l'existence de fonctions dérivées partielles continues **entraîne** l'existence d'une fonction dérivée totale, qui est alors aussi continue.

point (a_1, a_2).

Nous laissons au lecteur le soin d'étendre ceci au produit $E_1 \times E_2 \times \ldots \times E_n$ de n espaces affines. En notation différentielle, on aura

(III,3;32) $$\overrightarrow{df} = f'(x) \cdot \overrightarrow{dx} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) \cdot \overrightarrow{dx_j} ,$$

avec cette fois $\frac{\partial f}{\partial x_j}(x) \in \mathcal{L}(\vec{E}_j ; \vec{F})$, $\overrightarrow{dx_j} \in \vec{E}_j$,

$$\overrightarrow{dx} = (\overrightarrow{dx_1}, \overrightarrow{dx_2}, \ldots, \overrightarrow{dx_n}) \in \vec{E} = \vec{E}_1 \times \vec{E}_2 \times \cdots \times \vec{E}_n .$$

Si les facteurs $\vec{E}_j$ sont identiques au corps des scalaires K, on retrouve la formule (III,3;18) (avec l'identification, indiquée à la remarque 2° page 199, entre vecteur dérive et application dérivée).

Dérivée d'une application bilinéaire continue

La formule classique $d(xy) = y\,dx + x\,dy$, pour la différentielle d'un produit, se généralise comme suit.

Théorème 9 bis - Soit B une application bilinéaire continue de $\vec{E}_1 \times \vec{E}_2$ dans $\vec{F}$. Alors B est dérivable en tout point $(\vec{a}_1, \vec{a}_2)$ de $\vec{E}_1 \times \vec{E}_2$. Sa dérivée partielle $\frac{\partial B}{\partial x_1}(\vec{a}_1, \vec{a}_2)$ est l'application partielle $B_{\vec{a}_2} \in \mathcal{L}(\vec{E}_1 ; \vec{F})$, définie par :

(III,3;33) $$B_{\vec{a}_2}(\vec{X}_1) = B(\vec{X}_1, \vec{a}_2) \in \vec{F} \qquad \text{(voir page 117)}.$$

De même sa dérivée partielle $\frac{\partial B}{\partial x_2}(\vec{a}_1, \vec{a}_2)$ est l'application partielle $B_{\vec{a}_1} \in \mathcal{L}(\vec{E}_2 ; \vec{F})$ définie par

(III,3;34) $$B_{\vec{a}_1}(\vec{X}_2) = B(\vec{a}_1, \vec{X}_2) \in \vec{F} .$$

Sa dérivée totale est donc

(III,3;35) $$B'(a_1, a_2) \cdot (\vec{X}_1, \vec{X}_2) = B(\vec{X}_1, \vec{a}_2) + B(\vec{a}_1, \vec{X}_2)$$

ou, en notation différentielle :

(III,3;36) $$dB = B(\overrightarrow{dx_1}, \vec{x}_2) + B(\vec{x}_1, \overrightarrow{dx_2}) .$$

<u>Démonstration.</u> Les **résultats** relatifs aux dérivées partielles sont évidents. Car l'application partielle $B_{\vec{a}_2}$:
$\vec{x}_1 \longrightarrow B(\vec{x}_1, \vec{a}_2)$ est linéaire continue, donc sa dérivée en a, est cette application elle-même (théorème 8 bis), c'est-à-dire $\vec{X}_1 \longrightarrow B(\vec{X}_1, \vec{a}_2)$. Si alors B est dérivable, la formule (III,3;35) est le résultat de (III,3; 30). Mais la dérivabilité partielle de B n'entrafne pas sa dérlvabilité totale et nous sommes obligés de ddmontrer cette dernière * .

Donnons à $\vec{x}_1$, $\vec{x}_2$, des accroissements $\overrightarrow{dx_1}, \overrightarrow{dx_2}$ et soit $\overrightarrow{\Delta B}$ l'accroissement de B . On a

$$(\text{III},3;37) \qquad \overrightarrow{\Delta B} = B(\vec{a}_1 + \overrightarrow{dx_1}, \vec{a}_2 + \overrightarrow{dx_2}) - B(\vec{a}_1, \vec{a}_2)$$

$$= B(\vec{a}_1 + \overrightarrow{dx_1}, \overrightarrow{dx_2}) + B(\overrightarrow{dx_1}, \vec{a}_2)$$

Alors on a

$$(\text{III},3;38) \qquad \overrightarrow{\Delta B} - B(\overrightarrow{dx_1}, \vec{a}_2) - B(\vec{a}_1, \overrightarrow{dx_2}) = B(\overrightarrow{dx_1}, \overrightarrow{dx_2})$$

Le 2ème membre est majoré en norme par

$$(\text{III},3;39) \qquad \|B\| \, \|\overrightarrow{dx_1}\| \, \|\overrightarrow{dx_2}\| \leq \|B\| \left(\|dx_1\| + \|dx_2\|\right)$$

quantité infiniment petite devant $\|\overrightarrow{dx_1}\| + \|\overrightarrow{dx_2}\|$ quand $(\overrightarrow{dx_1}, \overrightarrow{dx_2})$ tend vers $\vec{O}$ dans $\vec{E}_1 \times \vec{E}_2$; si donc on prend sur $\vec{E}_1 \times \vec{E}_2$ la norme $\|\;\|_1 + \|\;\|_2$, on voit que, d'après la définition même de l'application dérivée (III,3; 13), la dérivée de B existe en $(\vec{a}_1, \vec{a}_2)$ et qu'elle est donnée par (III,3:35), ce qui démontre le théorème.

Nous laissons au lecteur le soin d'étendre à un produit dem, espaces vectoriels normés. On a cette fois.

$$(\text{III},3;40) \quad \overrightarrow{dB} = B(\overrightarrow{dx_1}, \vec{x}_2, \ldots, \vec{x}_n) + B(\vec{x}_1, \overrightarrow{dx_2}, \ldots, \vec{x}_n) + \ldots + B(\vec{x}_1, \vec{x}_2, \ldots, \overrightarrow{dx_n})$$

* Le théorème 15 nous permettrait de nous en dispenser, car les fonctions dérivées partielles sont continues.

Remarque . L'application B' qui, à $\vec{x} = (\vec{x}_1, \vec{x}_2) \in \vec{E}_1 \times \vec{E}_2$, fait correspondre la dérivée $B'(\vec{x}) = B'(\vec{x}_1, \vec{x}_2) \in \mathcal{L}(\vec{E}_1 \times \vec{E}_2 ; \vec{F})$ définie par (III,3;35), est linéaire et continue de $\vec{E}_1 \times \vec{E}_2$ dans $\mathcal{L}(\vec{E}_1 \times \vec{E}_2 ; \vec{F})$; elle est en effet trivialement linéaire, et on a la majoration :

$$(\text{III},3;40\,bis) \qquad \| B'(\vec{x}).\vec{X} \| = \| B'(\vec{x}_1, \vec{x}_2).(\vec{X}_1, \vec{X}_2) \| \leq \| B.(\vec{X}_1, \vec{x}_2) \| + \| B(\vec{x}_1, \vec{X}_2) \|$$

$$\leq 2 \| B \| \; \| \vec{x} \| \; \| \vec{X} \| \qquad , \text{ donc}$$

$$(\text{III},3;40\,ter) \qquad \| B'(\vec{x}) \| = \sup_{\|\vec{X}\| \leq 1} \| B'(\vec{x}).\vec{X} \| \leq 2 \|B\| \, \| \vec{x} \| ,$$

ce qui prouve bien que B' est continue, $B' \in \mathcal{L}\left(\vec{E}_1 \times \vec{E}_2 ; \mathcal{L}(\vec{E}_1 \times \vec{E}_2 ; \vec{F})\right)$ et l'on a :

$$(\text{III},3;40\,quarto) \qquad \| B' \| \leq 2 \| B \| .$$

Fonctions dérivables, fonctions continuement dérivables

Si f a une dérivée $f'(x)$ en tout point x de Ω , elle est dite dérivable dans Ω . L'application $x \longrightarrow f'(x)$ est sa fonction dérivée; elle se note f' ou $\frac{df}{dx}$ ou Df , c'est une application de Ω dans $\mathcal{L}(\vec{E} ; \vec{F})$. La fonction! est dite continuement dérivable ou de classe C^1 si l'application dérivée f' de Ω dans $\mathcal{L}(\vec{E};\vec{F})$ est continué.

Théorème 10 - Si f est une application continuement dérivable de Ω dans F , alors l'application $(x, \vec{X}) \longrightarrow f'(x).\vec{X}$ de $\Omega \times \vec{E}$ dans $\vec{F}$ est continue. Réciproquement, si cette application est continue, et si E est de dimension finie, alors f est continuement dérivable.

Démonstration Supposons$' continue de Ω dans $\mathcal{L}(\vec{E};\vec{F})$ • soit $x_0, x_1, x_2, \dots, x_n, \dots$ une suite convergeant vers x dans Ω et $\vec{X}_0, \vec{X}_1, \vec{X}_2, \dots, \vec{X}_n, \dots$ une suite convergeant vers $\vec{X}$ dans $\vec{E}$; alors

la suite des $f'(x_n)$ converge vers $f'(x)$ dans $\mathcal{L}(\vec{E};\vec{F})$, à cause de la continuité de f' , puis la suite des $(f'(x_n), \vec{X}_n)$ converge vers $(f'(x), \vec{X})$ dans $\mathcal{L}(\vec{E};\vec{F}) \times \vec{E}$ d'après le théorème 17 du chapitre II; alors, comme l'application canonique $(u, \vec{X}) \to u.\vec{X}$ de $\mathcal{L}(\vec{E};\vec{F}) \times \vec{E}$ dans? est continue (théorème 54 du

du chapitre II), la suite des $f'(x_n)\cdot\vec{X}_n$ converge vers $f'(x)\cdot\vec{X}$ dans $\vec{F}$, ce qui démontre bien que l'application $(x,\vec{X})\to f'(x)\cdot\vec{X}$ de $\Omega\times\vec{E}$ dans $\vec{F}$ est continue.
Réciproquement , supposons cette application continue. Choisissons une base $(\vec{e}_j)_{j\in J}$ de E . L hypothèse entraine à fortiori la continuité de chaque dérivée partielle $x\longrightarrow f'(x)\cdot\vec{e}_j = \frac{\partial f}{\partial x_j}(x)$. Or nous verrons plus tard (théorème **15)** que, même sans supposer f dérivable, l'existence et la continuité des dérivées partielles $\frac{\partial f}{\partial x_j}$ suffisent à entrainer l'existence et la continuité de la dérivée totale f' ce qui démontre la réciproque.(On s'assurera aisément que les théorèmes qui suivent, jusqu'à 15, n'utilisent pas le théorème 10) ·

Exemples de fonctions continuement dérivables

Une fonction affine est continuement dérivable, puisque sa dérivée est une constante (théorème 8 bis). Une fonction bilinéaire continue est continuement dérivable, puisque sa dérivée est une fonction linéaire continue (voir la remarque qui suit le théorème 9 bis).

Une application de Ω dans un produit $F_1\times F_2\times\cdots\times F_m$ est continuement dérivable si et seulement si chacune de ses composantes est continuement dérivable (théorème 8 quarto)

Nous verrons dans la suite beaucoup d'autres exemples (théorème **15).**

Espaces de fonctions dérivables

Nous avons vu au chapitre II, page **137,** que, si E est un ensemble quelconque et $\vec{F}$ un espace vectoriel, l'ensemble $\vec{F}^E$ des applications de **E** dans? était un espace vectoriel. Si,au lieu de cela, F est un espace affine, alors F^E est un espace affine, d'espace vectoriel associé $\vec{F}^E$; si f et g sont deux applications de **E** dans F , on définit en effet immédiatement l'application $\overrightarrow{g-f}$ de E dans $\vec{F}$, par $\overrightarrow{g-f}(x) = \overrightarrow{g(x)-f(x)}$ *
Si **F** est affine normé, l'espace $(F^E)_b$ des applications bornées de E dans **F** est affine normé, d'espace vectoriel associe $(\vec{F}^E)_b$, il est complet si F est complet (théorime 64 du chapitre II): de même pour $(F^E)_{cb}$ et $(\vec{F}^E)_{cb}$, **si E** est un espace topologique.

* Si E est un ensemble à deux éléments $\{1,2\}$, on retrouve le fait que $F^2 = F\times F$ est un espace affine, d'espace vectoriel associé $\vec{F}^2=\vec{F}\times\vec{F}$

Soient maintenant E , F, des espaces affines normés, Ω un ouvert de E . Nous conviendrons d'appeler $(F^{\Omega})_{\mathcal{b};1}$ (resp. $(F^{\Omega})_{c\mathcal{b};1}$) l'espace des fonctions sur Ω à valeurs dans F , dérivables, bornées ainsi que leur fonction dérivée (resp. continuement dérivables, bornées ainsi que leur fonction dérivée) * . C'est un espace affine, d'espace vectoriel associé $(\vec{F}^{\Omega})_{\mathcal{b};1}$ (resp. $(\vec{F}^{\Omega})_{c\mathcal{b};1}$) .

On en fait un espace affine normé, en mettant sur son espace vectoriel associé la norme

$$(\text{III},3;41) \quad |||\vec{f}|||_1 = \operatorname*{Max}_{x\in\Omega}\left(\|\vec{f}(x)\|,\|f'(x)\|\right) = \operatorname{Max}\left(|||\vec{f}|||,|||\vec{f}'|||\right).$$

(où $\|\vec{f}(x)\|$ est la norme dans F, $\|f'(x)\|$ la norme dans $\mathcal{L}(\vec{E};\vec{F})$)

Nous verrons plus tard (théorème 113 du Chapitre IV) que ces espaces sont complets si F est complet. Dire qu'une suite de fonctions dérivables f_n converge, pour n infini, vers une fonction dérivable f , au sens de $(F^{\Omega})_{\mathcal{b};1}$, c'est dire que les f_n convergent uniformément vers f , et les f'_n uniformément vers f' . C'est donc une convergence plus forte que la convergence dans $(F^{\Omega})_{\mathcal{b}}$. Si E est le corps des réels $\mathbb{R}$, F le corps $\mathbb{R}$ ou le corps $\mathbb{C}$, Ω un intervalle compact $[a,b]$ (**) (resp. la droite réelle $\mathbb{R}$ tout entière), l'espace $(F^{\Omega})_{c\mathcal{b};1}$ est l'espace vectoriel, très important dans les applications, des fonctions continuement dérivables, bornées ainsi que leur dérivée, sur $[a,b]$ (resp. sur $\mathbb{R}$), à valeurs réelles ou complexes.

<u>Théorème 10 bis</u> - <u>L'application qui, à chaque fonction f fait correspondre sa fonction dérivée f', est linéaire et continue de norme ≤ 1, de $(\vec{F}^{\Omega})_{\mathcal{b};1}$ (resp. $(\vec{F}^{\Omega})_{c\mathcal{b};1}$) dans $\left((\mathcal{L}(\vec{E};\vec{F}))^{\Omega}\right)_{\mathcal{b}}$ (resp. $\left((\mathcal{L}(\vec{E};\vec{F}))^{\Omega}\right)_{c\mathcal{b}}$).</u>

Evident (rappelons que, si f est une fonction sur Ω à valeurs dans F , f' est une fonction sur Ω à valeurs dans $\mathcal{L}(\vec{E};\vec{F})$; d'autre part $|||f'||| \leq |||f|||_1$).

* La raison d'être de cet indice 1 apparaîtra page 251 : il s'agit d'espaces de fonctions 1 fois dérivables.

(**) $[a,b]$ n'est pas ouvert. Voir page 184, début du § 2.

§ 4 THÉORÈME DES FONCTIONS COMPOSÉES

Théorême 11 - Soient E, F, G, des espaces affines normés, Ω un ouvert de E, Ω' un ouvert de F ; soit f une application de Ω dans Ω' et g une application de Ω' dans G. Si l'application? admet une dérivée $f'(a) \in \mathcal{L}(\vec{E};\vec{F})$ en un point a de Ω, et si l'application g admet une dérivée $g'(b) \in \mathcal{L}(\vec{F};\vec{G})$ au point $b = f(a)$ de Ω', alors l'application composée, & $= g \circ f$, admet une application dérivée au point a, et celle-ci est composée des applications dérivées :

$$h'(a) = g'(b) \circ f'(a) = g'(f(a)) \circ f'(a). \tag{III,4;1}$$

Démonstration - Avant même de donner la démonstration remarquons que, si $E = F = G = \mathbb{K}$, les applications dérivées sont des multiplications par des scalaires, qui sont les dérivées usuelles, et alors cette formule n'est autre que la formule de la dérivée des fonctions de fonctions, telle qu'elle est écrite habituellement : $(g(f(x)))' = g'(f(x))f'(x)$; précisément le mérite des notations générales que nous avons adoptées est de donner pour des espaces affines de dimension quelconque, finie ou infinie, le même formalisme que pour une fonction réelle d'une variable réelle. Choisissons un accroissement $\vec{dx}$ de manière que le point $a+\vec{dx}$ appartienne à Ω. Le8 accroissements correspondants de y et de z seront notés par $\vec{\Delta y}$ et $\vec{\Delta z}$. On a alors les formules :

$$\vec{\Delta y} = \overrightarrow{f(a+\vec{dx}) - f(a)} = f'(a)\cdot\vec{dx} + \vec{\alpha}\,\|\vec{dx}\|, \tag{III,4;2}$$

où $\|\vec{\alpha}\|$ tend vers 0 lorsque $\vec{dx}$ tend vers $\vec{0}$, et

$$\vec{\Delta z} = \overrightarrow{g(b+\vec{\Delta y}) - g(b)} = g'(b)\cdot\vec{\Delta y} + \vec{\beta}\,\|\vec{\Delta y}\| \tag{III,4;3}$$

où $\|\vec{\beta}\|$ tend vers 0 lorsque $\vec{\Delta y}$ tend vers $\vec{0}$. On en déduit :

$$\vec{\Delta z} = g'(b)\cdot f'(a)\cdot\vec{dx} + g'(b)\cdot\vec{\alpha}\,\|\vec{dx}\| + \vec{\beta}\,\|\vec{\Delta y}\|. \tag{III,4;4}$$

Alors l'expression :

$$\vec{\Delta z} - g'(b)\cdot f'(a)\cdot\vec{dx} = g'(b)\cdot\vec{\alpha}\,\|\vec{dx}\| + \vec{\beta}\,\|\vec{\Delta y}\|, \tag{III,4;5}$$

est majorée par

$$\begin{aligned} &\|g'(b)\|\,\|\vec{\alpha}\|\,\|\vec{dx}\| + \|\vec{\beta}\|\left(\|f'(a)\|\,\|\vec{dx}\| + \|\vec{\alpha}\|\,\|\vec{dx}\|\right) \\ &= \|\vec{dx}\|\left(\|g'(b)\|\,\|\vec{\alpha}\| + \|f'(a)\|\,\|\vec{\beta}\| + \|\vec{\alpha}\|\,\|\vec{\beta}\|\right). \end{aligned} \tag{III,4;6}$$

Lorsque $\overrightarrow{dx}$ tend vers $\vec{0}$, il en est de même de $\overrightarrow{\Delta y}$ parce que f, étant dérivable en a, est continue en a, donc $\vec{\alpha}$ et $\vec{\beta}$ tendent vers $\vec{0}$ et la dernière parenthèse tend vers 0. Comme $g'(b) \circ f'(a)$ est continue (composée de deux applications continues) cela prouve bien que h est dérivable en a, et de dérivée $g'(b) \circ f'(a)$.

En notation différentielle, le résultat précédent s'exprime de la manière suivante. La différentielle de f est donnée par l'expression :

(III,4;7) $$\overrightarrow{dy} = f'(x) \cdot \overrightarrow{dx} ;$$

La différentielle de g par l'expression :

(III,4;8) $$\overrightarrow{dz} = g'(y) \cdot \overrightarrow{dy} ;$$

alors la différentielle de la fonction composée s'obtient en écrivant la différentielle de g, et en y remplaçant? par $f(x)$ et dy par la différentielle $f'(x) \cdot \overrightarrow{dx}$, ce qui donne immédiatement :

(III,4;9) $$\overrightarrow{dz} = h'(x) \cdot \overrightarrow{dx} = g'(f(x)) \cdot (f'(x) \cdot \overrightarrow{dx}) .$$

Corollaire 1 (Permutabilité de la dérivation et d'une application linéaire continue). Si f est une application de $\Omega \subset E$ dans F, ayant une dérivée $f'(a)$ en $a \in \Omega$, et si g application affine continue de F dans G, l'application composée $g \circ f$ de Ω dans G a une dérivée en a, donnée par :

(III,4;10) $$(g \circ f)'(a) = \vec{g} \circ f'(a) .$$

Cela résulte des théorèmes Il et 8 bis (et c'est d'ailleurs évident directement)

Supposons en particulier que E soit un espace affine sur le corps des réels, $\vec{F}$ un espace vectoriel sur le corps des complexes. Dans toute la théorie, on doit les considérer tous deux comme espaces affines sur le corps des réels. Néanmoins la multiplication par $\lambda \in \mathbb{C}$ conserve un sens dans $\vec{F}$, et c'est toujours une application linéaire continue de $\vec{F}$ dans lui-même. Alors, si f est une application dérivable de Ω dans $\vec{F}$, il en est de même de $\lambda \vec{f}$, et l'on aura $(\lambda \vec{f})' = \lambda (\vec{f}')$. Par exemple, du fait que, sur $\mathbb{R}$, la dérivée de $\sin x$ est $\cos x$, la dérivée de $i \sin x$ est $i \cos x$.

<u>Corollaire 2 - Si E, F, G sont des espaces affines de dimension finie, dans lesquels on a choisi des référentiels, alors la matrice dérivée de l'application $h = g \circ f$ au point a est le produit de la matrice dérivée de l'application g au point $b = f(a)$, et de la matrice dérivée de l'application f au point a.</u>

Cela résulte immédiatement de ce que la matrice de la composée de 2 applications linéaires est le produit des matrices.

Nous écrirons cela sous la forme suivante. Un point dans E (resp. F, resp. G) sera défini par ses coordonnées $(x_k)_{k \in K}$, $(y_j)_{j \in J}$, $(z_i)_{i \in I}$. Pour simplifier, prenons $I = \{1, 2, \ldots \ell\}$, $J = \{1, 2, \ldots m\}$, $K = \{1, 2, \ldots n\}$. Alors f est définie par m fonctions de n variables, $y_j = F_j(x_1, x_2, \ldots x_n)$, pour $j = 1, 2, \ldots m$; g est définie par ℓ fonctions de m variables, $z_i = G_i(y_1, y_2, \ldots y_m)$, pour $i = 1, 2, \ldots \ell$; et $h = g \circ f$ est définie par ℓ fonctions de n variables, $z_i = H_i(x_1, x_2, \ldots x_n)$, pour $i = 1, 2, \ldots \ell$. On a d'ailleurs

$$H_i(x_1, x_2, \ldots x_n) = G_i\left(F_1(x_1, x_2, \ldots x_n), F_2(x_1, x_2, \ldots x_n), \ldots, F_m(x_1, x_2, \ldots x_n)\right) \tag{III,4;11}$$

On a alors

$$\begin{pmatrix} \frac{\partial H_1}{\partial x_1}(a) & \frac{\partial H_1}{\partial x_2}(a) \ldots & \frac{\partial H_1}{\partial x_n}(a) \\ \frac{\partial H_2}{\partial x_1}(a) & \frac{\partial H_2}{\partial x_2}(a) \ldots & \frac{\partial H_2}{\partial x_n}(a) \\ \cdots & \cdots & \cdots \\ \frac{\partial H_\ell}{\partial x_1}(a) & \frac{\partial H_\ell}{\partial x_2}(a) \ldots & \frac{\partial H_\ell}{\partial x_n}(a) \end{pmatrix} \tag{III,4;12}$$

$$= \begin{pmatrix} \frac{\partial G_1}{\partial y_1}(b) & \frac{\partial G_1}{\partial y_2}(b) \ldots & \frac{\partial G_1}{\partial y_m}(b) \\ \frac{\partial G_2}{\partial y_1}(b) & \frac{\partial G_2}{\partial y_2}(b) \ldots & \frac{\partial G_2}{\partial y_m}(b) \\ \cdots & \cdots & \cdots \\ \frac{\partial G_\ell}{\partial y_1}(b) & \frac{\partial G_\ell}{\partial y_2}(b) \ldots & \frac{\partial G_\ell}{\partial y_m}(b) \end{pmatrix} \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(a) & \frac{\partial F_1}{\partial x_2}(a) & \ldots & \frac{\partial F_1}{\partial x_n}(a) \\ \frac{\partial F_2}{\partial x_1}(a) & \frac{\partial F_2}{\partial x_2}(a) & \ldots & \frac{\partial F_2}{\partial x_n}(a) \\ \cdots & \cdots & & \cdots \\ \frac{\partial F_m}{\partial x_1}(a) & \frac{\partial F_m}{\partial x_2}(a) & \ldots & \frac{\partial F_m}{\partial x_n}(a) \end{pmatrix}$$

$$\frac{\partial H_i}{\partial x_k}(a) = \sum_{j\in J} \frac{\partial G_i}{\partial y_j}(b)\, \frac{\partial F_j}{\partial x_k}(a)\ ;\quad i\in I\ ;\quad k\in K \tag{III,4;13}$$

On a d'ailleurs aussi

$$\frac{\overrightarrow{\partial h}}{\partial x_k}(a) = \sum_{j\in J} \frac{\overrightarrow{\partial g}}{\partial y_j}(b)\, \frac{\partial F_j}{\partial x_k}(a),\ k\in K\ . \tag{III,4;13 bis}$$

La formule précédente s'écrit autrement, avec un abus de langage qui évidemment peut être dangereux, mais qui rend de grands services. Tout d'abord, on identifie h et g en disant qu'il s'agit toujours de la même fonction g , exprimée tantôt avec la variable x , tantôt avec la variable y . D'autre part, on écrit les dérivées partielles sans spécifier en quel point elles doivent être prises; il est bien évident que la dérivée de f doit être prise au point a , et la dérivée de g au point $f(a) = b$. Par ailleurs la fonction f n'est pas nommée, et au lieu de ses dérivées partielles $\frac{\partial F_j}{\partial x_k}$, on écrit les dérivées partielles $\frac{\partial y_j}{\partial x_k}$. La formule prend alors la forme assez couramment employée

$$\frac{\partial G_i}{\partial x_k} = \sum_{j\in J} \frac{\partial G_i}{\partial y_j}\, \frac{\partial y_j}{\partial x_k}\ ;\quad i\in I\ ;\quad k\in K\ . \tag{III,4;14}$$

On se permet même assez facilement de remplacer g par z, variable courante de l'espace H , ou même de supprimer g complètement :

$$\frac{\partial z_i}{\partial x_k} = \sum_{j\in J} \frac{\partial z_i}{\partial y_j}\, \frac{\partial y_j}{\partial x_k}\ ;\quad i\in I\ ;\quad k\in K\ , \tag{III,4;15}$$

ou

$$\frac{\partial}{\partial x_k} = \sum_{j\in J} \frac{\partial y_j}{\partial x_k}\, \frac{\partial}{\partial y_j}\ ,\quad k\in K\ . \tag{III,4;16}$$

<u>Corollaire 3</u> - <u>Dans les conditions de l'énoncé du corollaire 2, si E, F, G , ont la même dimension n , le déterminant jacobien de l'application composée h au point a est égal au produit du déterminant jacobien de f au point a et du déterminant jacobien de g au point $f(a) = b$</u> . <u>On a la formule</u> :

$$\frac{D(z_1, z_2, \dots z_n)}{D(x_1, x_2, \dots x_n)} = \frac{D(z_1, z_2, \dots z_n)}{D(y_1, y_2, \dots y_n)}\ \frac{D(y_1, y_2, \dots y_n)}{D(x_1, x_2, \dots x_n)}\ . \tag{III,4;16 bis}$$

C'est au fond cette formule qui justifie l'emploi de la notation $\frac{D(y_1, y_2, \cdots y_n)}{D(x_1, x_2, \cdots x_n)}$ pour désigner le déterminant jacobien de la fonction $y = f(x)$ par rapport aux référentiels considérés.

Corollaire 4 - Soit Ω un ouvert de E, Ω' un ouvert de F, et f une bijection de Ω sur Ω', qui soit dérivable en tout point ainsi que sa bijection réciproque f^{-1}. Alors l'application dérivée $f'(a)$ est une bijection de $\vec{E}$ sur $\vec{F}$, et sa bijection réciproque n'est autre que la dérivée, au point $b = f(a)$, de la bijection réciproque f^{-1}. Autrement dit on a la formule :

$$(f'(a))^{-1} = (f^{-1})'(b). \tag{III,4;17}$$

En particulier, si E est F sont de dimension finie, ces dimensions sont égales *

En effet, appliquons le théorème des fonctions composées aux applications f et f^{-1}, qui vérifient les deux formules : $f \circ f^{-1} = I_F$, $f^{-1} \circ f = I_E$; comme l'application dérivée de l'application identique, est l'application identique, on a

$$f'(a) \circ (f^{-1})'(b) = I_{\vec{F}}, \text{ et } (f^{-1})'(b) \circ f'(a) = I_{\vec{E}} ; \text{ ce qui}$$

prouve bien que $f'(a)$ est une bijection, et que sa bijection réciproque est $(f^{-1})'(b)$. Comme cas particulier, si E et F sont de dimension finie, ces dimensions sont égales; et si l'on en a choisi des référentiels, la matrice dérivée de la bijection réciproque f^{-1} au point $b = f(a)$ est la matrice inverse de la matrice dérivée de f au point a ; le déterminant jacobien de la bijection réciproque f^{-1} au point b est l'inverse du déterminant jacobien de f au point a. Cela

* On a vu en algèbre linéaire que, s'il existe une bijection linéaire d'un espace vectoriel de dimension finie sur un autre, les dimensions sont égales. Nous en déduisons ici quelque chose de beaucoup plus fort : s'il existe une bijection, dérivable ainsi que sa bijection réciproque, d'un ouvert d'un espace affine de dimension finie E sur un ouvert d'un espace affine de dimension finie F, alors E et F ont la même dimension. Par des méthodes inspirées de celles de la page 92, on peut démontrer plus: s'il existe un homéomorphisme d'un ouvert de E sur un ouvert de F, E et F ont la même dimension; il n'y a aucune hypothèse de dérivabilité.

s'écrit sous la forme condensée :

(III,4;18) $$\frac{D(y_1, y_2, \dots y_n)}{D(x_1, x_2, \dots x_n)} = \left(\frac{D(x_1, x_2, \dots x_n)}{D(y_1, y_2, \dots y_n)}\right)^{-1}$$

<u>Corollaire 5 - La application composée de 2 plications continuement dérivables est continuement dérivable</u> .

Nous devons montrer que l'application $x \longrightarrow h'(x) = g'(f(x)) \circ f'(x)$ est continue de $\Omega \subset E$ dans $\mathcal{L}(\vec{E};\vec{G})$.

Or $x \longrightarrow f'(x)$ est la fonction, supposée continue, f' ; $x \longrightarrow g'(f(x))$ est la fonction $g' \circ f$, or g' est'supposée continue,et f est continue parce que dérivable, donc $g' \circ f$ est continue d'après le théorème desfonctions composées (théorème 10 du chapitre II); donc $x \longrightarrow (f'(x), g'(f(x)))$ est continue de Ω dans le produit $(\mathcal{L}(\vec{E};\vec{F}) \times \mathcal{L}(\vec{F} \times \vec{G}))$ (remarque suivant le théorème 17 du chapitre II). Mais l'application $x \rightarrow h'(x)$ est composée de celle-ci et de la composition $(u, v) \longrightarrow v \circ u$, application bilinéaire continue de $\mathcal{L}(\vec{E};\vec{F}) \times \mathcal{L}(\vec{F};\vec{G})$ dans $\mathcal{L}(\vec{E};\vec{G})$; le même théorème 10 du chapitre II dit bien qu'elle est continue.

Exemples de calcul de quelques dérivées usuelles

<u>Exemple 1</u> - Soit u une application d'un ouvert Ω d'un espace affine normé E dans la droite réelle $\mathbb{R}$; faisons la suivre de l'application $x \longrightarrow \operatorname{Arc\,tg} x$ de $\mathbb{R}$ dans $\mathbb{R}$. On définit ainsi une nouvelle application de Ω dans $\mathbb{R}$, qu'on appellera tout simplement la fonction $\operatorname{Arc\,tg} u$. Alors, si l'on suppose u dérivable, et compte tenu de ce que la fonction $\operatorname{Arc\,tg} x$ est dérivable, et de dérivée $\frac{1}{1+x^2}$ (Rappelons que, pour des fonctions réelles d'une variable réelle, l'application dérivée est la multiplication par un scalaire, qui est la dérivée usuelle; remarque 2, page 199), le théorème des fonctions composées montre pue la fonction $\operatorname{Arc\,tg} u$ admet elle-même une application dérivée, donnée par la formule :

(III,4;19) $$(\operatorname{Arc\,tg} u)'(a) = \frac{u'(a)}{1+(u(a))^2}, \text{ ou } (\operatorname{Arc\,tg} u)' = \frac{u'}{1+u^2},$$

ou encore, en notation différentielle :

(III,4;20) $$d(\operatorname{Arc\,tg} u) = \frac{du}{1+u^2}.$$

Le même calcul est valable si l'on remplace $Arc\,tg$ par d'autres fonctions. Par exemple, avec Log, on a les formules suivantes (si $u(x) > 0$ pour $x \in \Omega$)' :

$$(III,4;21) \quad \begin{cases} (Log\, u)'(a) = \dfrac{u'(a)}{u(a)}, \text{ ou } (Log\, u)' = \dfrac{u'}{u}, \\ \text{ou} \quad d(Log\, u) = \dfrac{du}{u} \; * . \end{cases}$$

Exemple 2 - Dérivée d'un produit.

La formule classique pour la dérivée d'un produit de fonctions $(uv)' = u'v + uv'$, est un cas particulier de la dérivée d'une application bilinéaire continue.

Théorème 12 - Soient; et $\vec{v}$ des applications dérivables (resp. continuement dérivables) d'un ouvert Ω d'un espace affine normé E, dans des espaces vectoriels normés $\vec{F}_1$ et $\vec{F}_2$, et soit B une application bilinéaire continue de $\vec{F}_1 \times \vec{F}_2$ dans $\vec{G}$. Alors l'application $B(\vec{u}_1, \vec{u}_2) : x \to B(u_1(x), u_2(x))$ de E dans $\vec{G}$,

est dérivable (resp. continuement dérivable), et son application dérivée est donnee par la formule suivante, pour $\vec{X} \in \vec{E}$:

$$(III,4;22) \quad \left(B(\vec{u}_1, \vec{u}_2)\right)'(a) \cdot \vec{X} = B(u_1'(a)\cdot\vec{X}, \vec{u}_2(a)) + B\left(\vec{u}_1(a), u_2'(a)\cdot\vec{X}\right)$$

ou encore

$$(III,4;23) \quad \overrightarrow{D_{\vec{X}} B(\vec{u}_1, \vec{u}_2)}(a) = B\left(\overrightarrow{D_{\vec{X}} \vec{u}_1}(a), \vec{u}_2(a)\right) + B\left(\vec{u}_1(a), \overrightarrow{D_{\vec{X}} \vec{u}_2}(a)\right),$$

ou, en notation différentielle :

$$(III,4;24) \quad \begin{aligned} d(B(\vec{u}_1, \vec{u}_2)) &= B\left(u_1'(a)\cdot d\vec{x}, \vec{u}_2(a)\right) + B\left(\vec{u}_1(a), u_2'(a)\cdot d\vec{x}\right) \\ &= B(d\vec{u}_1, \vec{u}_2) + B(\vec{u}_1, d\vec{u}_2). \end{aligned}$$

* Dans la deuxième formule (III,4;19) ou (III,4;21), les 2 membres sont des applications linéaires continues de $\vec{E}$ dans $\mathbb{R}$, donc des éléments du dual $\vec{E}'$.

Démonstration - L'application considérée est la composée de deux applications; on prend d'abord l'application de E dans $\vec{F}_1 \times \vec{F}_2$ définie par : $x \to (\overrightarrow{u_1(x)}, \overrightarrow{u_2(x)})$, et ensuite l'application B de $\vec{F}_1 \times \vec{F}_2$ dans $\vec{G}$. Chacune de ces deux applications est dérivable (resp. continuement dérivable) (théorème 8 quarto et 9 bis); il suffit alors d'appliquer le théorème des fonctions composées (resp. son corollaire 5).

A titre d'application particulièrement simple supposons que, dans un problème de mécanique, $\vec{u}$ et $\vec{v}$ soient des vecteurs d'un espace affine euclidien à trois dimensions orienté, qui soient des fonctions dérivables du temps t, ayant pour dérivées $\vec{u}'(t)$ et $\vec{v}'(t)$, alors le produit scalaire et produit vectoriel de ces deux vecteurs sont aussi des fonctions dérivables du temps, et l'on a les formules *

(III,4;32)
$$\begin{cases} (\vec{u}\,|\,\vec{v})' = (\vec{u}'\,|\,\vec{v}) + (\vec{u}\,|\,\vec{v}') & \text{(produit scalaire);} \\ (\vec{u}\wedge\vec{v})' = \vec{u}'\wedge\vec{v} + \vec{u}\wedge\vec{v}' & \text{(produit vectoriel).} \end{cases}$$

Naturellement ce que nous venons de dire pour une application bilinéaire est valable pour une application multilinéaire : si l'on a m applications dérivables $\vec{u}_1, \vec{u}_2, \ldots \vec{u}_m$, d'un ouvert Ω de E dans des espaces vectoriels normés $\vec{F}_1, \vec{F}_2, \ldots \vec{F}_m$, et si B est une application m-linéaire continue de $\vec{F}_1 \times \vec{F}_2 \ldots \times \vec{F}_m$ dans $\vec{G}$, alors la fonction $B^* = B(\vec{u}_1, \vec{u}_2, \ldots \vec{u}_m) : x \to B(\overrightarrow{u_1(x)}, \overrightarrow{u_2(x)}, \ldots \overrightarrow{u_m(x)})$ est dérivable, et son application dérivée est donnée par la formule :

(III,4;33)
$$B^{*\prime}(a).\vec{X} = \sum_{j=1}^{m} B\left(\overrightarrow{u_1(a)}, \overrightarrow{u_2(a)}, \ldots \overrightarrow{u_{j-1}(a)}, u_j'(a).\vec{X}, \overrightarrow{u_{j+1}(a)}, \ldots \overrightarrow{u_m(a)}\right).$$

Par exemple, si $\vec{u}$, $\vec{v}$, $\vec{w}$, sont trois vecteurs d'un espace affine euclidien à 3 dimensions orienté, fonctions dérivables du temps, alors leur produit mixte : $(\vec{u}, \vec{v}, \vec{w})$ est une fonction scalaire dérivable du temps, et sa dérivée est donnée par la formule :

(III,4;34)
$$(\vec{u}, \vec{v}, \vec{w})' = (\vec{u}', \vec{v}, \vec{w}) + (\vec{u}, \vec{v}', \vec{w}) + (\vec{u}, \vec{v}, \vec{w}').$$

* Comme t parcourt la droite réelle $\mathbb{R}$, nous prenons, au lieu des applications dérivées au sens de (III,3;13), les vecteurs dérivés au sens de (III,3;1); il suffit pour cela de faire $\vec{X} = 1 \in \mathbb{R}$ dans (III,4;22).

Considdrons maintenant l'application $u \longrightarrow u \circ u \circ \ldots u = u^m$ de $\mathcal{L}(\vec{H};\vec{H})$ dans lui-même, $\vec{H}$ étant un espace vectoriel normé. Elle est composée de l'application $u \rightarrow (u, u, \ldots u)$ de $\vec{E} = \mathcal{L}(\vec{H};\vec{H})$ dans $\vec{F}_1 \times \vec{F}_2 \cdots \times \vec{F}_m = \mathcal{L}(\vec{H};\vec{H}) \times \mathcal{L}(\vec{H};\vec{H}) \ldots \times \mathcal{L}(\vec{H};\vec{H})$, et de la composition, $(u_1, u_2, \ldots u_m) \longrightarrow u_1 \circ u_2 \cdots \circ u_m$, application multilinéaire continue de $\vec{F}_1 \times \vec{F}_2 \ldots \times \vec{F}_m$ dans $\vec{G} = \mathcal{L}(\vec{H};\vec{H})$ (théorème 54 du chapitre II). On peut donc appliquer le théorème 12, et on obtient :

(III,4;35)
$$d(u^m) = du \circ u \circ \cdots \circ u + u \circ du \circ \cdots \circ u + u \circ u \cdots \circ du$$
$$= du \circ u^{m-1} + u \circ du \circ u^{m-2} + \ldots + u^{m-1} \circ du .$$

Si $\vec{H}$ est le corps des scalaires $\mathbb{K}$, il en est de même de $\mathcal{L}(\vec{H};\vec{H})$, qui est alors un corps commutatif, et on retrouve la différentielle classique des fonctions puissances :

$$d(x^m) = m\,x^{m-1}dx .$$

Exemple 3 - Dérivée d'un quotient

Soit u une application d'un ouvert Ω d'un espace affine normé E, dans un espace vectoriel normé $\vec{F}$, et soit v une fonction scalaire définie sur Ω, et partout $\neq 0$. Alors on peut définir le quotient $\frac{\vec{u}}{v} : x \longrightarrow \frac{\vec{u}(x)}{v(x)}$.

Si u et v sont des fonctions dérivables, ce quotient est aussi dérivable, et sa dérivée est donnée par la formule :

(III,4;36)
$$d\left(\frac{\vec{u}}{v}\right) = \frac{v\,d\vec{u} - dv\,\vec{u}}{v^2}$$

La démonstration est immédiate. L'application $\frac{\vec{u}}{v}$ est composée de l'application $x \longrightarrow \left(\vec{u}(x), \frac{1}{v(x)}\right)$ de Ω dans $F \times \mathbb{K}$, et de la multiplication $(\vec{u}, \lambda) \longrightarrow \lambda\vec{u}$ de $\vec{F} \times \mathbb{K}$ dans $\vec{F}$. Comme, en vertu de l'exemple 1, on a $d\left(\frac{1}{v}\right) = -\frac{dv}{v^2} = -\frac{v'(x)\cdot\vec{dx}}{v^2(x)}$, la formule (III,4;24) donne (III,4;36).

Application à la technique du changement de variables. Soit g une fonction d'une variable x, c'est-à-dire une application d'un ouvert Ω d'un espace affine normé E dans un espace affine normé G. Pratiquement, ces espaces seront de dimension finie, et auront même des référentiels donnés. On suppose alors que g satisfasse à une-certaine équation aux dérivées partielles ou à un système d'équations aux dérivées partielles. On se propose de faire le changement de variable : $y = f(x)$; si l'on admet qu'il s'agit d'une bijection, g devient une fonction de la variable y, c'est-&-dire une application d'un ouvert de F dans G . Alors g satisfait à une nouvelle équation ou à un nouveau système d'équations aux dérivées partielles par rapport à y , et c'est ce système que nous nous proposons de trouver; cela suppose qu'on sache à l'avance que la dérivabilité de g par rapport à x est équivalente à sa dérivabilité par rapport à y ; on supposera donc que f est dérivable ainsi que sa bijection réciproque f^{-1} . On appellera x la variable ancienne, y la variable nouvelle, et on cherchera la nouvelle équation aux dérivées partielles à partir de l'ancienne. Comme précisément l'ancienne équation fait intervenir la dérivée dg/dx , il suffira d'exprimer celle-ci à partir de la dérivée $\frac{dg}{dy}$, ce que l'on fait par la formule : $\frac{dg}{dx} = \frac{dg}{dy} \circ f'(x)$, et de porter le résultat dans l'équation aux dérivées partielles considérée . On obtient une nouvelle équation qui contient précisément la dérivée $\frac{dg}{dy}$, et qui est l'équation cherchée; à condition, naturellement, qu'elle ne fasse plus intervenir que la variable y , c'est-à-dire que, partout, la variable x ait été remplacée à partir de la variable y , par son expression $x = f^{-1}(y)$, faisant intervenir la bijection réciproque f^{-1} .

Exemple 1) Equation des cordes vibrantes

On appelle ainsi l'équation :

(III.4;37)
$$\frac{\partial^2 \vec{U}}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 \vec{U}}{\partial t^2} = \vec{0} \quad *$$

où x est une variable spatiale réelle, t une variable temporelle réelle, et $\vec{U}$ une fonction vectorielle des deux Variables x, t , application de $\mathbb{R} \times \mathbb{R}$ dans $\vec{F}$. Voici l'origine physique de ce problème. Considérons une corde homogène d extrémités fixées A et B , et susceptible de vibrer transversalement au voisinage de sa position d'équilibre. Un point de la corde, occupant à l'équilibre la position M_0, occupe dans le mouvement la position M .

* Nous ne devrions donner cet exemple qu'après l'étude des dérivées d'ordre > 1 . Mais celles-ci ont tout de même été étudiées antérieurement. On trouvera les justifications nécessaires au § 6; voir théorème 19.

On pourra, à chaque instant, le connaître par son déplacement, le vecteur $\vec{U} = \overrightarrow{M_0 M}$; si le point M_0 de la corde a sur celle-ci l'abcisse x, $\vec{U}$ devient bien une fonction de x et de t . Il s'agit essentiellement de <u>faibles vibrations</u> au voisinage d une position d'équilibre, <u>c'est-à-dire</u> pour une corde très fortement tendue presque rectiligne entre A et B ; faute de quoi, l'équation ne serait pas une éouation aux dérivées partielles aussi simple que la précédente. Alors $\vec{U}$ prend ses valeurs dans le sous-espace vectoriel à 2 dimensionsperpendiculaire à $\vec{AB}$. Ici v est la vitesse de propagation des vibrations transversales le long de la corde, elle est donnée par la formule :

(III,4;38) $$v = \sqrt{\frac{T}{\rho}}$$

où T est la tension moyenne de la corde, et ρ la masse spécifique linéaire, c'est-à-dire la masse de l'unité de longueur de la corde. Remarquons que l'équation est bien physiquement homogène, car la dimension de v en unités de masse, longueur, et temps, est $\sqrt{\frac{m\ell t^{-2}}{m\ell^{-1}}} = \ell t^{-1}$,

et par conséquent c'est bien la dimension d'une vitesse. *

Effectuons alors le changement de variables défini par

(III,4;39) $$\xi = x + vt, \qquad \eta = x - vt \,.$$

Il s'agit bien d'une bijection, dérivable ainsi que sa bijection réciproque; celle-ci est en effet donnée par

(III,4;40) $$x = \frac{\xi+\eta}{2}, \qquad t = \frac{\xi-\eta}{2v},$$

et les fonctions linéaires sont bien dérivables.

Si nous remplaçons x et t par leurs valeurs dans la fonction U , on obtient la fonction composée $\vec{U}^*$ définie par

(III,4;41) $$\xi, \eta \to \vec{U}^*(\xi,\eta) = \vec{U}\left(\frac{\xi+\eta}{2}, \frac{\xi-\eta}{2v}\right);$$

avec les abus de notation précédents, elle s'écrit encore $\vec{U}$,"supposée exprimée en fonction de ξ et de η ".

Inversement on a la formule réciproque :

(III,4;42) $$\vec{U}(x,t) = \vec{U}^*(x+vt, x-vt).$$

* Naturellement, au cours des oscillations, la longueur de la corde varie très légèrement.

Alors, si l'on calcule les dérivées anciennes, à partir des dérivées nouvelles, on a les formules

(III,4;43) $$\frac{\partial \vec{U}}{\partial x} = \frac{\partial \vec{U}}{\partial \xi} + \frac{\partial \vec{U}}{\partial \eta} \quad , \qquad \frac{\partial \vec{U}}{\partial t} = v\left(\frac{\partial \vec{U}}{\partial \xi} - \frac{\partial \vec{U}}{\partial \eta}\right)$$

ce qui donne, pour les dérivées partielles du second ordre, les formules :

(III,4;44) $$\frac{\partial^2 \vec{U}}{\partial x^2} = \frac{\partial^2 \vec{U}}{\partial \xi^2} + 2\frac{\partial^2 \vec{U}}{\partial \xi \partial \eta} + \frac{\partial^2 \vec{U}}{\partial \eta^2} \; , \quad \frac{1}{v^2}\frac{\partial^2 \vec{U}}{\partial t^2} = \frac{\partial^2 \vec{U}}{\partial \xi^2} - 2\frac{\partial^2 \vec{U}}{\partial \xi \partial \eta} + \frac{\partial^2 \vec{U}}{\partial \eta^2} .$$

On en déduit alors que l'on a :

(III,4;45) $$\frac{\partial^2 \vec{U}}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 \vec{U}}{\partial t^2} = 4\frac{\partial^2 \vec{U}}{\partial \xi \partial \eta} .$$

Il y a toujours la même confusion entre U et U^* ; cette formule devrait en réalité s'écrire :

(III,4;46) $$\frac{\partial^2 \vec{U}}{\partial x^2}(x,t) - \frac{1}{v^2}\frac{\partial^2 \vec{U}}{\partial t^2}(x,t) = 4\frac{\partial^2 \vec{U}^*}{\partial \xi \partial \eta}(x+vt,\; x-vt) .^*$$

La nouvelle équation obtenue, pour $\vec{U}$ exprimée en fonction de ξ et de η (en réalité pour $\vec{U}^*$), est donc l'équation aux dérivées partielles beaucoup plus simple :

(III,4;47) $$\frac{\partial^2 \vec{U}}{\partial \xi \partial \eta} = 0 .$$

Il est trés facile de résoudre cette équation. En effet, si nous considérons la dérivée partielle $\frac{\partial \vec{U}}{\partial \xi}$, on voit que sa dérivée partielle par rapport à η est nulle; il en résulte qu'elle est une constante par rapport à η, c'est-à-dire une fonction "arbitraire' de ξ. On en déduit alors que $\vec{U}$ est la somme d'une primitive de cette fonction de ξ, qui est elle-même une fonction "arbitraire" de ξ, augmentée d'une constante par rapport à ξ, c'est-à-dire d'une fonction "arbitraire" de η.

On a finalement la formule :

(III,4;47bis) $$\vec{U} = \vec{f}(\xi) + \vec{g}(\eta) ,$$

d'où l'on déduit la solution cherchée de l'équation aux dérivées partielles des cordes vibrantes :

(III,4;48) $$\vec{U}(x,t) = \vec{f}(x+vt) + \vec{g}(x-vt) ,$$

* Les 2 membres sont des fonctions de x et t. Comme toujours, $\frac{\partial^2 \vec{U}^*}{\partial \xi \partial \eta}(x+vt, x-vt)$ veut dire la valeur de $\frac{\partial^2 \vec{U}^*}{\partial \xi \partial \eta}$ au point de coordonnées $x+vt,\; x-vt$.

$\vec{f}$ et $\vec{g}$ **étant** des fonctions "arbitraires" d'une **variable réelle.** Quand nous avons **dit**"**fonction** arbitraire' , ce n'est **évidemment** pas exact; **il** faut que tous les calculs **précédents** soient justifies dans le changement de variables (voir théorème 19); pour cela on supposera que la fonction $\vec{U}$ admet une dérivée totale du 2ème ordre; cela suppose les fonctions $\vec{f}$ **et** $\vec{g}$ d'une variable deux fois dérivables.

Exercice - **Résoudre** de la même manière l'équation des cordes vibrantes avec second membre :

(III,4;49) $$\frac{\partial^2 \vec{U}}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 \vec{U}}{\partial t^2} = \vec{g}(x,t),$$

où le second membre est une fonction continue donnée de x **et de** t **,** cette **équation** s'introduit lorsque la corde est **supposée soumise à** d'autres forces que sa **tension,transversales** et ne dépendant que de x et de t .

Exemple 2) : **Dérivées** partielles en coordonnées **cartésiennes** et en **coordonnées** polaires planes.

1°/ Dérivées en polaires **à** partir des **dérivées** en **cartésiennes.**

D'après les formules de transformation :

(III,4;50) $$x = r\cos\varphi, \qquad y = r\sin\varphi,$$

et en tenant compte de l'abus de langage **déjà signalé,** on a **immédiatement** la formule de transformation :

(III,4;51) $$\frac{\partial}{\partial r} = \cos\varphi\frac{\partial}{\partial x} + \sin\varphi\frac{\partial}{\partial y}, \quad \frac{\partial}{\partial\varphi} = -r\sin\varphi\frac{\partial}{\partial x} + r\cos\varphi\frac{\partial}{\partial y}.$$

Que signifie exactement cette formule **?**

Soit g une application du **plan**$\mathbb{R}^2$ **dans** un espace affine sur le corps des **réels.** Elle peut **s'écrire** $(x,y) \longrightarrow g(x,y)$.

Soit d'autre part P l'application **de** $\mathbb{R}^2$ **dans** lui-même, définie par

(III,4;52) $$(r,\varphi) \longrightarrow (x,y), \quad x = r\cos\varphi, \qquad y = r\sin\varphi.$$

(Attention **!** on appelle (r,φ) le point objet **de** $\mathbb{R}^2$**,donc** r **et** φ sont ses coordonnées **cartésiennes** usuelles. Son image $P(r,\theta)$ **a** pour **coordonnées cartésiennes** x, y , données par **(III,4;52), donc** (r,φ) est un système de coordonnées polaires pour ce point Image).

L'application P est **dérivable,** et sa matrice **dérivée** est donnée par la formule : (*)

(III,4;53) $$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial\varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial\varphi} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

(*) Voir ce renvoi **à** la page suivante.

Il en résulte que, si g est une fonction dérivable, l'application composée $g^* \stackrel{d}{=} g \circ P$ est aussi dérivable, et elle vérifie la formule :

$$
(\text{III } 4;54) \quad \begin{cases} \dfrac{\overrightarrow{\partial g^*}}{\partial r}(r,\varphi) = \dfrac{\overrightarrow{\partial g}}{\partial x}(r\cos\varphi, r\sin\varphi)\cos\varphi + \dfrac{\overrightarrow{\partial g}}{\partial y}(r\cos\varphi, r\sin\varphi)\sin\varphi \\[2ex] \dfrac{\overrightarrow{\partial g^*}}{\partial \varphi}(r,\varphi) = -\dfrac{\overrightarrow{\partial g}}{\partial x}(r\cos\varphi, r\sin\varphi)\, r\sin\varphi + \dfrac{\overrightarrow{\partial g}}{\partial y}(r\cos\varphi, r\sin\varphi)\, r\cos\varphi \end{cases}
$$

qu'on écrira, avec l'abus de notation signalé antérieurement :

$$
(\text{III } 4;55) \quad \frac{\overrightarrow{\partial g}}{\partial r} = \cos\varphi \frac{\overrightarrow{\partial g}}{\partial x} + \sin\varphi \frac{\overrightarrow{\partial g}}{\partial y} \; ; \; \frac{\overrightarrow{\partial g}}{\partial \varphi} = -r\sin\varphi \frac{\overrightarrow{\partial g}}{\partial x} + r\cos\varphi \frac{\overrightarrow{\partial g}}{\partial y}
$$

(comme toujours, on identifie g^* et g). Cela donne bien (III,4,51).

2°/ <u>Dérivées en cartésiennes à partir des dérivées en polaires,</u>

On peut aussi se proposer de calculer les dérivées partielles en x et y, , à partir des dérivées partielles en r et φ . Mais il faut remarquer-que P n'est pas une bijection. Elle est bien surjective, mais n'est pas injective,

* Renvoi de la page 226 -

- Il résulte du théorème 8 quarto que P est dérivable, si les deux fonctions scalaires $(r,\varphi) \to r\cos\varphi, (r,\varphi) \to r\sin\varphi$ sont dérivables. Or chacune des deux est un produit, donc est dérivable si les facteurs sont dérivables. Tout revient donc à montrer que $(r,\varphi) \to r$ et $(r,\varphi) \to \cos\varphi$ ou $\sin\varphi$ sont ddrivables. Or $(r,\varphi) \to r$ est linéaire; $(r,\varphi) \to \cos\varphi$ est composée de $(r,\varphi) \to \varphi$ et $\varphi \to \cos\varphi$, la première est linéaire et la seconde est connue comme fonction réelle dérivable d'une variable réelle.

On pourra aussi utiliser le corollaire 2 du théorème 15 : $(r,\varphi) \to r\cos\varphi$ et $(r,\varphi) \to r\sin\varphi$ ont des dérivées partielles continues.

car un point (x,y) a une infinité de systèmesde coordonnées polaires. C'est pourquoi il faut s'attendre à ce que l'application dérivée de P ne soit pas elle-même une bijection linéaire. Toutefois, au voisinage d'un point (r,φ) où le déterminant jacobien de P , c'est-à-dire $\begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r$ est différent de 0 autrement dit d'un point $(x,y) \neq (0,0)$, il sera possible de résoudre * les équations (III,4;54) et de trouver les dérivées partielles $\frac{\partial \vec{g}}{\partial x}$, $\frac{\partial \vec{g}}{\partial y}$, à partir des dérivées partielles : $\frac{\partial \vec{g}^*}{\partial r}$, $\frac{\partial \vec{g}^*}{\partial \varphi}$. On obtient immédiatement :

$$(\text{III},4;56)\quad \begin{cases} \dfrac{\partial \vec{g}}{\partial x}(r\cos\varphi, r\sin\varphi) = \cos\varphi\, \dfrac{\partial \vec{g}^*}{\partial r}(r,\varphi) - \dfrac{\sin\varphi}{r}\, \dfrac{\partial \vec{g}^*}{\partial \varphi}(r,\varphi) \\[2ex] \dfrac{\partial \vec{g}}{\partial y}(r\cos\varphi, r\sin\varphi) = \sin\varphi\, \dfrac{\partial \vec{g}^*}{\partial r}(r,\varphi) + \dfrac{\cos\varphi}{r}\, \dfrac{\partial \vec{g}^*}{\partial \varphi}(r,\varphi) \end{cases}$$

On peut ensuite, dans cette formule, remplacer r et φ par l'un quelconque de leurs systèmes de valeurs possibles en fonction de x , y

On peut aussi voir la chose d'une autre manière. Soit (r_0, φ_0) un point de $\mathbb{R}^2$, $r_0 > 0$ par exemple; et soit (x_0, y_0) son image par P . Si l'on restreint P à l'ouvert Ω de $\mathbb{R}^2$ défini par les inégalités $r > 0$, $\varphi_0 - \pi < \varphi < \varphi_0 + \pi$, alors c'est une bijection P_Ω (et même un homéomorphisme) de Ω sur $P(\Omega)$, et elle a donc une bijection réciproque P^{-1}. Tant que (x,y) varie dans $P(\Omega)$, on pourra lui affecter un système unique de coordonnées polaires variant continuement, c'est-a-dire une image réciproque unique dans Ω , et on aura $g = g^* \circ P_\Omega^{-1}$.

Les formules (III,4;56), où r et φ sont remplacées par leurs valeurs en fonction de x , y , ont alors la même signification que les formules (III,4;54).

* Mous utilisons ici un cas particulier, évident directement, du théorème de fonctions implicites (§ 8; voir en particulier remarque page 296).

Elles s'écrivent

$$\text{(III, 4 ; 56}^{\text{bis}}\text{)}\quad \begin{cases} \dfrac{\overrightarrow{\partial g}}{\partial x}(x,y) = \dfrac{x}{\sqrt{x^2+y^2}}\,\dfrac{\overrightarrow{\partial g}^*}{\partial r}\left(\sqrt{x^2+y^2}\,,\varphi(x,y)\right) - \dfrac{y}{x^2+y^2}\,\dfrac{\overrightarrow{\partial g}^*}{\partial \varphi}\left(\sqrt{x^2+y^2}\,,\varphi(x,y)\right) \\ \dfrac{\overrightarrow{\partial g}}{\partial y}(x,y) = \dfrac{y}{\sqrt{x^2+y^2}}\,\dfrac{\overrightarrow{\partial g}^*}{\partial r}\left(\sqrt{x^2+y^2}\,,\varphi(x,y)\right) + \dfrac{x}{x^2+y^2}\,\dfrac{\overrightarrow{\partial g}^*}{\partial \varphi}\left(\sqrt{x^2+y^2}\,,\varphi(x,y)\right) \end{cases}$$

On a des formules analogues si $r_0 < 0$, en prenant $r < 0$, et $r = -\sqrt{x^2+y^2}$. Quant à la fonction continue $\varphi : (x,y) \longrightarrow \varphi(x,y)$, elle n'a pas une expression analytique simple, on a évidemment $\varphi(x,y) = \text{Arctg}\,\frac{y}{x} + k\pi$, mais k dépend de x, y , et n'est pas toujours le même pour $(x,y) \in P(\Omega)$

On aurait aussi pu, pour résoudre les deux mêmes problèmes, utiliser la notation différentielle. Il suffit en effet d'écrire la différentielle sous la forme

(III,4;57) $$\vec{dq} = \frac{\partial \vec{q}}{\partial x}\, dx + \frac{\partial \vec{q}}{\partial y}\, dy ,$$

puis d'utiliser les formules :

(III,4;58) $$dx = \cos\varphi\, dr - r\sin\varphi\, d\varphi \quad , \quad dy = \sin\varphi\, dr + r\cos\varphi\, d\varphi ,$$

ce qui donne la nouvelle expression de la différentielle

(III,4;59) $$\vec{dq}^* = \left(\cos\varphi \frac{\partial \vec{q}}{\partial x} + \sin\varphi \frac{\partial \vec{q}}{\partial y}\right) dr + \left(- r\sin\varphi \frac{\partial \vec{q}}{\partial x} + r\cos\varphi \frac{\partial \vec{q}}{\partial y}\right) d\varphi .$$

Les coefficients de dr et $d\varphi$ ne sont autres que $\frac{\partial \vec{q}^*}{\partial r}$, $\frac{\partial \vec{q}^*}{\partial \varphi}$, ce qui redonne la solution du 1er problème et la formule (III,4;54).

Pour retrouver (III,4;56), on fait le même raisonnement en partant de

(III,4;60) $$\vec{dq}^* = \frac{\partial \vec{q}^*}{\partial r}\, dr + \frac{\partial \vec{q}^*}{\partial \varphi}\, d\varphi .$$

On calcule dr et $d\varphi$ en fonction de dx et dy par résolution du système (III,4;58), c'est-à-dire

(III,4;61) $$dr = \cos\varphi\, dx + \sin\varphi\, dy \ , \ d\varphi = -\frac{1}{r}\sin\varphi\, dx + \frac{1}{r}\cos\varphi\, dy ,$$

d'où, en portant dans (III,4;60):

(III,4;62) $$\vec{dq} = \left(\cos\varphi \frac{\partial \vec{q}^*}{\partial r} - \frac{\sin\varphi}{r} \frac{\partial \vec{q}^*}{\partial \varphi}\right) dx + \left(\sin\varphi \frac{\partial \vec{q}^*}{\partial r} + \frac{\cos\varphi}{r} \frac{\partial \vec{q}^*}{\partial \varphi}\right) dy ,$$

d'où l'on déduit de nouveau la solution du 2ème problème et (III,4;56).

A titre d'application, calculons par exemple le <u>laplacien en coordonnées polaires planes</u>.

Le laplacien d'une fonction $\vec{q}$ de x et y est défini par :

(III,4;63) $$\Delta \vec{q} = \frac{\partial^2 \vec{q}}{\partial x^2} + \frac{\partial^2 \vec{q}}{\partial y^2}$$

On a d'abord, pour les dérivées partielles du premier ordre, les formules. (III,4;56). En calculant alors les dérivées partielles du second ordre comme dérivées de celles du premier, on a

(III,4;64)
$$\frac{\partial^2 \vec{g}}{\partial x^2} = \left(\cos\varphi \frac{\partial}{\partial r} - \frac{\sin\varphi}{r}\frac{\partial}{\partial\varphi}\right)\left(\cos\varphi \frac{\partial \vec{g}^*}{\partial r} - \frac{\sin\varphi}{r}\frac{\partial \vec{g}^*}{\partial\varphi}\right)$$

qui est une expression de la forme

(III,4;65)
$$\vec{\alpha}\cos^2\varphi + \vec{\beta}\sin\varphi\cos\varphi + \vec{\gamma}\sin^2\varphi \ .$$

On passe des dérivées partielles en x aux dérivées partielles en y en remplaçant φ par $(\varphi - \frac{\pi}{2})$, c'est-à-dire $\cos\varphi$, $\sin\varphi$ par $\sin\varphi$, $-\cos\varphi$ respectivement. Alors l'expression de $\frac{\partial^2 \vec{g}}{\partial y^2}$ est $\vec{\alpha}\sin^2\varphi - \vec{\beta}\sin\varphi\cos\varphi + \vec{\gamma}\cos^2\varphi$, et le laplacien vaut $\vec{\alpha} + \vec{\gamma}$, c'est-à-dire la somme des coefficients de $\cos^2\varphi$ et de $\sin^2\varphi$ dans (III,4;65).

Le calcul est alors immédiat et donne :

(III,4;66)
$$\vec{\Delta g}\,(r\cos\varphi, r\sin\varphi) = \frac{\partial^2 \vec{g}^*}{\partial r^2}(r,\varphi) + \frac{1}{r}\frac{\partial \vec{g}^*}{\partial r}(r,\varphi) + \frac{1}{r^2}\frac{\partial^2 \vec{g}^*}{\partial\varphi^2}(r,\varphi)$$

qu'on écrit, avec l'abus de langage habituel :

$$\vec{\Delta g} = \frac{\partial^2 \vec{g}}{\partial r^2} + \frac{1}{r}\frac{\partial \vec{g}}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \vec{g}}{\partial\varphi^2} \qquad \text{ou}$$

(III,4;67)
$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\varphi^2}$$

<u>Remarques</u> 1°) Cette formule vérifie bien les règles de l'homogénéité. Si les quantités x, y, r, ont la dimension ℓ de longueurs, et si φ est sans dimension, $\vec{\Delta g}$ a pour dimension $g\,\ell^{-2}$, et il en est de même des termes du 2ème membre.

2°) Tout ce calcul suppose que g ait une dérivée totale du 2ème ordre (voir § 6, théorème 19)

<u>Exemple 3 - Calcul du Laplacien d'une fonction définie sur $\mathbb{R}^n$ et ne dépendant que de la distance à l'origine.</u>

Soit g une application de la demi-droite $\mathbb{R}_+$ (ensemble des nombres réels ≥ 0) dans un espace affine normé $\mathcal{G}$

On considère ensuite l'application r de $\mathbb{R}^n$ dans $\mathbb{R}_+$ définie par :

(III,4;68)
$$r(x_1, x_2, \ldots, x_n) = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

L'application composée $g^* = g \circ r$ qu'on écrira plus brièvement $g(r)$, est alors une application de $\mathbb{R}^n$ dans G, ne dépendant que de la distance du point objet à l'origine. Cette fonction possède un Laplacien $\vec{\Delta} g^* = \sum_{i=1}^{n} \dfrac{\overrightarrow{\partial^2 g^*}}{\partial x^2}$ au point $(x_1, x_2, \ldots x_n)$, si g est 2 fois dérivable, et $(x_1, x_2, \ldots x_n) \neq (0,0,\ldots 0)$ (pour que r soit aussi 2 fois dérivable).

Ce Laplacien se calcule par les méthodes précédentes. D'abord :

(III,4;69)
$$\begin{cases} \dfrac{\overrightarrow{\partial g^*}}{\partial x_i} = \vec{g}' \dfrac{\partial r}{\partial x_i} = \vec{g}' \dfrac{x_i}{r} \\[2ex] \dfrac{\overrightarrow{\partial^2 g^*}}{\partial x_i^2} = \vec{g}'' \dfrac{x_i^2}{r^2} + \vec{g}' \dfrac{1}{r} - \vec{g}' \dfrac{x_i^2}{r^3} \ ; \end{cases}$$

En ajoutant ensuite les résultats obtenus pour $i = 1, 2, \ldots, n$, on obtient immédiatement :

(III,4;70)
$$\vec{\Delta} g^*(x_1, x_2, \ldots x_n) = \vec{g}''\left(\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}\right) + \frac{n-1}{\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}} \vec{g}'\left(\sqrt{x_1^2 + x_2^2 + \cdots x_n^2}\right),$$

qu'on écrira plus brièvement

(III,4;71)
$$\vec{\Delta} g = \vec{g}'' + \frac{n-1}{r} \vec{g}' = \frac{\overrightarrow{d^2 g}}{dr^2} + \frac{n-1}{r} \frac{\overrightarrow{dg}}{dr}$$

Dans le cas particulier $n = 2$, cette formule est un cas particulier de (III,4;66), à condition d'intervertir les rôles de g et g^* (Si, dans (III,4;66), on suppose que g^* ne dépend que de r, on a $\dfrac{\overrightarrow{\partial^2 g^*}}{\partial \varphi^2} = 0$, et on obtient $\vec{\Delta} g = \dfrac{\overrightarrow{\partial^2 g^*}}{\partial r^2} + \dfrac{1}{r} \cdot \dfrac{\overrightarrow{\partial g^*}}{\partial r}$), ce qui est l'équivalent de (III,4;71) pour $n = 2$.

(On retiendra facilement la présence du facteur $n-1$: pour $n = 1$, on a exactement $\vec{\Delta} g = \vec{g}''$. On remarquera aussi que la formule vérifie les règles de l'homogénéité : les 2 membres ont la dimension $g\,\ell^{-2}$ (voir fin de l'exemple 2')).

§ 5 FORMULE DES ACCROISSEMENTS FINIS

Supposer que f ait une dérivée en $a \in \Omega$, c'est donner, d'après (III,3;13), une estimation de $f(a+\vec{h}) - f(a)$ pour $\vec{h}$ infiniment petit; on pourrait appeler (III,3;13) formule des accroissements infiniment petits. Mais (III,3;13) né donne aucune estimation de $f(a+\vec{h}) - f(a)$ pour une fonction f et un accroissement $\vec{h}$ déterminés; le but de la formule des accroissements finis est de donner de telles estimations, en introduisant la dérivée de aux points voisins de a.

La formule (III,2;2) se généralise immédiatement à une application de E dans F, sif est le corps des réels lui-même.

Théorème 13 A.

Soit f une fonction réelle continue sur un ouvert Ω d'un espace affine normé E sur le corps des réels. Si le segment $[x, x+\vec{h}]$ appartient tout entier à Ω, et si f a une application dérivée en tout point du segment ouvert $]x, x+\vec{h}[$, alors on a

$$f(x+\vec{h}) - f(x) = f'(x+\theta\vec{h}) \cdot \vec{h} \in \mathbb{R}, \quad 0 < \theta < 1. \tag{III,5;0}$$

Démonstration -

Considérons en effet l'application $\Phi : t \longrightarrow f(x+t\vec{h})$ de l'intervalle réel $[0,1]$ dans $F = \mathbb{R}$. Elle est composée de l'application $t \longrightarrow x+t\vec{h}$ de $[0,1]$ dans Ω, et de l'application f, de Ω dans $\mathbb{R}$. D'après le théorème des fonctions composées, elle est donc continue dans $[0,1]$ (théorème 10 du chapitre II), et dérivable dans $]0,1[$ (théorème 11 du chapitre III), de dérivée

$$\Phi'(t) = f'(x+t\vec{h}) \cdot \vec{h} \in \mathbb{R} \tag{III,5;0 bis}$$

Il suffit alors d'appliquer à Φ la formule (III,2;2), pour l'intervalle $[0,1]$, et comme $\Phi(1) - \Phi(0) = f(x+\vec{h}) - f(x)$, on obtient le résultat.

Par contre (III,2;2) ne peut pas s'étendre telle quelle, si F est de dimension > 1, ou s'il s'agit d'espaces affines sur le corps des complexes; elle n'est déjà plus vraie pour une fonction complexe d'une variable réelle *. Considérons en effet la fonction complexe: $x \longrightarrow e^{2i\pi x}$ définie sur le segment $[0,1]$ de $\mathbb{R}$. Cette fonction prend la

* C'est-à-dire une application de $\mathbb{R}$ dans $\mathbb{C}$, considérés comme espaces vectoriels respectivement à 1 et 2 dimensions sur le corps des réels $\mathbb{R}$.

même valeur pour $x = 0$ et pour $x = 1$. Si la formule des accroissements finis était vraie sous la forme du théorème (III,2;2), il devrait exister un point c de l'intervalle $]0,1[$ où sa dérivée serait nulle; or sa dérivée, qui est la fonction $x \longrightarrow 2i\pi e^{2i\pi x}$, est partout différente de 0.

Nous allons cependant donner la formule des accroissements finis, sous une forme un peu différente, dans le cas le plus général :

Théorème 13 - Soit f une application continue d'un ouvert Ω d'un espace affine normé E, dans un espace affine normé F. Alors, si le segment $[x, x+\vec{h}]$ appartient tout entier à Ω, et si f a une application dérivée en tout point du segment ouvert $]x, x+\vec{h}[$, de norme $\leq M$, on a la majoration :

(III,5;1) $$\left\|\overrightarrow{f(x+\vec{h})-f(x)}\right\| \leq M \|\vec{h}\|$$

Démonstration - Nous allons d'abord démontrer un lemme.

Lemme : Soient f une application du segment $[0,1]$ de la droite réelle $\mathbb{R}$, dans un espace affine normé F, et g une fonction réelle sur $[0,1]$. On suppose f et g continues sur le segment fermé $[0,1]$, et dérivables sur le segment ouvert $]0,1[$. Dans ces conditions, si l'on a la majoration :

(III,5;2) $$\left\|\overrightarrow{f'(x)}\right\| \leq g'(x), \quad \text{pour } 0 < x < 1,$$

on a aussi la majoration :

(III,5;3) $$\left\|\overrightarrow{f(1)-f(0)}\right\| \leq g(1)-g(0).$$

Démontrons d'abord ce lemme :

Soit $\varepsilon > 0$. Désignons par A_ε l'ensemble des points x de l'intervalle $[0,1]$ tels qu'on ait la majoration

(III,5;4) $$\left\|\overrightarrow{f(x)-f(0)}\right\| \leq g(x)-g(0)+\varepsilon x + \varepsilon.$$

La fonction

(III,5;4 bis) $$x \to \left\|\overrightarrow{f(x)-f(0)}\right\| - g(x) + g(0) - \varepsilon x - \varepsilon$$

est continue; l'ensemble A_ε des points où elle est ≤ 0 est fermé. En particulier A_ε contient sa borne supérieure β, qui est donc un maximum. Tout d'abord on ne peut pas avoir $\beta = 0$; en effet, f et g étant continues, on a, pour

x suffisamment petit, l'inégalité $\|\overrightarrow{f(x)-f(0)}\| \leq \varepsilon$, donc à fortiori (III,5;4) : A_ε contient un voisinage de 0. Mais on ne peut pas avoir non plus $0 < \beta < 1$. En effet, d'après la définition même de la dérivée, (formule (III,3;13)), il existerait alors un nombre $\delta > 0$ tel que l'on ait les inégalités

$$(\text{III},5;5)\quad \begin{cases} \|\overrightarrow{f(\beta+\delta)-f(\beta)}\| \leq \|\overrightarrow{f'(\beta)}\|\delta + \frac{\varepsilon}{2}\delta \leq g'(\beta)\delta + \frac{\varepsilon}{2}\delta \\ g(\beta+\delta)-g(\beta) \geq g'(\beta)\delta - \frac{\varepsilon}{2}\delta \end{cases}$$

d'où

$$(\text{III},5;6)\quad \|\overrightarrow{f(\beta+\delta)-f(\beta)}\| \leq g(\beta+\delta) - g(\beta) + \varepsilon\delta .$$

Mais, puisque $\beta \in A_\varepsilon$, on a

$$(\text{III},5;6\,bis)\quad \|\overrightarrow{f(\beta)-f(0)}\| \leq g(\beta) - g(0) + \varepsilon\beta + \varepsilon ,$$

donc, en additionnant :

$$(\text{III},5;6\,ter)\quad \|\overrightarrow{f(\beta+\delta)-f(0)}\| \leq g(\beta+\delta) - g(0) + \varepsilon(\beta+\delta) + \varepsilon ;$$

cela entraînerait $\beta+\delta \in A_\varepsilon$, ce qui est absurde puisque β est le maximum de A_ε.

On a donc $\beta = 1$. (III,5;6bis) s'écrit donc

$$(\text{III},5;6\,quarto)\quad \|\overrightarrow{f(1)-f(0)}\| \leq g(1) - g(0) + 2\varepsilon .$$

Cette inégalité étant vraie quelle que soit $\varepsilon > 0$ on en déduit bien l'inégalité (III,5;3), et le lemme est démontré'.

Remarques sur le lemme. 1°/ Au lieu de supposer f et g dérivables, on peut se borner à les supposer dérivables à droite, avec $\|\overrightarrow{f'_d}\| \leq g'_d$; la démonstration le montre immédiatement. En remplaçant $f(x)$, $g(x)$, par $-f(1-x)$, $-g(1-x)$, on peut supposer la dérivabilité à gauche.

2°/ supposons qu'en outre, en au moins un point c de $]0,1[$, on ait une inégalité stricte $\|\overrightarrow{f'_d(c)}\| < g'_d(c)$. Alors on a nécessairement $\|\overrightarrow{f(1)-f(0)}\| < g(1) - g(0)$.

Si en effet $\delta = g'_d(c) - \|\overrightarrow{f'_d(c)}\| > 0$, on a, pour $h > 0$ assez petit,

$$\|\overrightarrow{f(c+h)-f(c)}\| \leq \|\overrightarrow{f'_d(c)}\|h + \frac{\delta}{3}h \leq g'_d(c) - \frac{2\delta}{3}h ,$$

$$g(c+h)-g(c) \geq g'_d(c) - \frac{\delta}{3}h , \quad \text{d'où}$$

$$(\text{III},5;6\,quinto)\quad \|\overrightarrow{f(c+h)-f(c)}\| \leq g(c+h) - g(c) - \frac{\delta}{3}h < g(c+h) - g(c) .$$

Mais le lemme, appliqué aux intervalles $[0,c]$, $[c+h,1]$, au lieu de $[0,1]$, donne

(III,5;6 sexto) $$\|\overrightarrow{f(c)-f(0)}\| \leq g(c)-g(0)$$

$$\|\overrightarrow{f(1)-f(c+h)}\| \leq g(1)-g(c+h)$$

L'addition de (III,5;6 quinto et sexto) donne le résultat.

3°/ Si f est aussi réelle, on peut supposer qu'elle a partout sur $]0,1[$ une dérivée finie ou égale à $-\infty$, et qu'on a seulement une inégalité du type $f'_d \leq g_d$. Alors on aura une inégalité $f(1)-f(0) \leq g(1)-g(0)$, avec inégalité stricte si, en au moins un point c de $]0,1[$, on a $f'_d(c) < g'_d(c)$.

4°/ De là on déduit les extensions signalées aux remarques suivant les théorèmes 4,5,7. Prenons le cas du théorème 4 (de Rolle), qui entraîne les autres. On ne peut pas avoir partout $f'_d(x) < 0$; en effet, en prenant $g = 0$, cela entraînerait $f(1)-f(0) < 0$, contrairement à l'hypothèse. Pour la même raison on ne peut pas avoir partout $f'_d(x) > 0$; donc il existe c_1 tel que $f'_d(c) \leq 0$ et c_2 tel que $f'_d(c_2) \geq 0$.

Demontrons maintenant le théorème. Considérons l'application $\Phi : t \longrightarrow f(x+th)$ de $[0,1]$ dans F, déjà considérée dans la démonstration du théorème 13 A.

Sa dérivée s'écrit donc sous la forme :

(III,5;7) $$\overrightarrow{\Phi'(t)} = f'(x+th).\vec{h} \quad .$$

Il en résulte que, dans les hypothèses du théorème, la norme de cette dérivée est majorée par $M\|\vec{h}\|$. Si donc on lui applique le lemme, en prenant pour g la fonction linéaire $t \longrightarrow M\|\vec{h}\|t$, on a bien la majoration (III,5;1).

Le théorème des accroissements finis admet l'interessant complément suivant :

<u>Corollaire 1</u> - <u>Dans les conditions de l'énoncé du théorème 13, si L est une application linéaire continue de E dans F, on a la majoration</u> :

(III,5;8) $$\left\|\overrightarrow{f(x+\vec{h})-f(x)} - L\vec{h}\right\| \leq \omega \|\vec{h}\| ,$$

<u>où ω est la borne supérieure de la quantité</u> $\|f'(\xi)-L\|$, <u>lorsque</u> ξ <u>varie dans l'intervalle</u> $]x, x+h[$. Assez généralement, on suppose que f admet en outre une dérivée en x, et on prend $L = f'(x)$.

<u>Démonstration</u> - Il suffit-pliquer le théorème à la fonction $\xi \longrightarrow f(\xi) - L\cdot(\xi - x)$, dont la dérivée au point ξ est $f'(\xi) - L$.

<u>Corollaire 2</u> - <u>Si Ω est un ouvert convexe* de l'espace affine normé E , et si f est une application dérivable de Ω dans F , dont l'application dérivée en tout point de Ω est majorée en norme par une constante M , alors la fonction f est lipschitzienne, et par conséquent uniformément continue.</u>

<u>Démonstration</u> - Du fait que Ω est convexe, si x' et x'' sont deux points quelconques de Ω, tout le segment $[x', x'']$ appartient à Ω ; on peut donc lui appliquer la formule des accroissements finis, qui donne la majoration

$$(\mathrm{III},5;9) \qquad \| \overrightarrow{f(x') - f(x'')} \| \leq M \| \overrightarrow{x' - x''} \| ,$$

ce qui prouve le théorème.

<u>Corollaire3</u> - <u>Soit f une application dérivable d'un espace affine normé E dans un espace affine normé F ; on suppose en outre que sa dérivee f' ~~soit une~~ application uniformement continue de E dans $\mathcal{L}(\vec{E};\vec{F})$. Dans ces conditions, la fonction $x \longrightarrow \overrightarrow{f(x+\vec{h}) - f(x)} - f'(x).\vec{h}$ converge uniformément vers $\vec{0}$, lorsaue $\vec{h} \neq \vec{0}$ tend vers $\vec{0}$.</u>

<u>Démonstration</u> Etant donné $\varepsilon > 0$, on peut déterminer $\eta > 0$ de manière que l'inégalité $\| \overrightarrow{x' - x''} \| \leq \eta$ entraîne l'inégalité : $\| f'(x') - f'(x'') \| \leq \varepsilon$, en vertu de l'hypothèse de continuité uniforme de la dérivée f' . Alors, dès que $\| \vec{h} \|$ est $\leq \eta$, la quantité ω , qui intervient dans la formule (III,5;8) relative à $L = f'(x)$, est majorée par ε , ce qui prouve bien le théorème.

<u>Remarques</u> : 1°/ Supposons que E soit le corps des scalaires. Alors le <u>corollaire 3</u> revient à dire que la fonction $x \longrightarrow \dfrac{\overrightarrow{f(x+h) - f(x)}}{h}$ converge uniformément vers la fonction dérivée f' , lorsque $h \neq 0$ tend vers $\vec{0}$.

On dit encore que f est <u>uniformément dérivable</u>.

* Cette condition de convexité est absolument essentielle.

2°/ On applique souvent ce théorème dans un cas un peu différent. On suppose que f est seulement dérivable sur un ouvert Ω de E (éventuellement un intervalle semi-ouvert ou fermé si $E = \mathbb{R}$); alors la fonction $x \to \frac{f(x+\vec{h})-f(x)-f'(x)\cdot\vec{h}}{\|\vec{h}\|}$ n'est évidemment pas définie sur Ω tout entier. On se bornera à considérer l'ensemble des valeurs de $\vec{h}$ qui sont majorées en norme par un nombre $\delta > 0$ fixe; alors les fonctions considérées seront toutes définies sur le même ensemble Ω_δ, à savoir l'ensemble (ouvert) des points de E dont la distance à $\complement\Omega$ est $> \delta$. Si la dérivée f' est uniformément continue, on en déduit que, dans l'ensemble Ω_δ, la fonction

$$x \longrightarrow \frac{f(x+\vec{h})-f(x)-f'(x)\cdot\vec{h}}{\|\vec{h}\|}$$ converge uniformément vers $\vec{0}$ lorsque $\vec{h} \neq \vec{0}$ tend vers $\vec{0}$.

Théorème 14 - Soit Ω un ouvert d'un espace affine normé E, soit c un point de Ω, et soit Ω_0 le complémentaire de c dans Ω. Soit f une application continue de Ω dans F, partout dérivable dans Ω_0. Si la dérivée $f'(x)$ tend vers une limite L dans $\mathcal{L}(\vec{E};\vec{F})$ lorsque $x \in \Omega_0$ tend vers c, alors f est dérivable au point c, et sa dérivée en c est l'application linéaire continue L.

Démonstration - Il suffit d'appliquer la majoration (III,5;8), pour $x = c$: en vertu des hypothèses, ω tend vers 0 quand $\vec{h}$ tend vers $\vec{0}$, d'où le résultat. On voit pourquoi, dans le théorème 13 et son corollaire 1, il est essentiel de ne pas supposer f dérivable aux extrémités de l'intervalle.

Dérivabilité totale et dérivabilité partielle

Théorème 15 - Soient E_1, E_2, F, des espaces affines normés, f une application d'un ouvert Ω de $E_1 \times E_2$ dans F. Pour que f soit continuement dérivable dans Ω il faut et il suffit qu'elle admette des fonctions dérivées partielles $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$, continues dans Ω.

Démonstration - Si f est dérivable dans Ω, le théorème 9 entraîne l'existence des dérivées partielles, et on a (III,3;29).

En outre, les inégalités (II,13;28) et (II,13;30) nous disent que, si une application linéaire continue u de $\vec{E}_1 \times \vec{E}_2$ dans $\vec{F}$ est représentée, suivant (II,13;26), par des applications linéaires continues u_1, u_2, de $\vec{E}_1$, $\vec{E}_2$ respectivement, dans $\vec{F}$, on a :

(III,5;10)
$$\begin{cases} \| u_1 \| \leq \| u \| , \quad \| u_2 \| \leq \| u \| , \\ \| u \| \leq \| u_1 \| + \| u_2 \| . \end{cases}$$

Alors, si f est continuement dérivable, lorsque x tend vers x_0, $\| f'(x) - f'(x_0) \|$ tend vers 0 , et la première inégalité (III,5;10) montre que $\left\| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(x_0) \right\| (i = 1,2)$ tend vers 0 , donc les dérivées partielles de f sont bien continues.

Inversement, supposons que f ait des dérivées partielles. On a vu que cela n'impliquait pas l'existence de la dérivée totale*(exemple (III,3;9)). Mais l'existence de dérivées partielles continues entraîne l'existence et la continuité de la dérivée totale (donc a fortiori la continuité de f).

Soient en effet $\vec{h}_1 \in \vec{E}_1$, $\vec{h}_2 \in \vec{E}_2$; si $a = (a_1, a_2) \in \Omega$, comme Ω est ouvert, il existe un nombre $\rho > 0$ tel que $\| \vec{h}_1 \| \leq \rho$, $\| \vec{h}_2 \| \leq \rho$ entraîne $(a_1 + \vec{h}_1 , a_2 + \vec{h}) \in \Omega$. Alors

(III,5;11)
$$\begin{aligned} &\overline{f(a_1 + \vec{h}_1 , a_2 + \vec{h}_2) - f(a_1 , a_2)} \\ &\overrightarrow{= f(a_1 + \vec{h}_1 , a_2 + \vec{h}_2) - f(a_1 + \vec{h}_1 , a_2)} \\ &\overline{+ f(a_1 + \vec{h}_1 , a_2) - f(a_1 , a_2)} \end{aligned}$$

Appliquons le corollaire 1 du théorème 13 a la fonction (continuement dérivable) $x_2 \to f(a_1 + \vec{h}_1 , x_2)$, avec $L = \frac{\partial f}{\partial x_2}(a_1, a_2)$.

(III,5;12)
$$\begin{aligned} &\overrightarrow{f(a_1 + \vec{h}_1 , a_2 + \vec{h}_2) - f(a_1 + \vec{h}_1 , a_2)} \\ &= \frac{\partial f}{\partial x_2}(a_1 , a_2) \cdot \vec{h}_2 + \vec{\omega} \, \| \vec{h}_2 \| \end{aligned}$$

* ni même la continuité de f .

avec la majoration

(III,5;13) $$\left\| \vec{\omega} \right\| \leq \sup_{0<\theta<1} \left\| \frac{\partial f}{\partial x_2}(a_1 + \vec{h}_1 \, , a_2 + \theta \vec{h}_2) - \frac{\partial f}{\partial x_2}(a_1 \, , a_2) \right\|$$

Si $\vec{h} = (\vec{h}_1 , \vec{h}_2)$ tend vers $\vec{0}$, cette expression tend vers 0 en vertu de la continuité de $\frac{\partial f}{\partial x_2}$ au point (a_1 , a_2).

D'autre part, la définition même de l'application dérivée, (III,3;13), montre que

(III,5;14) $$\overrightarrow{f(a_1 + \vec{h}_1 , a_2) - f(a_1 , a_2)} = \frac{\partial f}{\partial x_1}(a_1 , a_2) \vec{h}_1 + \vec{\alpha} \, \|\vec{h}_1\| ,$$

où $\vec{\alpha}$ tend vers $\vec{0}$ avec $\vec{h}_1$.

Alors finalement

(III,5;15) $$\left\{ \begin{aligned} & \overrightarrow{f(a_1 + \vec{h}_1 \, , a_2 + \vec{h}_2) - f(a_1 , a_2)} \\ & = \left(\frac{\partial f}{\partial x_1}(a_1 , a_2) \cdot \vec{h}_1 + \frac{\partial f}{\partial x_2}(a_1 , a_2) \cdot \vec{h}_2 \right) \\ & + \left(\vec{\alpha} \, \|\vec{h}_1\| + \vec{\omega} \, \|\vec{h}_2\| \right) \end{aligned} \right.$$

où $\vec{\alpha}$ et $\vec{\omega}$ tendent vers $\vec{0}$ avec $\vec{h}$.

La première parenthèse définit une application linéaire continue u de $\vec{E}_1 \times \vec{E}_2$ dans $\vec{F}$, par des applications linéaires continues u_1 et u_2 de $\vec{E}_1$ et $\vec{E}_2$ dans $\vec{F}$; la deuxième parenthèse est majorée en norme par $(\|\vec{\alpha}\| + \|\vec{\omega}\|) \|\vec{h}\|$, où $\|\vec{\alpha}\| + \|\vec{\omega}\|$ tend vers 0 avec $\|\vec{h}\|$. Donc f admet bien une dérivée totale $f'(a) \in \mathcal{L}(\vec{E}_1 \times \vec{E}_2 ; \vec{F})$ en a , définie par $\frac{\partial f}{\partial x_1}(a_1 , a_2) \in \mathcal{L}(\vec{E}_1 ; \vec{F})$ et $\frac{\partial f}{\partial x_2}(a_1 , a_2) \in \mathcal{L}(\vec{E}_2 ; \vec{F})$. Il reste à montrer que la continuité de $\frac{\partial f}{\partial x_1}$ et $\frac{\partial f}{\partial x_2}$ entraîne la continuité de f'. Si x tend vers x_0 dans Ω, $\left\| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(x_0) \right\|$ tend vers 0 ($i = 1,2$) , donc la deuxième inégalité (III,5;10) montre que $\| f'(x) - f'(x_0) \|$ tend vers 0 , et ceci achève la démonstration.

<u>Remarque</u> 1°/ La continuité de $\frac{\partial f}{\partial x_2}$ et <u>l'existence</u> seule de $\frac{\partial f}{\partial x_1}$ suffisent à entrainer l'existence de f' comme l'a montre le raisonnement (mais n'entrainent évidemment pas la continuité de f').

2°/ Le théorème que nous venons de voir pour un produit de 2 espaces F_1, F_2, est évidemment vrai pour un produit de n espaces $E_1, E_2, \ldots E_n$ (démonstration directe, ou par récurrence sur n). En outre la continuité de toutes les dérivées partielles, sauf peut-être une, et l'existence seule de cette dernière, suffisent à entrainer l'existence de la dérivée totale.

3°/ La même utilisation des inégalités (III,5;10) montre que, s'il s'agit de fonctions f dépendant d'un paramètre λ parcourant un espace topologique Λ, et si $\frac{\partial f}{\partial x_1}(x_1, x_2, \lambda)$, $\frac{\partial f}{\partial x_2}(x_1, x_2, \lambda)$, dépendent continuement de x_1, x_2, λ, il en est de même de $f'(x_1, x_2, \lambda)$.

<u>Corollaire 1</u> - <u>Soient E F des espaces affines, E de dimension finie, et supposons choisi dans E un référentiel $0, (\vec{e}_j)_{j\in J}$. Pour qu'une application f d'un ouvert Ω de E dans F soit continuement dérivable, il faut et il suffit qu'elle ait des dérivées partielles $\frac{\partial f}{\partial x_i}$ continues dans Ω.</u>

Car le référentiel identifie E à K^J.

C'est ce qui explique qu'en mathématiques spéciales, pour appliquer le théorème des fonctions composées (théoreme 11) on ait toujours supposé que les fonctions considérées avaient des dérivées partielles <u>continues</u>.

En réalité, cette hypothèse était trop forte puisqu'elle impliquait l'existence et la continuité de la dérivée totale; l'existence d'une dérivée totale est la seule chose à exiger. Mais la continuité des dérivées-partielles l'entraîne, alors que la seule existence des dérivées partielles ne l'entraine pas.

<u>Corollaire 2</u> - <u>Si, dans les conditions du corollaire 1, F est en outre de dimension finie, et qu'on y a choisi un référentiel $(f_i)_{i\in I}$, alors f, définie par les fonctions $F_i\,(i\in I)$ des variables $x_j\,(j\in J)$</u> *, <u>est continuement dérivable, si et seulement si les F_i ont des dérivées partielles usuelles $\frac{\partial F_i}{\partial x_j}$ continues.</u>

* notation de la formule (III,3;16).

§ 6 DÉRIVÉES D'ORDRE SUPÉRIEUR

Soit f une application d'un ouvert Ω d'un espace affine normé E dans un espace affine normé F. Si elle est partout dérivable, alors sa fonction dérivée f' est une application de Ω dans l'espace vectoriel normé $\mathcal{L}(\vec{E};\vec{F})$. Il est donc normal de chercher si cette application, à son tour, est dérivable. S'il en est bien ainsi, la dérivée en un point a, notée $f''(a)$, est un élément de $\mathcal{L}(\vec{E};\mathcal{L}(\vec{E};\vec{F}))$, et, s'il existe une dérivée partout, la fonction dérivée $f'' : x \longrightarrow f''(x)$ est une application de Ω dans $\mathcal{L}(\vec{E};\mathcal{L}(\vec{E};\vec{F}))$.

Soit alors $\vec{X}$ un vecteur de $\vec{E}$. Par définition de $f''(a)$, $f''(a)\cdot\vec{X}$ représente un élément de $\mathcal{L}(\vec{E};\vec{F})$, c'est-à-dire une application linéaire continue de $\vec{E}$ dans $\vec{F}$. Si alors $\vec{Y}$ est un autre vecteur de $\vec{E}$, on pourra parler de $(f''(a)\cdot\vec{X})\cdot\vec{Y}$, qui est un vecteur de $\vec{F}$. Nous allons donner une autre interprétation de ce vecteur.

Pour $\vec{Y}$ fixé, l'application $u \longrightarrow u\cdot\vec{Y}$ est linéaire continue de $\mathcal{L}(\vec{E};\vec{F})$ dans $\vec{F}$, d'après le théorème 54 du chapitre II; par suite, d'après le corollaire 1 du théorème 11, cette application permute avec la dérivation partielle $D_{\vec{X}}$ suivant le vecteur $\vec{X}$. Autrement dit :

(III,6;1) $$((D_{\vec{X}} f')(a))\cdot\vec{Y} = (D_{\vec{X}}(f'(x)\cdot\vec{Y}))_{x=a} \in \vec{F},$$

le second membre existant toutes les fois que le premier existe. Mais, si $f''(a)$ existe, $D_{\vec{X}} f'(a)$ n'est autre que $f''(a)\cdot\vec{X}$, donc le premier membre est le vecteur $(f''(a)\cdot\vec{X})\cdot\vec{Y}$ introduit plus haut; d'autre part $f'(x)\cdot\vec{Y}$ est aussi $\overrightarrow{D_{\vec{Y}} f(x)}$, donc le second membre existe et vaut $(\overrightarrow{D_{\vec{X}}(D_{\vec{Y}} f)})(a)$.

Finalement, si $f''(a)$ existe, $\overrightarrow{D_{\vec{X}}(D_{\vec{Y}} f)}(a)$ existe et on a :

(III,6;2) $$(f''(a)\cdot\vec{X})\cdot\vec{Y} = \overrightarrow{D_{\vec{X}}(D_{\vec{Y}} f)}(a) ;$$

le 2ème membre est aussi ce qu'on appelle la dérivée partielle du second ordre $\overrightarrow{D_{\vec{X}} D_{\vec{Y}} f}(a) = \overrightarrow{D^2_{\vec{X},\vec{Y}} f}(a)$ (étant entendu que, dans la dérivation partielle $D_{\vec{X}} D_{\vec{Y}}$, on fait d'abord la dérivation partielle $D_{\vec{Y}}$, ensuite la dérivation partielle $D_{\vec{X}}$)

Par ailleurs, nous avons vu au chapitre II, théorème 54 bis qu'on peut identifier une application linéaire continue de $\vec{E}$ dans $\mathcal{L}(\vec{E};\vec{F})$ avec une application bilinéaire continue de $\vec{E}\times\vec{E}$ dans $\vec{F}$. D'après la définition même de cette identification, cela revient à dire qu'à l'application linéaire continue U de $\vec{E}$ dans $\mathcal{L}(\vec{E};\vec{F})$, on fait correspondre l'application bilinéaire continue u de $\vec{E}\times\vec{E}$ dans $\vec{F}$ définie par :

$$u.(\vec{X},\vec{Y}) = (U.\vec{X}).\vec{Y} \quad (*) \tag{III,6;3}$$

C'est souvent sous cette forme que l'on considérera la dérivée seconde .

Dans ce cas, ce que nous notions $(f''(a).\vec{X}).\vec{Y}$ pourra s'écrire $f''(a).(\vec{X},\vec{Y})$, $f''(a)$ devient un élément de l'espace $\mathcal{L}_2(\vec{E},\vec{E};\vec{F})$ ou $\mathcal{L}_2(\vec{E}^2;\vec{F})$ des applications bilinéaires continues de $\vec{E}\times\vec{E}=\vec{E}^2$ dans F , et,si $f''(x)$ existe pour tout x de Ω , $f'':x \longrightarrow f''(x)$ sera une fonction définie sur $\Omega\subset E$ à valeurs dans $\mathcal{L}_2(\vec{E}^2;\vec{F})$.
(III ,6;2) devient

$$f''(a).(\vec{X},\vec{Y}) = \overrightarrow{D_{\vec{X}}\,D_{\vec{Y}}}f(a) \tag{III,6;3bis}$$

le second membre existant dès que le premier, c'est-à-dire $f''(a)$, existe (**)

* Il n'y a qu'une seule identification possible de $\mathcal{L}(\vec{E}_1;\mathcal{L}(\vec{E}_2;\vec{F}))$ avec $\mathcal{L}_2(\vec{E}_1,\vec{E}_2;\vec{F})$, si $\vec{E}_1\neq\vec{E}_2$. Comme ici $\vec{E}_1=\vec{E}_2=\vec{E}$, il y a deux identifications possibles,et nous avons choisi l'une d'elles

** Rappelons une fois encore que l'existence de $f''(a)$ implique l'existence de f' , sinon dans Ω tout entier, tout au moins dans tout un voisinage de a.

Définition. - On dit qu'une application bilinéaire d'un produit $\vec{E} \times E$ d'espaces vectoriels dans un espace vectoriel $\vec{F}$, est symétrique, si, quels que soient les éléments $\vec{X}$ et $\vec{Y}$ de $\vec{E}$, on a :

(III,6;4) $$u(\vec{X},\vec{Y}) = u(\vec{Y},\vec{X}).$$

Théorème 16 - Soit f une application d'un ouvert Ω d'un espace affine normé E dans un espace affine normé F . On suppose qu'elle admette une dérivée seconde $f''(a)$ en un point a de Ω. Alors $f''(a)$ est une application bilinéaire continue symétri[redacted] de $\vec{E} \times \vec{F}$ dans $\vec{F}$:

(III,6;5) $$f''(a).(\vec{X},\vec{Y}) = f''(a).(\vec{Y},\vec{X})$$

Si $E = \mathbb{K}^n$, en prenant pour $\vec{X}$, Y, des vecteurs de base, on retrouve la relation connue
$$\frac{\partial^2 \vec{f}}{\partial x_i \partial x_j}(a) = \frac{\partial^2 \vec{f}}{\partial x_j \partial x_i}(a)$$

Démonstration Nous considérerons l'expression :

(III,6;5bis) $$\vec{A} = \overrightarrow{f(a+t(\vec{X}+\vec{Y})) - f(a+t\vec{X}) - f(a+t\vec{Y}) + f(a)} \in \vec{F}.$$

On suppose $a \in \Omega$ fixé, ainsi que $\vec{X} \in \vec{E}$ et $\vec{Y} \in \vec{E}$ et on fera tendre t (scalaire $\neq 0$) vers 0 . Supposons aussi $\vec{X} \neq \vec{0}$, $\vec{Y} \neq \vec{0}$, sans quoi (III,6;5) est évidente. Comme Ω est ouvert, il contient une boule de centre a et de rayon $\rho > 0$ convenable; nous nous bornerons à prendre t tel que $|t| < \dfrac{\rho}{\|\vec{X}\| + \|\vec{Y}\|}$, de sorte que les points a, $a+t\vec{X}$, $a+t\vec{Y}$, $a+t(\vec{X}+\vec{Y})$ seront dans cette boule, donc dans Ω, ainsi que tout segment joignant deux de ces points par suite de la convexité de la boule, de sorte qu'on pourra appliquer la formule des accroissements finis (théorème 13).

L'expression $\vec{A}$ peut s'écrire comme différence :

(III,6;6) $$\vec{A} = \vec{g}(\vec{X}) - \vec{g}(\vec{0}),$$

si l'on définit l'application $\vec{g}$ de la boule $\{\vec{\xi}\,;\|\vec{\xi}\| < \|\vec{X}\|\}$ de $\vec{E}$ dans $\vec{F}$, par la formule :

(III,6;7) $$\vec{g}(\vec{\xi}) = \overrightarrow{f(a+t(\vec{\xi}+\vec{Y})) - f(a+t\vec{\xi})} \in \vec{F} \quad (a, \vec{Y}, t, \text{fixes})$$

La fonction g est dérivable, Sa dérivée est

(III,6;8) $$g'(\xi) = t f'(a+t(\vec{\xi}+\vec{Y})) - t f'(a+t\xi)$$

On peut lui appliquer le corollaire 1 du théorème 13; encore faut-il choisir L convenablement. Raisonnons intuitivement. On a

(III,6;9)
$$\vec{g}(\vec{X}) - \vec{g}(\vec{0}) \# g'(0)\cdot\vec{X}$$
$$\# \ t\left(f'(a+t\vec{Y}) - f'(a)\right)\cdot\vec{X} \# t^2\left(f''(a)\cdot\vec{Y}\right)\cdot\vec{X}$$

Nous allons donc prendre

(III,6;10)
$$L = t^2\left(f''(a)\cdot\vec{Y}\right) \in \mathcal{L}(\vec{E};\vec{F})$$

Alors la formule (III,5;8) donne

(III,6;11)
$$\vec{A} = \left(\vec{g}(\vec{X}) - \vec{g}(\vec{0})\right) = L\cdot\vec{X} + \vec{\omega}\,\|\vec{X}\|,$$

où $\vec{\omega} \in \vec{F}$ admet la majoration

(III,6;12)
$$\|\vec{\omega}\| \leq \sup_{\vec{\xi}\in]\vec{0},\vec{X}[} \|g'(\xi) - L\|$$
$$= \sup_{\vec{\xi}\in]\vec{0},\vec{X}[} \left\| t\,f'(a+t(\xi+Y)) - t\,f'(a+t\xi) - t^2(f''(a)\,\vec{Y})\right\|$$

Appliquons alors à f' la définition (III,3,13) de la dérivée :

(III,6;13)
$$\begin{cases} f'(a+t(\vec{\xi}+\vec{Y})) = f'(a) + f''(a)\cdot\left(t(\vec{\xi}+\vec{Y})\right) + \alpha\,\|t(\vec{\xi}+\vec{Y})\| \\ f'(a+t\vec{\xi}) = f'(a) + f''(a)\cdot t\vec{\xi} + \beta\,\|t\vec{\xi}\| \end{cases}$$

où, pour $a, \vec{X}$ et $\vec{Y}$ fixés, et $\vec{\xi} \in]\vec{0},\vec{X}[$ $\alpha = \alpha(t,\vec{\xi})$ et $\beta = \beta(t,\vec{\xi})$ tendent uniformément vers 0 dans $\mathcal{L}(\vec{E};\vec{F})$ quand t tend vers 0 . En portant dans (III,6;12), on obtient

(III,6;14)
$$\|\vec{\omega}\| \leq t^2 \sup_{\vec{\xi}\in]\vec{0},\vec{X}[} \left\|\alpha\cdot(\vec{\xi}+\vec{Y}) + \beta\cdot\vec{\xi}\right\| \leq \sup_{\vec{\xi}\in]\vec{0},\vec{X}[} t^2\left[(\|\alpha\|+\|\beta\|)(\|\vec{X}\|+\|\vec{Y}\|)\right]_1;$$

et le dernier crochet tend vers 0 avec t . Alors, compte tenu de la valeur de L, (III,6;11) montre que $\frac{\vec{A}}{t^2}$ tend vers $(f''(a)\cdot\vec{Y})\cdot\vec{X} = f''(a)\cdot(\vec{Y},\vec{X})$ quand t tend vers 0

Mais l'expression $\vec{A}$ est symétrique en $\vec{X}$ et $\vec{Y}$, donc aussi la limite de $\frac{\vec{A}}{t^2}$, ce qui démontre le théorème.

Remarques 1°/ Sous la forme modifiée

(III,6;17) $$\overrightarrow{D_{\vec{X}} D_{\vec{Y}} f}(a) = \overrightarrow{D_{\vec{X}} D_{\vec{Y}} f}(a) ,$$

on a une relation qui ne fait intervenir que la restriction de f à l'intersection de Ω avec la variété linéaire menée par a parallèlement aux vecteurs $\vec{X}$, $\vec{Y}$. Donc il est tout à fait inutile de supposer.$ deux fois dérivable : il suffit de supposer que la restriction de f à cette variété admet une dérivée seconde en a .

2°/ Il existe une variante de ce théorème, qui ne suppose pas que $f''(a)$ existe, mais qui lit que, si $\overrightarrow{D_{\vec{X}} D_{\vec{Y}} f}$ et $\overrightarrow{D_{\vec{X}} D_{\vec{Y}} f}$ existent dans tout un voisinage de a et sont continues en a alors elles sont égales en a . L'énoncé ainsi obtenu n'est ni plus fort ni moins fort que celui que nous donnons, il est différent : si nous prenons, pour simplifier $\vec{E} = \mathbb{R}^2$, l'existence de la dérivée seconde totale en un point implique l'existence de $\overrightarrow{\frac{\partial^2 f}{\partial x \partial y}}$ et de $\overrightarrow{\frac{\partial^2 f}{\partial y \partial x}}$ en ce point, mais non leur existence dans un voisinage ni leur continuitd'; inversement la continuité de ces dérivées partielles ne suffit pas a entrainer l'existence de la dérivée seconde totale (qui implique celle des dérivées partielles $\overrightarrow{\frac{\partial^2 f}{\partial x^2}}$, $\overrightarrow{\frac{\partial^2 f}{\partial y^2}}$). On sait par ailleurs que la seule hypothèse de l'existence des dérivées $\overrightarrow{\frac{\partial^2 f}{\partial x \partial y}}$, $\overrightarrow{\frac{\partial^2 f}{\partial y \partial x}}$ est insuffisante à entrainer leur égalité.

Dérivées successives

On définit de la même manière les dérivées suivantes. La dérivée d'ordre m en $a \in \Omega$ (qu'on ne définit que si les dérivées d'ordre $\leq m-1$ existent dans tout un voisinage de a) peut être identifiée à une application m- linéaire continue de $\vec{E}^m$ dans $\vec{F}$; si $f^{(m)}(x)$ existe pour tout x de Ω, $f^{(m)} : x \to f^{(m)}(x)$ est une application de Ω dans l'espace $\mathcal{L}_m(\vec{E}^m ; \vec{F})$ de ces applications m- linéaires continues. Plus précisément, $f^{(m)}$ se définit comme suit. Soit $f^{(m-1)}$ la dérivée d'ordre $m-1$, identifiée à une fonction sur Ω à valeurs dans $\mathcal{L}_{m-1}(\vec{E}^{m-1} ; \vec{F})$. Alors

$(f^{(m-1)})'(a) = f^{(m)}(a)$, sa dérivée en a est une application linéaire continue de $\vec{E}$ dans $\mathcal{L}_{m-1}(\vec{E}^{m-1} ; \vec{F})$, qui s'identifie à un élément de $\mathcal{L}_m(\vec{E}^m ; \vec{F})$, noté encore $f^{(m)}(a)$, par

(III,6;18) $$f^{(m)}(a).(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m) = ((f^{(m-1)})'(a).\vec{X}_1).(\vec{X}_2, \dots, \vec{X}_m).$$

On montre alors aussitôt, par récurrence sur m, que :

1°/ si $p \leqslant m$, et si on considère $f^{(p)}$ comme fonction sur Ω à valeurs dans $\mathcal{L}_p(\vec{E}^p\,;\vec{F})$, sa dérivée d'ordre $m-p$ en a, si elle existe, est un élément de $\mathcal{L}_{m-p}(\vec{E}^{m-p}\,;\mathcal{L}_p(\vec{E}^p;\vec{F}))$; cet élément peut s'identifier à un élément de $\mathcal{L}_m(\vec{E}^m;\vec{F})$, et il en est de même de $f^{(m)}(a)$, si elle existe; ces deux éléments existent en même temps l'un que l'autre, et

$$\text{(III,6;19)}\quad \left\{\begin{aligned} &(f^{(p)})^{(m-p)}(a) = f^{(m)}(a) \in \mathcal{L}_m(\vec{E}^m;\vec{F}) \\ &\left[\left((f^{(p)})^{(m-p)}(a)\right)\cdot(\vec{X}_1,\vec{X}_2,\dots,\vec{X}_{m-p})\right]\cdot(\vec{X}_{m-p+1},\dots,\vec{X}_m) \\ &\qquad = f^{(m)}(a)\cdot(\vec{X}_1,\vec{X}_2,\dots,\vec{X}_m) \in \vec{F} \end{aligned}\right.$$

2°/ Si $f^{(m)}(a)$ existe, alors $\overrightarrow{D_{\vec{X}_1} D_{\vec{X}_2} \dots D_{\vec{X}_m} f(a)} \in \vec{F}$ existe, et

$$\text{(III,6;20)}\quad f^{(m)}(a)\cdot(\vec{X}_1,\vec{X}_2,\dots,\vec{X}_m) = \overrightarrow{D_{\vec{X}_1} D_{\vec{X}_2} \dots D_{\vec{X}_m} f}(a) \in \vec{F}$$

<u>Définition</u> - On dit qu'une application m-linéaire u de $\vec{E}\times\vec{E}\times\dots\times\vec{E} = \vec{E}^m$ dans $\vec{F}$ est <u>symétrique</u>, si, quelle que soit la permutation $\sigma : k \longrightarrow \sigma_k$ de l'ensemble $\{1,2,\dots,m\}$, et le système dent vecteurs $\vec{X}_1,\vec{X}_2,\dots,\vec{X}_m$ de $\vec{E}$, on a la formule :

$$\text{(III,6;21)}\quad u(\vec{X}_{\sigma_1},\vec{X}_{\sigma_2},\dots,\vec{X}_{\sigma_m}) = u(\vec{X}_1,\vec{X}_2,\dots,\vec{X}_m)$$

Le théorème 16 s'étend alors aux dérivées d'ordre quelconque :

<u>Théorème 16 bis - Si une application f de $\Omega \subset E$ dans F a une dérivée d'ordre m en $a \in \Omega$, celle-ci, $f^{(m)}(a)$, est une application m-linéaire continue symétrique de $\vec{E}^m$ dans $\vec{F}$: pour toute permutation σ de $\{1,2,\dots,m\}$, ~~et tout~~ $(\vec{X}_1,\vec{X}_2,\dots,\vec{X}_m) \in \vec{E}^m$, on a :</u>

$$\text{(III,6;22)}\quad \left\{\begin{aligned} &f^{(m)}(a)\cdot(\vec{X}_1,\vec{X}_2,\dots,\vec{X}_m) = f^{(m)}(a)\cdot(\vec{X}_{\sigma_1},\vec{X}_{\sigma_2},\dots,\vec{X}_{\sigma_m}) \\ &\overrightarrow{D_{\vec{X}_1} D_{\vec{X}_2} \dots D_{\vec{X}_m} f}(a) = \overrightarrow{D_{\vec{X}_{\sigma_1}} D_{\vec{X}_{\sigma_2}} \dots D_{\vec{X}_{\sigma_m}} f}(a) \end{aligned}\right.$$

Démonstration

Le théoreme étant démontré pour les dérivées secondes, supposons-le démontré pour les dérivées d'ordre $\leq m-1$, et démontrons-le pour la dérivée d'ordre $m \geq 3$. Supposons d'abord que $\sigma_1 = 1$, et que σ échange donc entre eux les entiers $2,\ldots,m$. Alors

$\overrightarrow{D_{\vec{X}_2},\ldots, D_{\vec{X}_m} f}$ et $\overrightarrow{D_{\vec{X}_{\sigma_2}},\ldots,D_{\vec{X}_{\sigma_m}} f}$ sont des fonctions identiques (définies sur Ω, à valeurs dans $\vec{F}$), d'après l'hypothèse de récurrence. Donc leurs dérivées en a suivant le vecteur $\vec{X}_1 = \vec{X}_{\sigma_1}$ sont les mêmes, et on a bien (III,6;22)

Supposons maintenant que $\sigma_1 = 2$, $\sigma_2 = 1$, et que σ conserve chaque entier $3,\ldots m$. Alors

$\overrightarrow{D_{\vec{X}_3},\ldots D_{\vec{X}_m} f} = \overrightarrow{D_{\vec{X}_{\sigma_3}},\ldots,D_{\vec{X}_{\sigma_m}} f}$ est une fonction définie sur Ω admettant une dérivée première partout et une dérivée seconde en a. On peut donc lui appliquer le théorème 16 relativement aux dérivations suivant $\vec{X}_1$ et $\vec{X}_2$, et on aura encore (III,6;22).

Mais toute permutation de $\{1,2,\ldots,m\}$ est composée de permutationsde ces 2 types (c'est vrai si $\sigma_1 = 1$, sinon, on passe de $1,2,\ldots,m$ à $\sigma_1,\sigma_2,\ldots\sigma_m$, en passant d'abord de $1,2,\ldots m$ à $1,\sigma_1,\ldots,m$; puis de $1,\sigma_1,\ldots m$ à $\sigma_1, 1,\ldots m$; puis de $\sigma_1,1,\ldots m$ à $\sigma_1,\sigma_2,\ldots,\sigma_m$); le théorème est donc démontré.

Exercice. On pourra démontrer de proche en proche ce qui suit. Soit L une application m-linéaire continue symétrique de $\vec{E}^m$ dans $\vec{F}$. Elle définit une application p-linéaire continue L_p de $\vec{E}^p$ dans $\mathcal{L}_{m-p}(\vec{E}^{m-p};\vec{F})$; L_p est encore symétrique *

Alors la fonction dérivée $k^{\text{ième}}$ de la fonction "monome" $\vec{x} \longrightarrow L.\vec{x}^m$ de $\vec{E}$ dans $\vec{F}$ est la fonction

(III,6;22bis) $$\vec{x} \longrightarrow m(m-1)\ldots(m-k+1)\, L_{m-k}.\vec{x}^{m-k} \in \mathcal{L}_k(\vec{E}^k;\vec{F}).$$

* L_m et L_0 seront, par convention, L elle-même, et $\vec{x}^0$ sera $1 \in \mathbb{K}$.

Nous appelons $L_p.\vec{x}^p$ l'expression $L_p.\underbrace{(\vec{x},\vec{x},\ldots,\vec{x})}_{p}$

En prenant $E = F = \mathbb{K}$, on retrouve la formule élémentaire de dérivation des monomes.

Remarque - Le théorème 8 ter (linéarité de la dérivation), 8 quarto (dérivée d'une fonction à valeurs dans un produit), et le corollaire 1 du théorème 11 (permutabilité de la dérivation avec les applications linéaires continues) s'étendent immédiatement, par récurrence sur m, aux dérivées d'ordre m.

Cas où $E = \mathbb{K}^n$.

Supposons que E soit de dimension finie, ce qui est le cas le plus fréquent dans la pratique, et qu'on y ait choisi un référentiel 0; $(\vec{e}_i)_{i\in I}$. Alors on a, pour la dérivée seconde totale, la formule suivante, si $\vec{X} = \sum_{i\in I} X_i \vec{e}_i$, $\vec{Y} = \sum_{i\in I} Y_i \vec{e}_i$:

(III,6;23)
$$f''(a).(\vec{X},\vec{Y}) = \sum_{i,j\in I} X_i Y_j f''(a).(\vec{e}_i,\vec{e}_j) = \sum_{i,j\in I} X_i Y_j \overrightarrow{\frac{\partial^2 f}{\partial x_i \partial x_j}}(a) \in \vec{F}$$

Pour la dérivée 3ème, on aura

(III,6; 24)
$$f'''(a).(\vec{X},\vec{Y},\vec{Z}) = \sum_{i,j,k\in I} X_i Y_j Z_k \overrightarrow{\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}}(a)$$

Plus généralement, si $I = \{1,2,\dots n\}$:

(III,6 ; 25)
$$f^{(m)}(a).(\vec{X}_1,\vec{X}_2,\dots,\vec{X}_m) = \sum_{j} \overrightarrow{\frac{\partial^m f}{\partial x_{j_1} \partial x_{j_2} \cdots \partial x_{j_m}}}(a)\, X_{1,j_1} X_{2,j_2} \cdots X_{m,j_m} \in \vec{F},$$

où $X_{\alpha,\beta}$ est la $\beta^{\text{ième}}$ coordonnée de $\vec{X}_\alpha$ ($\alpha = 1,2,\dots,m$; β $1,2,\dots,n$), et où j parcourt l'ensemble de toutes les applications

$$\alpha \longrightarrow j_\alpha \text{ de } \{1,2,\dots,m\} \text{ dans } \{1,2,\dots,n\}$$

Si $E = \mathbb{K}$, corps des scalaires, le vecteur dérivé d'ordre m, $\overrightarrow{f^{(m)}(a)} \in \vec{F}$, défini page 192, est donc relié à l'application dérivée d'ordre m, $f^{(m)}(a) \in \mathcal{L}_m(\mathbb{K}^m; \vec{F})$, par

(III,6 ; 25bis)
$$\overrightarrow{f^{(m)}}(a) = f^{(m)}(a).(1,1,\dots,1)$$

On adopte des notations abrégées pour les dérivées partielles, en réunissant ensemble tous les $\frac{\partial}{\partial x_k}$ qui sont les mêmes. Soit $\vec{p}$ * un élément de $\mathbb{N}^n$, c'est-à-dire un système de n entiers ≥ 0, $p_1, p_2, \ldots, p_n$. Alors $D^{\vec{p}}$ ou $\left(\frac{\partial}{\partial x}\right)^{\vec{p}}$ désignera la dérivée partielle

(III,6;26) $D^{\vec{p}} = \left(\frac{\partial}{\partial x_1}\right)^{p_1} \left(\frac{\partial}{\partial x_2}\right)^{p_2} \cdots \left(\frac{\partial}{\partial x_n}\right)^{p_n}$.

Par exemple, si $n = 3$, $\vec{p} = (1,0,2)$, $D^{\vec{p}}$ est la dérivée $\frac{\partial}{\partial x_1}\left(\frac{\partial}{\partial x_3}\right)^2$ si $\vec{p} = \vec{0} = (0,0,\ldots,0)$, $D^{\vec{p}} f = f$. La quantité $|\vec{p}| = p_1 + p_2 + \cdots + p_n$ s'appelle l'ordre de $\vec{p}$ ou de la dérivation partielle. Si $\vec{q}$ est un autre élément de $\mathbb{N}^n$, on appellera $\vec{p} + \vec{q}$ l'élément $(p_1 + q_1, p_2 + q_2, \ldots, p_n + q_n)$, de sorte que $D^{\vec{p}+\vec{q}} f = D^{\vec{p}}(D^{\vec{q}} f)$. On dira que $\vec{p} \geq \vec{q}$ si $p_1 \geq q_1$, $p_2 \geq q_2, \ldots, p_n \geq q_n$ (c'est là une relatinn d'ordre dans $\mathbb{N}^n$); alors on appellera $\vec{p} - \vec{q}$ l'élément $(p_1 - q_1, p_2 - q_2, \ldots, p_n - q_n)$, et $\vec{p} = \vec{q} + (\vec{p} - \vec{q})$.

Pour des raisons qui apparaîtront plus loin, on est également amené à poser :

(III,6;27)
$$
\begin{cases}
\vec{p}\,! = p_1!\, p_2! \cdots p_n! \\[2ex]
\begin{pmatrix} \vec{p} \\ \vec{q} \end{pmatrix} = \begin{pmatrix} p_1 \\ q_1 \end{pmatrix} \begin{pmatrix} p_2 \\ q_2 \end{pmatrix} \cdots \begin{pmatrix} p_n \\ q_n \end{pmatrix}^{**} \quad \text{, pour } \vec{q} \leq \vec{p}. \\[2ex]
\vec{X}^{\vec{p}} = X_1^{p_1} X_2^{p_2} \cdots X_n^{p_n} \in \mathbb{K} \text{ , pour } \vec{X} \in \mathbb{K}^n,\ \vec{p} \in \mathbb{N}^n.
\end{cases}
$$

<u>Définition - On dit que f est m fois continuement dérivable, ou de classe C^m, dans Ω, si elle a des dérivées d'ordre $\leq m$ continues dans Ω. Toute applicationnt fois Qerivable dans Ω est au moins de classe C^{m-1}.</u>

* Bien que $\mathbb{N}^n$ ne soit pas un espace vectoriel, il est commode d'écrire $\vec{p}$ plutôt que p, pour rappeler que ce n'est pas un entier, mais un système de n entiers.

** Rappelons que $\begin{pmatrix} r \\ s \end{pmatrix} = C_r^s = \frac{r!}{s\,!\,(LA)!}$

Cas d' espaces produits. Dérivabilité totale et dérivabilité partielle

Le théorème **15** s'étend comme suit :

Théorème **17** - Pour qu'une application f d'un ouvert Ω ù'un produit $E_1 \times E_2 \times \dots \times E_n$ dans ~~F soit~~ de classe $\mathcal{C}^m$, il faut et il suffit qu'elle ait des dérivées partielles d'ordre $\leqslant m$ continues sur Ω .

Démonstration 1°/ !Jn élément u de $\mathcal{L}(E;F)$ étant équivalent à un système d'éléments u_i de $\mathcal{L}(E_i;F)$, $i = 1,2,\dots,n$,on met en correspondance biunivoque $\mathcal{L}(E_\cdot, F)$ et le produit des $\mathcal{L}(E_i;F)$. Cette correspondance respecte la structure vectorielle. D'autre part les inégalités $\|u_i\| \leqslant \|u\| \leqslant \sum_{j=1}^{n} \|u_j\|$, $i = 1,2,\dots,n$, montrent qu'à la norme de $\mathcal{L}(E;F)$ correspond ainsi, sur le produit des $\mathcal{L}(E_i;F)$, une norme équivalente à l'une de ses normes naturelles. Alors. d'après la remarque de la page **248** (extension-du théorème 8 quartoj, une application de $\Omega \subset E$ dans $\mathcal{L}(E;F)$ est de classe $\mathcal{C}^k$, si et seulement si les applications de Ω dans les $\mathcal{L}(\vec{E}_i;\vec{F})$ qui la définissent sont de classe $\mathcal{C}^k$.

2°/ Le théorème **17** étant alors vrai pour $m=1$ (théorème 15), démontrons -le par récurrence; supposons-le vrai pour $m-1$, démontrons-le pour $m \geqslant 2$. Soit f de classe $\mathcal{C}^m$; alors f' est de classe $\mathcal{C}^{m-1}$; comme f' est une application de Ω dans $\mathcal{L}(\vec{E};\vec{F})$ définie par les applications $\frac{\partial f}{\partial x_i}$ de Ω dans les $\mathcal{L}(\vec{E}_i;\vec{F})$, chaque $\frac{\partial f}{\partial x_i}$ est de classe $\mathcal{C}^{m-1}$ d'après 1°/); d'après l'hypothèse de récurrence, $\frac{\partial f}{\partial x_i}$ a donc des dérivées partielles d'ordre $\leqslant m-1$ continues sur Ω, et f a donc des dérivées partielles d'ordre $\leqslant m$ continues sur Ω . Inversement, si f a des dérivées partielles d'ordre $\leqslant m$ continues sur Ω , en particulier les $\frac{\partial f}{\partial x_i}$ existent et ont des dérivées partielles d'ordre $\leqslant m-1$ continues sur Ω. Alors, d'après l'hypothèse de récurrence, elles sont de classe $\mathcal{C}^{m-1}$.

En particulier, elles sont continues, donc f' existe d'après le théorème 15, et c'est la fonction sur Ω à valeurs dans $\mathcal{L}(\vec{E};\vec{F})$ définie par les fonctions $\frac{\partial f}{\partial x_i}$ sur Ω à valeurs dans les $\mathcal{L}(\vec{E}_i;\vec{F})$; comme les $\frac{\partial f}{\partial x_i}$ sont de classe $\mathcal{C}^{m-1}$, f' est de classe $\mathcal{C}^{m-1}$ d'après 1°, et f est bien de classe $\mathcal{C}^m$.

Espaces de fonctions m fois dérivables

Généralisant ce que nous avons Introduit à la fin du § 3, nous appellerons $(F^{\Omega})_{\ell;m}$ (resp. $(F^{\Omega})_{c\ell;m}$) l'espace des applications m fois dérivables (resp. m fois continuement dérivables), bornées ainsi que chacune de leurs dérivées d'ordre $\leqslant m$, de Ω dans F.

C'est un espace affine, d'espace vectoriel associé $(\vec{F}^{\Omega})_{\ell;m}$ (resp. $(\vec{F}^{\Omega})_{c\ell;m}$). On en fait un espace affine normé, en mettant sur l'espace vectoriel associé la norme :

$$(\text{III},6;27^{bis}) \quad \|\vec{f}\|_m = \operatorname*{Sup}_{\substack{x\in\Omega\\0\leqslant k\leqslant m}} \left(\| f^{(k)}(x) \| \right) ;$$

naturellement $\| f^{(k)}(x) \|$ est la norme dans $\mathcal{L}_k(\vec{E}^k;\vec{F})$.

On voit ici la raison d'être de l'indice 1 introduit dans $(F^{\Omega})_{\ell;1}$ ou $(F^{\Omega})_{c\ell;1}$. Avec ces notations, $(F^{\Omega})_{\ell}$ (resp. $(F^{\Omega})_{c\ell}$) peut aussi s'écrire $(F^{\Omega})_{\ell;0}$ (resp. $(F^{\Omega})_{c\ell;0}$) ; et $\|\vec{f}\|$ peut aussi s'écrire $\|\vec{f}\|_0$

<u>Théorème 17 bis</u> - <u>L'application qui, à chaque, fonction f, fait correspondre sa fonction derivee d'ordre $k \leqslant m$, est linéaire et continue, de norme $\leqslant 1$, de $(\vec{F}^{\Omega})_{\ell;m}$ (resp. $(\vec{F}^{\Omega})_{c\ell,m}$) dans $\left((\mathcal{L}_k(\vec{E}^k;\vec{F}))^{\Omega}\right)_{\ell;m-k}$ (resp. $\left((\mathcal{L}_k(\vec{E}^k;\vec{F}))^{\Omega}\right)_{c\ell;m-k}$)</u>

Malgré l'écriture un peu rébarbative des formules, c'est évident.

Très souvent, on met sur les espaces précédents d'autres normes équivalentes, par exemple on définit $\|\vec{f}\|_m$ par $\sum_{k=0}^{m}\left(\operatorname*{Sup}_{x\in\Omega}\| f^{(k)}(x)\|\right)$. Si E est de dimension finie n, et si on en a choisi un référentiel, on prend souvent la norme :

$$\|\vec{f}\|_m = \operatorname*{Sup}_{\substack{x\in\Omega\\|\vec{p}|\leqslant m}} \left\| \overrightarrow{D^{\vec{p}} f(x)} \right\| \quad \text{ou} \quad \sum_{|\vec{p}|\leqslant m} \left(\operatorname*{Sup}_{x\in\Omega} \left\| \overrightarrow{D^{\vec{p}} f(x)} \right\| \right)$$

Dérivées d'un produit (formule de Leibnitz)

Théorème 18 - Soient E un espace affine normé, F_1 F_2, G : des espaces vectoriels normes. Soit B une application bilinéaire continue de $\vec{F_1} \times \vec{F_2}$ dans $\vec{G}$. Si $\vec{u_1}$ (resp. $\vec{u_2}$) est une application m fois dérivable ou m fois continuement dérivable d'un ouvert Ω de E dans F_1 (resp. F_2), il en est de même de l'application $B(\vec{u_1}, \vec{u_2}) : x \to B(\vec{u_1}(x), \vec{u_2}(x))$ de Ω dans G, Les dérivées de cette fonction se calculent par la formule de Leibnitz.

Démonstration - Utilisons les notations du théorème 9 bis. La formule (III,4;22) s'écrit alors

$$(III,6;28) \quad (B(u_1, u_2))'(x) \cdot \vec{X} = B_{\vec{u_2}(x)} \cdot (u_1'(x) \cdot \vec{X}) + B_{\overrightarrow{u_1(x)}} \cdot (u_2'(x) \cdot \vec{X}) \in \vec{G}$$

on a donc aussi

$$(III,6;29) \quad (B(u_1, u_2))'(x) = B_{\overrightarrow{u_2(x)}} \circ u_1'(x) + B_{\overrightarrow{u_1(x)}} \circ u_2'(x) \in \mathcal{L}(\vec{E}; \vec{G})$$

Utilisons alors une récurrence sur m. Le théorème est démontré pour $m = 1$ (théorème 12); supposons-le démontré pour la dérivée d'ordre $m-1$, démontrons-le pour la dérivée d'ordre m. Soient donc $\vec{u_1}$ et $\vec{u_2}$ m fois dérivables, ou m fois continuement dérivables.

Alors $x \longrightarrow u_1'(x)$ est $m-1$ fois dérivable ou continuement dérivable, à valeurs dans $\mathcal{L}$ (Ë; $\vec{E_1}$). D'autre part, $x \longrightarrow B_{+-(\bar{x})}$, fonction définie sur Ω avaleurs dans $\mathcal{L}(\vec{F_1}; \vec{G})$, est composée de $x \longrightarrow \overrightarrow{u_2(x)}$, de Ω dans F_2, m fois dérivable ou continuement dérivable, et de $\vec{V_2} \longrightarrow B_{\vec{V_2}}$, application linéaire continue de $\vec{F_2}$ dans $\mathcal{L}(\vec{F_1}; \vec{G})$; d'après la permutabilité de la dérivation et des applications linéaires continues (corollaire 1 du théorème 11), $x \longrightarrow B_{\overrightarrow{u_2(x)}}$ est donc m fois dérivable ou continuement dérivable à valeurs dans $\mathcal{L}(\vec{F_1}; \vec{G})$. Comme alors $(U_1, U_2) \longrightarrow U_2 \circ U_1$ est une application bilinéaire continue de $\mathcal{L}(\vec{E}; \vec{F_1}) \times \mathcal{L}(\vec{F_1}; \vec{G})$ dans $\mathcal{L}(\vec{E}; \vec{G})$, l'hypothèse de récurrence nous permet d'affirmer que $x \longrightarrow B_{\overrightarrow{u_2(x)}} \circ u_1'(x)$ est $m-1$ fois dérivable ou continuement dérivable de Ω dans $\mathcal{L}(\vec{E}; \vec{G})$. Comme il en est de même pour

$x \longrightarrow B_{\overrightarrow{u_1(x)}} \circ u_2'(x)$, on voit que $(B(u_1, u_2))'$ est $m-1$ fois dérivable ou continuement dérlvable, donc $(B(u_1, u_2))$ l'est blenn fois * .

La formule de Leibnitz est d'une inutile complication dans le cas général. Nous ne la donnerons que dans le cas particulier où E est de dimension finie n, et muni d'un référentiel $\mathcal{O}, \vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$. Utilisons les notations de la formule (III,6;26). La formule de Leibnitz s'écrit alors :

$$D^{\vec{p}}\left(B(u_1, u_2)\right) = \sum_{\vec{q} \leq \vec{p}} \binom{\vec{p}}{\vec{q}} B\left(\overrightarrow{D^{\vec{q}} u_1}, \overrightarrow{D^{\vec{p}-\vec{q}} u_2}\right) \qquad \text{(III,6;36)}$$

Le cas le plus important est celui où $\vec{F}_1 = \vec{F}_2 = \vec{G} = \mathbb{K}$, corps des scalaires, et où $B(u_1, u_2) = u_1 u_2$, produit ordinaire. On obtient alors :

$$D^{\vec{p}} u_1 u_2 = \sum_{\vec{q} \leq \vec{p}} \binom{\vec{p}}{\vec{q}} D^{\vec{q}} u_1 \; D^{\vec{p}-\vec{q}} u_2 . \qquad \text{(III,6;36bis)}$$

Ici encore, l'intérêt des notations employées est que, dans le cas de n variables, les formules sont les mêmes que dans le cas de fonctions d'une variable.

La démonstration de cette formule est très simple. S'il n'y a qu'une variable $(n = 1)$, elle s'écrit

$$(u_1 u_2)^{(m)} = \sum_{k \leq m} \binom{m}{k} u_1^{(k)} \; u_2^{(m-k)} \qquad \text{(III,6;36ter)}$$

Elle est connue pour $m = 1$ et se prouve trivialement, dans le cas général, par récurrence sur m. Si en eff'et elle est supposée vraie pour $m-1$, on aura

* Dans les cas usuels, E est de dimension finie, et on peut en choisir un référentiel. Alors le théorème 17 nous dit que, pour montrer que $B(u_1, u_2)$ est de classe C^m lorsque u_1 et u_2 le sont, il suffit de montrer que ses dérivées partielles en $x_1, x_2, \dots, x_n$, d'ordre $\leq m$ sont continues. Or ceci est très simple par récurrence. C'est vrai pour $m = 1$, supposons-le démontré pour $m-1$, démontrons-le pour $m \geq 2$. On a $\frac{\partial}{\partial x_i}\left(B(\vec{u}_1, \vec{u}_2)\right) = B\left(\frac{\partial \vec{u}_1}{\partial x_i}, \vec{u}_2\right) + B\left(\vec{u}_1, \frac{\partial u_2}{\partial x_i}\right)$.
Mals $\overrightarrow{\frac{\partial u_1}{\partial x_i}}$ et $\vec{u}_2$ ont leurs dérivées partielles d'ordre $\leq m-1$ continues, donc aussi $B\left(\frac{\partial \vec{u}_1}{\partial x_i}, \vec{u}_2\right)$ d'après l'hypothèse de récurrence: de même pour $B\left(u_1, \frac{\partial u_2}{\partial x_i}\right)$; donc $B(u_1, u_2)$ a bien des dérivées partielles d'ordre $\leq m$ continues, d'où le résultat.

$$(u_1 u_2)^{(m)} = \left((u_1 u_2)^{(m-1)}\right)' = \left(\sum_{k' \leq m-1} \binom{m-1}{k'} u_1^{(k')} u_2^{(m-k'-1)}\right)'$$

$$= \sum_{k' \leq m-1} \left[\binom{m-1}{k'} u_1^{(k'+1)} u_2^{(m-k'-1)} + \binom{m-1}{k'} u_1^{(k')} u_2^{(m-k')}\right]$$

$$= \sum_{k \leq m} u_1^{(k)} u_2^{(m-k)} \left(\binom{m-1}{k-1} + \binom{m-1}{k}\right)$$

$$= \sum_{k \leq m} \binom{m}{k} u_1^{(k)} u_2^{(m-k)}$$

On fait ensuite une récurrence sur le nombren de variables. Supposons la formule démontrée pour $n-1$, démontrons la pour n .

$$D^{\vec{p}}(u_1 u_2) = \left(\frac{\partial}{\partial x_1}\right)^{p_1} D^{\vec{p}'}(u_1 u_2),$$

avec $\vec{p}' = (p_2, \dots, p_n) \in \mathbb{N}^{n-1}$.

Alors $D^{\vec{p}'}(u_1 u_2)$ est, pour x_1 fixé, une dérivée par rapport aux $(n-1)$ variables $x_2, \dots, x_n$; ensuite $\left(\frac{\partial}{\partial x_1}\right)^{p_1}$ est une dérivée par rapport a 1 variable.

$$D^{\vec{p}}(u_1 u_2) = \left(\frac{\partial}{\partial x_1}\right)^{p_1} \sum_{\vec{q}' \leq \vec{p}'} \binom{\vec{p}'}{\vec{q}'} D^{\vec{q}'} u_1 \, D^{\vec{p}'-\vec{q}'} u_2$$

$$= \sum_{q_1 \leq p_1} \sum_{\vec{q}' \leq \vec{p}'} \binom{p_1}{q_1} \binom{\vec{p}'}{\vec{q}'} \left(\frac{\partial}{\partial x_1}\right)^{q_1} \cdot D^{\vec{q}'} u_1$$

$$\left(\frac{\partial}{\partial x_1}\right)^{p_1 - q_1} D^{\vec{p}'-\vec{q}'} u_2$$

$$= \sum_{\vec{q} \leq \vec{p}} \binom{\vec{p}}{\vec{q}} D^{\vec{q}} u_1 \, D^{\vec{p}-\vec{q}} u_2$$

Théorème 9 - **Soient** E, F, G, trois espaces affines normés, Ω et Ω' des ouverts de E et F respectivement. Soit f une application de Ω dans Ω', g une application de Ω' dans G. Si f et g sont m fois dérivables ou m fois continuement dérivables, il en est de même de l'application composée $h = g \circ f$

Démonstration - Le théorème est vrai pour la dérivation d'ordre 1 (théorème 11 et son corollaire 5). Supposons-le vrai pour la dérivation d'ordre $\leqslant m-1$.-montrons-le pour la dérivation d'ordre m, $m \geqslant 2$.

Récrivons (III,4:1) :

$$h'(x) = g'(f(x)) \circ f'(x) \tag{III,6;37}$$

D'après l'hypothèse, f est m fois dérivable, donc f' l'est m -1 fois. D'autre part, f et g' sont toutes deux $m-1$ fois dérivables, donc, d'après l'hypothèse de récurrence, $g' \circ f : x \to g'(f(x))$ est aussi $m-1$ fois dérivable. Alors les fonctions $x \to f'(x)$ de Ω dans $\mathcal{L}(\vec{E};\vec{F})$ et $x \to g'(f(x))$ de Ω dans $\mathcal{L}(\vec{F};\vec{G})$ sont toutes les deux $m-1$ fois dérivables. et $(u,v) \longrightarrow v \circ u$ est une application bilinéaire continue de $\mathcal{L}(\vec{E};\vec{F}) \times \mathcal{L}(\vec{E};\vec{G})$ dans $\mathcal{L}(\vec{E};\vec{G})$ (théorème 54 du chapitre II); le théorème 18 nous dit donc que $x \to g'(f(x)) \circ f'(x)$ est aussi $m-1$ fois dérivable. Ainsi h' est $m-1$ fois dérivable, et cela signifie bien que h est m fois dérivable. Même démonstration pour la continue dérivabilité !

Remarques 1°/ On aurait pu démontrer les théorèmes 18 et 19 par une seule récurrence.

En effet le théorème 19 pour l'entier m n'utilise le théorème 18 que pour l'entier $m-1$. D'autre part, si on veut prouver le théorème 18 relativement à $m-1$, la fonction $x \longrightarrow B(u_1(x),u_2(x))$ est la composée de $x \to (u_1(x),u_2(x))$, qui est de classe $\mathcal{C}^{m-1}$, et de $(u_1,u_2) \to B(u_1,u_2)$, qui est bilinéaire continue donc Indéfiniment dérivable, donc on peut utiliser le theorème 19 pour l'entier $m-1$.

Or les théorèmes 18 et 19 sont prouvés pour $m = 1$. On en déduit 19 pour $m = 2$, donc 18 pour $m = 2$, donc 19 pour $m = 3$, donc 18 pour $m = 3$, etc....

2°/ Le calcul des dérivées successives de $h = g \circ f$ peut se faire avec des formules explicites, mais il est assez compliqué. Par exemple, si $\vec{X}$ et $\vec{Y}$ sont des vecteurs de $\vec{E}$, on trouve

$$(\text{III},6;38)\quad h''(x)\cdot(\vec{X},\vec{Y}) = g''(f(x))\cdot(f'(x)\cdot\vec{X}, f'(x)\cdot\vec{Y}) + g'(f(x))\cdot(f''(x)\cdot(\vec{X},\vec{Y}))\in\vec{G}.$$

Si E, F, G sont le corps des scalaires $\mathbb{K}$, et si on prend $X = Y = 1$, on trouve la formule élémentaire

$$(\text{III},6;39)\quad h''(x) = g''(f(x))\, f'^2(x) + g'(f(x))\, f''(x).$$

Il est souvent utile de connaître l'expression de la dérivée d'ordre m de $h = g\circ f$ dans ce cas particulier $E = F = G = \mathbb{K}$: nous la donnerons au théorème 21 ter.

§ 7 FORMULE DE TAYLOR. MAXIMA ET MINIMA

Du moment que la dérivée d'ordre m, $f^{(m)}(a)$, est une application $m$linéaire, il est possible de calculer sa valeur sur le système $(\vec{X},\vec{X},\ldots\vec{X})\in\vec{E}^m$, défini à partir d'un même vecteur $\vec{X}$ de $\vec{E}$. Il sera commode de représenter l'expression $f^{(m)}(a)\cdot(\vec{X},\vec{X},\ldots\vec{X})$ par le symbole abrégé :

$$(\text{III},7;1)\quad f^{(m)}(a)\cdot\vec{X}^m = f^{(m)}(a)\cdot(\vec{X},\vec{X},\ldots\vec{X}).$$

Il y a alors autant de formules de Taylor qu'il y a de formules des accroissements finis.

Au théorème 13 A, correspond une formule de Taylor qui étend le théorème 6 :

Théorème 20 A - soit f une fonction réelle m fois dérivable sur l'ouvert $\Omega\subset E$. Supposons que ~~le~~ segment fermé $[x, x+\vec{h}]$ soit tout entier dans Ω, et que f admette une dérivée d'ordre $m+1$ en tout point du segment ouvert $]x, x+\vec{h}[$. Alors on a la formule :

$$(\text{III},7;1\,bis)\quad f(x+\vec{h}) = f(x) + \frac{f'(x)}{1!}\cdot\vec{h} + \frac{f''(x)}{2!}\cdot\vec{h}^2 + \cdots + \frac{f^{(m)}(x)}{m!}\cdot\vec{h}^m + \frac{f^{(m+1)}(x+\theta h)}{(m+1)!}\cdot\vec{h}^{m+1},$$

où θ est un nombre réel, $0<\theta<1$.

Démonstration - Appelons Φ la fonction réelle sur $[0,1]$ définie par la formule (III,5:). On voit aussitôt, par récurrence sur k, que sa dérivée d'ordre $k \leqslant m$ existe sur $[0,1]$, et sa dérivée d'ordre $m+1$ sur $]0,1[$, avec :

(III,7;1 ter) $$\Phi^{(k)}(t) = f^{(k)}(x+t\vec{h}).\vec{h}^{k} .$$

Il suffit alors de lui appliquer le théorème 6 (formule de Taylor pour une fonction réelle d'une variable réelle) pour l'intervalle $[0,1]$, et l'on obtient le résultat.

Ensuite, au théorème 13, correspond la formule de Taylor suivante :

Théorème 20 - Soit f une application m fois dérivable de l'ouvert $\Omega \subset E$ dans F. Supposons que le segment fermé $[x, x+\vec{h}]$ soit tout entier dans Ω, et que f admette une dérivée d'ordre $m+1$ en tout point du segment ouvert $]x, x+\vec{h}[$, majorée en norme par M ; alors on a la formule:

(III,7;2) $$\|\vec{A}\| = \left\|\overrightarrow{f(x+\vec{h})-f(x)} - f'(x).\vec{h} - \frac{f''(x)}{2!}.\vec{h}^{2} \dots - \frac{f^{(m)}(x)}{m!}.\vec{h}^{m}\right\| \leqslant \frac{M\|\vec{h}\|^{m+1}}{(m+1)!}.$$

Démonstration - Le théorème a déjà été démontré pour $m=0$ au théoreme 13, nous allons donc faire une récurrence. Supposons-le démontré pour la formule de Taylor d'ordre $m-1$, et démontrons-le pour la formule de Taylor d'ordre $m \geqslant 1$.

• Considérons la fonction $\vec{g}$ définie par

(III,7;3) $$\overrightarrow{g(t)} = \overrightarrow{f(x+t\vec{h})-f(x)} - f'(x).t\vec{h} \dots - \frac{f^{(m)}(x)}{m!}(t\vec{h})^{m}$$

C'est une application du segment fermé $[0,1]$ de $\mathbb{R}$ dans $\vec{F}$, et $\vec{A} = \overrightarrow{g(1)-g(0)}$. Elle est continue sur ce segment, et en outre, elle admet une dérivée première continue sur tout ce segment, donnée par :

(III,7;4) $$\overrightarrow{g'(t)} = f'(x+t\vec{h}).\vec{h} - f'(x).\vec{h} \dots - \frac{t^{m-1}}{(m-1)!} f^{(m)}(x).\vec{h}^{m} .$$

Mais, d'après la 2ème formule (III,6;19) :

(III,7;5) $$f^{(k)}(x).\vec{h}^{k} = \left((f')^{(k-1)}(x)\cdot\vec{h}^{k-1}\right)\cdot\vec{h}$$

(III,7;4) peut donc s'écrire :

(III,7;6) $$g'(t) = \left[f'(x+t\vec{h}) - f'(x) - \dots - \frac{(f')^{(m-1)}}{(m-1)!}(x).(t\vec{h})^{m-1}\right].\vec{h} \in \vec{F}$$

où le crochet est un éldment de $\mathcal{L}(\vec{E};\vec{F})$; appliqué à $\vec{h} \in \vec{E}$, il donne bien un élément de $\vec{F}$.

Cherchons une majoration de cette dérivée. Nous voyons que le crochet n est pas autre chose que l'expression de Taylor analogue a $\vec{A}$, mais relative à l'accroissement $t\vec{h}$ et à l'ordre $m-1$, et appliquée à la fonction f' , définie sur Ω à valeurs dans $\mathcal{L}(\vec{E};\vec{F})$. Grâce à l'hypothèse de récurrence, et compte tenu de ce que $(f')^{(m)} = f^{(m+1)}$, nous pouvons donc écrire la majoration valable pour $0 \leqslant t \leqslant 1$:

$$(\text{III},7;7) \qquad \|\overrightarrow{g'(t)}\| \leqslant M \frac{\|t\vec{h}\|^m}{m!} \|\vec{h}\| = M\|\vec{h}\|^{m+1} \frac{t^m}{m!} .$$

Utilisant alors le lemme démontré, à l'occasion du théorème 13, (en y remplaçant $f(x)$ par $g(x)$ et $g(x)$ par $M\|\vec{h}\|^{m+1}\frac{x^{m+1}}{(m+1)!}$), nous voyons que la 'onction g , dans l'intervalle $[0,1]$, vérifie la majoratinn :

$$(\text{III},7;8) \qquad \|\overrightarrow{g(1)-g(0)}\| \leqslant M\frac{\|\vec{h}\|^{m+1}}{(m+1)!} ,$$

et par conséquent nous avons bien la formule III,7,2).

Au corollaire du tnéor me 13 correspond ici le suivant :

<u>Corollaire. Si</u>, <u>dans les conditions de l'énoncé du théorème, L est une application</u> $(m+1)$-<u>linéaire continue de</u> $\vec{E}^{m+1}$ <u>dans</u> $\vec{F}$, <u>on a la formule</u> :

$$(\text{III},7;9) \qquad \Big\|\overrightarrow{f(x+\vec{h}) - f(x)} - f'(x).\vec{h} - \frac{f''(x)}{2!}.\vec{h}^2 - \dots - \frac{f^{(m)}(x)}{m!}.\vec{h}^m - \frac{L}{(m+1)!}.\vec{h}^{m+1}\Big\| \leqslant \omega\frac{\|\vec{h}\|^{m+1}}{(m+1)!} ,$$

<u>où</u> ω <u>est la borne supérieure de</u> $\|f^{(m+1)}(\xi) - L\|$, <u>lorsque</u> ξ <u>parcourt le segment</u> $]x, x+\vec{h}[$.

<u>Démonstration</u> - On se ramène facilement au cas où L est une forme $(m+1)$- linéaire symétrique, qui est d'ailleurs le seul Intéressant. On applique alors le théorème à la fonction $\xi \longrightarrow f(\xi) - \frac{L}{(m+1)!}.(\overrightarrow{\xi - x})^{m+1}$, dont les dérivées se calculent par (III,6;22 bis)

Pour terminer, remarquons qu'il existe une formule de Taylor qui correspond à la définition de la dérivée (III,3; 13) et qui peut s'appeler formule de Taylor pour des accroissements Infiniment petits. Il est plus commode ici de remplacer $m+1$ par m :

<u>Théorème 21</u> - <u>Si</u> f <u>a des dérivées d'ordre</u> $\leqslant m-1$ <u>dans</u> Ω <u>et une dérivée d'ordre</u> m en a, on a :

$$(\mathrm{III},7;10)\qquad f(a+\vec{h}) = f(a) + f'(a).\vec{h} + \dots + \frac{f^{(m)}(a)}{m!}\cdot\vec{h}^m + \vec{\alpha}\,\|\vec{h}\|^m,$$

où $\vec{\alpha}$ tend vers $\vec{0}$ avec $\vec{h}$.

Démontrons ce résultat encore par récurrence sur m; il est vrai pour $m = 1$, supposons-le **vrai** pour l'entier $m-1$, et montrons le pour l'entier $m \geqslant 2$. Alors la fonction de $\vec{\xi}$

$$(\mathrm{III},7;11)\qquad \overrightarrow{g(\vec{\xi})} = \overrightarrow{f(a+\vec{\xi}) - f(a)} - f'(a)\cdot\vec{\xi}\dots - \frac{f^{(m)}(a)}{m!}\cdot\vec{\xi}^m$$

est définie au voisinage de $\vec{0}$ dans $\vec{E}$, à valeurs dans? , et dérlvable. Sa dérivée est, d'aprés (**III**,6:22 **bis**)

$$(\mathrm{III},7;12)\qquad g'(\vec{\xi}) = f'(a+\vec{\xi}) - f'(a) - \dots - \frac{(f')^{(m-1)}(a)}{(m-1)!}\cdot\vec{\xi}^{m-1} \in \mathcal{L}(\vec{E};\vec{F}).$$

Appliquons à g' la !'ormule (**III**,7:10) pour l'entlerm-1:

$$(\mathrm{III},7;12\,bis)\qquad g'(\xi) = \overrightarrow{\beta(\xi)}\,\|\vec{\xi}\|^{m-1},$$

où $\vec{\beta}(\vec{\xi})$ tend **vers** $\vec{0}$ **avec** $\vec{\xi}$.

Alors la formule des accroissements finis (théorème 13) donne :

$$(\mathrm{III},7;12\,ter)\qquad \|\overrightarrow{g(\vec{h})}\| = \|\overrightarrow{g(\vec{h})} - \overrightarrow{g(\vec{0})}\| \leqslant \|\vec{h}\| \sup_{\vec{\xi}\in]\vec{0},\vec{h}[} \|\overrightarrow{g'(\vec{\xi})}\|$$

$$\leqslant \|\vec{h}\| \sup_{\vec{\xi}\in]\vec{0},\vec{h}[} \|\overrightarrow{\beta(\vec{\xi})}\|\,\|\vec{\xi}\|^{m-1} \leqslant \varepsilon\,\|\vec{h}\|^m,$$

où ε tend vers o avec $\|\vec{h}\|$, ce **qui** donne le résultat voulu pour l'entier-.

Application de la formule de Taylor au calcul de dérivées de fonctions

Il arrive qu'il **soit** plus facile de trouver un développement de Taylor que de calculer **des** dérivées successives; alors c'est ce développement qui donne les dérivées, autrement **dit** on a la réciproque du théorème 21 :

Théorème 21 **bis** - SI f est une applicationmfols **dérivable** de l'ouvert Ω de $\vec{E}$ **dans** $\vec{F}$, et si on a trouvé des applications L_k , k **linéaires continues symétriques de** $\vec{E}^*$ **dans** $\vec{F}$, $k = 1,2\dots m$, et un élément L_0 de $\vec{F}$, **tels que** l'on **ait** :

(III,7;12quarto) $$f(a+\vec{h}) = L_0 + L_1 \cdot \vec{h} + \frac{L_2}{2!} \cdot \vec{h}^2 + \dots + \frac{L_m}{m!} \vec{h}^m + \vec{\alpha} \|\vec{h}\|^m \quad ,$$

où $\vec{\alpha} = \vec{\alpha}(\vec{h})$ tend vers $\dot{0}$ avec $\vec{h}$, alors on a **nécessairement** $L_k = f^{(k)}(a)$, $k = 1, 2, \dots m$, et $L_0 = f(a)$.

Démonstration . Il résulte du **théorème** 21 que si nous posons

(III,7;12quinto) $$A_k = f^{(k)}(a) - L_k \in \mathcal{L}_k(\vec{E}^k; \vec{F}) , \quad \vec{A}_0 = \overrightarrow{f(a) - L_0} \in \vec{F} \quad ,$$

on a

(III,7,12sexto) $$\sum_{k=1}^{m} \frac{A_k}{k!} \cdot \vec{h}^k = \vec{\beta} \|\vec{h}\|^m \quad ,$$

où $\vec{\beta}$ tend **vers** $\vec{0}$ avec $\vec{h}$.

$\vec{A}_0$ est nul, **comme** on le voit en faisant tendre h vers $\acute{0}$. Supposons démontré que $A_0, A_1, A_2, \dots A_{k-1}$ sont nuls, $k \leq m$, . montrons **que** A_k est nul; il en **résultera** que tous les A_k , $k = 0, 1, 2, \dots m$, **seront** nuls, et le **théorème** sera démontré. Posons $\vec{h} = t\vec{X}$, $\vec{X}$ fixé, t scalaire qu'on fera tendre vers 0 . Alors

(III,7;12 octavo) $$t^k\left(\frac{A_k}{k!} \cdot \vec{X}^k + t \sum_{l=k+1}^{m} \frac{A_l}{l!} \cdot t^{l-k-1} \vec{X}^l\right) ,$$

est infiniment petit devant t^m donc devant t^k quand t tend vers 0 : donc son **quotient par** t^k tend vers 0 , donc $A_k \cdot \vec{X}^k = 0$, quel que soit $\vec{X} \in \vec{E}$. Alors, la démonstration par récurrence sera achevée et le théorème démontré, quand nous aurons démontré le **lemme** :

Lemme - Soit **A** une **application** k-**linéaire** symétrique de $\vec{E}^k$ **dans** $\vec{F}$. Si, pour tout $\vec{X}$ de $\vec{E}$, on a :

(III,7;13) $$A \cdot \vec{X}^k = \vec{0} \in \vec{F} \quad ,$$

alors A est nulle.

Ce **lemme est** bien connu pour $m = 2$, car-alors on a (**à** cause de **la symétrie** de A), pour $\vec{X} \in \vec{E}$, $Y \in \vec{E}$:

$$A.(\vec{X} + \vec{Y}, \vec{X} + \vec{Y}) - \mathbf{A(?,?)} - A(\vec{Y},\vec{Y}) = 4A(\vec{X},\vec{Y}),$$

et l'hypothèse (III,7;13) **entraîne** la nullité du premier membre, donc du deuxième, donc de A . Etendons cette **démonstration** au cas général.

Considérons, pour $\vec{X}_1, \vec{X}_2, \ldots \vec{X}_k$ fixés, la fonction h de $t = (t_1, t_2, \ldots t_k) \in \mathbb{R}^k$

$$h(t) = A \cdot \left(t_1 \vec{X}_1 + t_2 \vec{X}_2 + \cdots + t_k \vec{X}_k\right)^k . \tag{III,7;13 bis}$$

Par hypothèse elle est identiquement nulle. Or c'est un polynôme en $t_1, t_2, \ldots t_k$; donc chacun des coefficients de ce polynôme est nul. Mais, d'après la symétrie de A, le coefficient de $t_1 t_2 \ldots t_k$ n'est autre que $k!\, A(\vec{X}_1, \vec{X}_2, \ldots \vec{X}_k)$, ce qui démontre le lemme.

Remarque - Nous avons dû supposer f dérivable jusqu'à l'ordre m ; c'est essentiel. Le fait que f admette au voisinage de a un développement du type (III,7;12 quarto) n'entraîne absolument pas qu'elle ait de dérivée d'ordre > 1. Par exemple la fonction réelle de variable réelle f, définie par $f(0) = 0$, $f(x) = x^{m+1} \sin \frac{1}{x^{m+1}}$ pour $x \neq 0$, admet un tel développement à l'origine avec $L_k = 0$. Elle admet des dérivées de tout ordre dans le complémentaire de l'origine; elle a une dérivée première nulle à l'origine, puisque $|f(x) - f(0)| = |f(x)| \leq |x^{m+1}|$. Mais sa dérivée première en $x \neq 0$, égale à $(m+1) x^m \sin \frac{1}{x^{m+1}} - \frac{(m+1)}{x} \cos \frac{1}{x^{m+1}}$, prend des valeurs arbitrairement grandes dans tout voisinage de 0. Donc la dérivée première est discontinue à l'origine, et il n'y a surement pas de dérivée seconde.

Voici une application du théorème 21 bis au calcul des dérivées successives d'une fonction composée.

Théorème 21 ter - Si g et f sont des fonctions scalaires d'une variable scalaire, m fois dérivables, la dérivée d'ordre m de $h = g \circ f$ au point a est donnée par la formule suivante, où $f^{(p)}$ veut dire $f^{(p)}(a)$, et $g^{(q)}$ veut dire $g^{(q)}(f(a))$:

$$h^{(m)}(a) = \sum_{k_1 + 2k_2 + \cdots + mk_m = m} \frac{m!}{k_1!\, k_2! \ldots k_m!\, (1!)^{k_1} (2!)^{k_2} (3!)^{k_3} \ldots (m!)^{k_m}} \times g^{(k_1 + k_2 + \cdots + k_m)} (f')^{k_1} (f'')^{k_2} \ldots (f^{(m)})^{k_m} . \tag{III,7;13 ter}$$

Démonstration - Utilisons le développement de Taylor. On a

$$h(x) = g(f(x)) = \sum_{0 \leq \ell \leq m} \frac{(f(x) - f(a))^\ell}{\ell!} g^{(\ell)} + R_m . \tag{III,7;13 quarto}$$

Mais

(III,7;13 quinto) $$f(x) - f(a) = \sum_{1 \le k \le m} \frac{(x-a)^k}{k!} f^{(k)} + \rho_m .$$

On a donc, à des termes pris qui sont infiniment petits devant $(x-a)^m$ quand x tend vers a :

(III,7;13 sexto) $$\left(f(x) - f(a)\right)^\ell \# \left(\sum_{1 \le k \le m} \frac{(x-a)^k}{k!} f^{(k)}\right)^\ell .$$

La formule du développement de la puissance ℓ-ième d'une somme est parfois donnée en Mathématiques Spéciales; elle généralise le développement du binôme de Newton. De toute façon elle sera donnée à la formule (III,7 26). On a donc :

(III,7;13 septimo) $$\left(f(x) - f(a)\right)^\ell = \sum_{k_1+k_2+\cdots+k_m=\ell} \frac{\ell!}{k_1!\,k_2!\ldots k_m!\,(1!)^{k_1}(2!)^{k_2}\ldots(m!)^{k_m}} (x-a)^{k_1+2k_2+\cdots+mk_m} (f')^{k_1}(f'')^{k_2}\ldots(f^{(m)})^{k_m} .$$

On en déduit $\dfrac{\left(f(x)-f(a)\right)^\ell}{\ell!}$, d'où l'expression de $h(x)$, à des infiniment petits près devant $(x-a)^m$ quand x tend vers a :

(III,7,13 octavo) $$h(x) \# \sum_{\substack{0 \le \ell \le m \\ k_1+k_2+\cdots+k_m=\ell}} \frac{(x-a)^{k_1+2k_2+\cdots+mk_m}}{k_1!\,k_2!\ldots k_m!\,(1!)^{k_1}(2!)^{k_2}\ldots(m!)^{k_m}} \times (f')^{k_1}(f'')^{k_2}\ldots(f^{(m)})^{k_m} \times g^{(\ell)} .$$

Comme on sait que h est m fois dérivable (théorème 19), la quantité $h^{(m)}(a)$ ne peut être que le produit par $m!$ du coefficient de $(x-a)^m$ dans le développement précédent, d'où le résultat.

On voit, sur cet exemple, le grand rôle que peut jouer le développement de Taylor pour calculer des dérivées d'ordre > 1, à condition d'avoir démontré, par une autre méthode, l'existence des dérivées à calculer.

Formule de Taylor par rapport à un système de coordonnées

Théorème 21 quarto

L'expression $\displaystyle\sum_{k=0}^{m} \frac{f^{(k)}(x)}{k!} \cdot \vec{h}^k$ qui Intervient dans les formules de Taylor s'écrit, si E est de dimension finie et qu'on y a choisi un **référentiel,** si $\vec{h}$ a les **coordonnées** $h_1, h_2, \ldots h_n$, et si on utilise **la notation (III,6; 26-27)** :

$$(\text{III},7;14) \qquad \sum_{k=0}^{m} \frac{f^{(k)}(x)}{k!} \cdot \vec{h}^{k} = \sum_{|\vec{p}|\leq m} \frac{\overrightarrow{D^{\vec{p}} f(x)}}{\vec{p}\,!}\, \vec{h}^{\vec{p}} \in F$$

Démonstration - Nous avons vu à la formule (III,6;25) que l'on a :

$$(\text{III},7;14\,bis) \qquad \frac{f^{(k)}(x)}{k!} \cdot \vec{h}^{k} = \frac{1}{k!} \sum_{j_1, j_2, \cdots j_k} \frac{\overrightarrow{\partial^k f}}{\partial x_{j_1} \partial x_{j_2} \cdots \partial x_{j_k}}(x)\, h_{j_1} h_{j_2} \ldots h_{j_k} \,{}^{*} \in \vec{F}.$$

Dans cette formule, $j_1, j_2, \ldots, j_k$ sont des indices qui prennent, indépendamment les uns des autres, toutes les valeurs possibles, parmi les entiers 1 , 2 , --- , n . Réunissons entre elles toutes les suites finies $j_1, j_2, \cdots, j_k$, pour lesquelles p_1 des indices sont égaux à 1, p_2 des indices égaux à 2 ,..., p_n des indices égaux à n .

Elles donnent, dans la somme (III,7;13), le **même** élément, qui **s'écrit,** avec la notation (III,6;26 et 27).

$$(\text{III},7;14\,ter) \qquad \frac{\overrightarrow{D^{\vec{p}} f(x)}\, \vec{h}^{\vec{p}}}{k!} = \overrightarrow{\left(\frac{\partial}{\partial x_1}\right)^{p_1} \left(\frac{\partial}{\partial x_2}\right)^{p_2} \cdots \left(\frac{\partial}{\partial x_n}\right)^{p_n}} f(x)\, \frac{h_1^{p_1} h_2^{p_2} \ldots h_n^{p_n}}{k!} \in \vec{F}.$$

Combien y a-t-il de termes de ce type ? Autant qu'il y a d'applications j **de** $\{1, 2, \ldots k\}$ **dans** $\{1, 2, \ldots n\}$ **, où** $\{1\}$ **a une image réciproque à** p_1 **éléments,** $\{2\}$ **une image réciproque à** p_2 **éléments ,......**$\{n\}$ **une image réciproque à** p_n **éléments,** $|\vec{p}| = p_1 + p_2 + \ldots + p_n = k$ **. Ce** nombre $\gamma_{p_1, p_2 \ldots p_n} = \gamma_{\vec{p}}$ s'appelle aussi le nombre des "permutations avec répétitions" **des** n **objets** $\{1, 2, \ldots n\}$, où **l'objet** 1 **est pris** p_1 fois, **l'objet** 2 , p_2 **fois,.... l'objet** n, p_n **fois.** Il y a bien des manières de le **calculer. Pour déterminer une** application j qui nous **intéresse, nous** devons d'abord choisir une partie de $\{1, 2, \ldots k\}$ à p_1 **éléments,** qui sera $j^{-1}\{1\}$: il y **a** un nombre de choix possibles qui est le **nombre de** parties **à** p_1 éléments de $\{1, 2, \ldots k\}$, soit $\frac{k!}{(p_1)!\,(k-p_1)!}$. Ceci choisi, nous devons choisir, dans le complément $\{1, 2, \ldots k\} - j^{-1}\{1\}$

* Dans (III,6;25) on avait des vecteurs $\vec{X}_1, \vec{X}_2, \ldots$ quelconques **ici** , **ils** sont tous égaux **à** $\vec{h}$.

ensemble à $k - p_1$ éléments, une partie arbitraire à p_2 éléments, qui sera $j^{-1}\{2\}$; le nombre des choix possibles est $\dfrac{(k-p_1)!}{(p_2)!\,(k-p_1-p_2)!}$. Et ainsi de suite. Le nombre de ces applications j est donc

$$\gamma_{\vec{p}} = \frac{k!}{p_1!\,(k-p_1)!}\;\frac{(k-p_1)!}{p_2!\,(k-p_1-p_2)!}\cdots\frac{(k-p_1-p_2\cdots-p_{n-1})!}{p_n!\,(k-p_1-p_2\cdots-p_n)!} = \frac{k!}{p_1!\,p_2!\cdots p_n!} = \frac{k!}{\vec{p}\,!}\,, \tag{III,7;15}$$

donc

$$\frac{\gamma_{\vec{p}}}{k!} = \frac{1}{\vec{p}\,!}\,, \tag{III,7;16}$$

et le théorème est démontré.

Il est remarquable qu'avec les notations employées, la formule soit identique à celle qui correspond aux fonctions d'une variable. Dans le premier membre, $f(x+\vec{h})$ est dans l'espace affine F ; dans le 2ème membre, tous les termes sont dans l'espace vectoriel $\vec{F}$ sauf le premier, correspondant à $\vec{p} = \vec{0}$, soit $f(x)$, qui est dans l'espace affine F .

Corollaire. Dans les conditions du théorème, si on a trouvé des éléments $\vec{c}_{\vec{p}}$ ($\vec{c}_{\vec{p}} \in \vec{F}$ pour $\vec{p} \neq \vec{0}$, $c_{\vec{0}} \in F$), $|\vec{p}| \leqslant m$, tels que

$$f(a+\vec{h}) = \sum_{\vec{p}} \vec{c}_{\vec{p}}\,\frac{\vec{h}^{\vec{p}}}{\vec{p}\,!} + \vec{\alpha}\,\|\vec{h}\|^m\,, \tag{III,7;17}$$

où $\vec{\alpha} \in \vec{F}$ tend vers $\vec{0}$ avec $\vec{h}$, et si l'on sait que f est m fois dérivable en a , on a nécessairement

$$\vec{c}_{\vec{p}} = \overrightarrow{D^{\vec{p}} f(a)}\,.$$

Démonstration - Par différence, si on pose $\vec{a}_{\vec{p}} = \overrightarrow{D^{\vec{p}} f(a)} - \vec{c}_{\vec{p}}$, (ici $\vec{a}_{\vec{0}} \in \vec{F}$ aussi) on voit que le "polynome" * (à coefficients vectoriels) en h_1 , h_2 , ... h_n :

$$\sum_{|\vec{p}|\leqslant m} \vec{a}_{\vec{p}}\,\vec{h}^{\vec{p}} = \sum \vec{a}_{p_1,p_2,\dots p_n}\,h_1^{p_1}\,h_2^{p_2}\cdots h_n^{p_n} \tag{III,7;18}$$

* Voir plus de renseignements sur les polynomes au théorème 22 -

a une norme infiniment petite devant $\|\vec{h}\|^m$ quand $\vec{h}$ tend vers $\vec{0}$. En posant $\vec{h} = t\vec{X}$, et en faisant tendre t vers 0 on voit, par une récurrence analogue à celle qui est utilisé; dans la démonstration du théorème 21 bis (mals sans utiliser le lemme), que les parties homogènes de degrés successifs $0, 1, 2, \ldots m$, du polynome $\sum_{|\vec{r}| \leq m} \vec{a}_{\vec{r}}\, \vec{X}^{\vec{r}}$ sont identiquement nulles. Mais un polynome identiquement nul a tous ses coefficients nuls (*), ce qui démontre le corollaire.

Exercice - Appliquer le corollaire pour démontrer de nouveau la formule de Leibnitz (III,6;36) dans l'esprit du théorème 21 ter.

Théorème 22 - Pour qu'une application f d'un ouvert connexe Ω de E dans F ait une dérivée d ordre $m+1$ nulle dans Ω, il faut et il suffit que f soit un polynome de degré $\leq m$.

Démonstration - Soit d'abord $m = 0$. On a vu que, même si Ω n'est pas connexe, une constante a une dérivée nulle. Nous devons montrer que, si Ω est connexe, une application f de $\Omega \subset E$ dans F, de dérivée partout nulle, est une constante.

Soit x un point quelconque de Ω. Il existe un nombre $\rho_x > 0$ tel que la boule ouverte B_x de centre x et de rayon ρ_x soit dans Ω. Alors, pour $\|\vec{h}\| < \rho_x$, on peut appliquez la formule des accroissements finis au segment $[x, x+\vec{h}]$, tout entier dans Ω, et on trouve :

(III,7;20) $$\|f(x+\vec{h}) - f(x)\| \leq \sup_{\xi \in]x, x+\vec{h}[} \|f'(\xi)\| = 0 .$$

Donc f est égale à la constante $f(x)$ dans B_x.

Soit alors un point a quelconque fixe de Ω. Appelons A l'ensemble des points ξ de Ω tels que $f(\xi) = f(a)$. A est évidemment fermé, car il est l'image réciproque de $\{f(a)\} \in F$ par l'application continue f. Il est aussi ouvert, car, si $x \in A$, toute la boule B_x définie ci dessus appartient aussi à A Comme Ω est connexe et que A n'est pas vide ($a \in A$), A est Ω tout entier, ce qui démontre le théorème pour $m = 0$.

Soit maintenant $m \geq 1$. Il faudrait donner d'abord la définition d'un polynome sur E à valeurs dans F. Nous nous bornerons au cas particulier où E est de dimension finie n.

* Ceci n'a été démontré en Mathématiques Spéciales, par récurrence sur le nombre n des variables X_1, X_2, . . . X_n que pour un polynome à coefficients scalaires, mais le résultat subsiste pour des coefficients vectoriels, et la démonstration est la même.

Si on choisit un référentiel $O, \vec{e}_1, \vec{e}_2, \dots \vec{e}_n$, on Sait ce qu'est un polynome en $x_1, x_2, \dots x_n$, de degré $\leq n$ à valeurs dans F : c'est une fonction de la forme

$$(\text{III},7;21) \quad f(x) = \sum_{|\vec{p}|\leq m} \vec{a}_{\vec{p}}\, x^{\vec{p}} = \sum_{p_1+p_2+\dots+p_n\leq m} a_{p_1,p_2,\dots p_n}\, x_1^{p_1} x_2^{p_2} \dots x_n^{p_n} ,$$

où les $\vec{a}_{\vec{p}}$ sont dans $\vec{F}$, sauf $a_{\vec{0}}$ qui est dans F .

Une telle fonction garde la même forme dans tout autre référentiel, et s'appelle polynome sur E à valeurs dans F, de degré $\leq m$. Sa dérivée d'ordre $m+1$ est nulle, que Ω soit connexe ou non, puisque, d'après (III,7;21), ses dérivées partielles d'ordre $m+1$ sont nulles.

Réciproquement, soit Ω connexe, et soit f une application de Ω dans F , de dérivée d'ordre $m+1$ nulle. Soit $a \in \Omega$. La formule de Taylor d'ordre m, par rapport à un référentiel quelconque, n'a pas de terme complémentaire $\vec{R}_m$, et montre que f est, dans la boule B_a signalée plus haut, un polynome P de degré $\leq m$. Appelons alors k un entier tel que $(f-P)^{(k)} \equiv 0$ dans Ω ; de tels entiers existent, par exemple $k = m+1$. Alors, si $k \geq 1$, $(f-P)^{(k-1)}$ est une fonction dont la dérivée première est nulle dans Ω connexe, donc elle est constante dans Ω : mais elle est nulle dans B_a , donc elle est nulle dans Ω , et $k-1$ a la même propriété que k . Alors, de proche en proche, on descend jusqu'à $k=0$, et $f \equiv P$ dans Ω, ce qui démontre le théorème.

Remarques -1°/ Bien entendu cette conclusion serait complètement fausse si Ω n'était pas connexe. Il suffit de prendre pour Ω le complémentaire de l'origine sur la droite réelle $\mathbb{R}$, et de prendre $m=0$; une fonction définie sur Ω à valeurs réelles et à dérivée partout nulle, n'est pas nécessairement une constante dans Ω ; elle peut être égale à une certaine constante dans la demi droite $]-\infty, 0[$, et à une constante différente dans la demi droite $]0, +\infty[$.

2°/ Supposons que nous n'ayons pas, à la formule (III,7;15), déterminé $\gamma_{\vec{p}}$. On peut le faire maintenant très facilement comme suit.

Ecrivons la formule de Maclaurin pour un polynome de degré m, $\gamma_{\vec{p}}$ n'étant pas supposé connu :

$$(\text{III},7;22) \quad \begin{cases} \vec{f}(\vec{x}) = \displaystyle\sum_{|\vec{p}|\leq m} \gamma_{\vec{p}} \frac{\overrightarrow{D^{\vec{p}} f(0)}}{k!} \vec{x}^{\vec{p}} = \sum_{|\vec{p}|\leq m} c_{\vec{p}} \frac{\overrightarrow{D^{\vec{p}} f(0)}}{\vec{p}\,!} \vec{x}^{\vec{p}} \\ \text{avec} \\ c_{\vec{p}} = \dfrac{\vec{p}\,!}{k!} \gamma_{\vec{p}} , \end{cases}$$

et nous voulons montrer que $c_{\vec{p}} = 1$

soit $\vec{q} \in \mathbb{N}^n$, $\vec{q} \leqslant \vec{p}$. On a **immédiatement** :

$$D^{\vec{q}}\,\frac{\vec{x}^{\vec{p}}}{\vec{p}\,!} = \left(\left(\frac{\partial}{\partial x_1}\right)^{q_1}\frac{x_1^{p_1}}{p_1!}\right)\left(\left(\frac{\partial}{\partial x_2}\right)^{q_2}\frac{x_2^{p_2}}{p_2!}\right)\cdots\left(\left(\frac{\partial}{\partial x_n}\right)^{q_n}\frac{x_n^{p_n}}{p_n!}\right) \tag{III,7;23}$$

$$= \begin{cases} \dfrac{\vec{x}^{\vec{p}-\vec{q}}}{(\vec{p}-\vec{q})!}\,, & \text{si } \vec{q} \leqslant \vec{p} \\ 0 & \text{si } \vec{q} \nleqslant \vec{p} \end{cases} \quad *$$

Alors la dérivation $D^{\vec{q}}$ de (III,7:22), pour $|\vec{q}\,| \leqslant m$, donne :

$$\overrightarrow{D^{\vec{q}} f(x)} = \sum_{\substack{|\vec{p}|\leqslant m \\ \vec{p} \geqslant \vec{q}}} c_{\vec{p}}\,\frac{\overrightarrow{D^{\vec{p}} f(0)}}{(\vec{p}-\vec{q}\,)!}\,x^{\vec{p}-\vec{q}} \tag{III,7;24}$$

Faisons $\vec{x} = \vec{0}$. Alors $\vec{x}^{\,\vec{p}-\vec{q}} = 0$ pour $\vec{p} \neq \vec{q}$, $= 1$ pour $\vec{p} = \vec{q}$. On obtient donc

$$\overrightarrow{D^{\vec{q}} f(0)} = c_{\vec{q}}\,\overrightarrow{D^{\vec{q}} f(0)}\ , \tag{III,7;25}$$

d'où $c_{\vec{q}} = 1$, comme nous voulions le montrer.

Prenons en particulier $f(x) = (x_1 + x_2 + \cdots + x_n)^k$.

Le théorème des fonctions **composées** montre que toute dérivée partielle d'ordre ℓ **est** $k(k-1)\cdots(k-\ell+1)(x_1+x_2+\cdots+x_n)^{k-\ell}$.

Une telle dérivée est toujours nulle pour $x_1 = x_2 \cdots = x_n = 0$, sauf si $\ell = k$, auquel cas elle vaut $k\,!$ Alors la formule de Maclaurin donne

$$(x_1 + x_2 + \cdots + x_n)^k = \sum_{|\vec{p}|=k} \gamma_{\vec{p}}\,\vec{x}^{\vec{p}} = \sum_{|\vec{p}|=k} \frac{k!}{\vec{p}\,!}\,\vec{x}^{\vec{p}} \ . \tag{III,7;26}$$

Ainsi $\gamma_{p_1,p_2,\dots p_n} = \dfrac{(p_1+p_2+\cdots+p_n)!}{p_1!\,p_2!\cdots p_n!}$ est le coefficient de $x_1^{p_1}\,x_2^{p_2}\dots x_n^{p_n}$ dans le **développement** de $(x_1 + x_2 + \cdots + x_n)^{p_1+p_2+\cdots+p_n}$, ce qui **généralise** la formule du binome. C'était d'ailleurs évident, d'après la définition combinatoire de $\gamma_{\vec{p}}$ donnée après (III,7 14).

Les coefficients du binome sont très simples, ce qui permet d'écrire immédiatement la formule de Maclaurin pour une fonction de 2 variables scalaires x, y . Si on pose

2 * La relation d'ordre $\vec{q} \leqslant \vec{p}$ **dans** $\mathbb{N}^n$ est partielle et non totale, et $\vec{q} \nleqslant \vec{p}$ n'est pas synonyme de $\vec{q} > \vec{p}$. Ainsi, **pour** $n = 2$, $(1,1) \nleqslant (0,2)$ **mais on n'a pas** $(1,1) > (0,2)$.

$$\vec{p} = \overrightarrow{\frac{\partial f}{\partial x}}(0,0)\,, \qquad \vec{q} = \overrightarrow{\frac{\partial f}{\partial y}}(0,0)$$

(III,7;26 bis) $$\vec{r} = \overrightarrow{\frac{\partial^2 f}{\partial x^2}}(0,0)\,, \qquad \vec{s} = \overrightarrow{\frac{\partial^2 f}{\partial x\,\partial y}}(0,0)\,, \qquad \vec{t} = \overrightarrow{\frac{\partial^2 f}{\partial y^2}}(0,0)$$

$$\vec{\alpha} = \overrightarrow{\frac{\partial^3 f}{\partial x^3}}(0,0)\,, \qquad \vec{\beta} = \overrightarrow{\frac{\partial^3 f}{\partial x^2\,\partial y}}(0,0)\,, \qquad \vec{\gamma} = \overrightarrow{\frac{\partial^3 f}{\partial x\,\partial y^2}}(0,0)\,, \qquad \vec{\delta} = \overrightarrow{\frac{\partial^3 f}{\partial y^3}}(0,0),$$

on aura

(III,7;26 ter) $$\begin{aligned} f(x,y) = {} & f(0,0) + \vec{p}\,x + \vec{q}\,y \\ & + \frac{1}{2}\left(\vec{r}\,x^2 + 2\vec{s}\,xy + \vec{t}\,y^2\right) \\ & + \frac{1}{6}\left(\vec{\alpha}\,x^3 + 3\vec{\beta}\,x^2 y \;t\; 3\vec{\gamma}\,x y^2 + \vec{\delta}\,y^3\right) + \cdots \end{aligned}$$

Application à l'étude des maxima et des minima. Définitions

Soit f une fonction définie sur un ensemble E et à valeurs réelles. Nous avons défini ce qu'étalent pour f, un maximum ou un minimum. Si maintenant E est un espace topologique, on dit que f admet, en un point a de E, un maximum relatif *, s'il existe un voisinage $\mathcal{V}$ de a tel que la restriction de f à ce voisinage admette au point a un maximum, c'est-à dire si, pour tout x de $\mathcal{V}$, on a l'inégalité : $f(x) \leqslant f(a)$. On dit qu'il s'agit d'un maximum relatif strict, si l'on peut choisir $\mathcal{V}$ de manière que l'on ait, pour tout $x \neq a$ de $\mathcal{V}$, l'inégalité stricte $f(x) < f(a)$.

Définition analogue pour un minimum relatif et un minimum relatif strict.

Bien entendu un maximum de f, c'est à dire un maximum absolu, est nécessairement a fortiori un maximum relatif, alors que le contraire n'est pas nécessairement vrai.

On dit que f présente en a un extremum si elle Présente un maximum ou un minimum relatif.

* Au lieu de dire que f admet en a un maximum, ou maximum relatif, etc..., on dit aussi que a est pour f un maximum, ou maximum relatif, etc... C est assez incorrect, car le maximum est $f(a)$ et non a !

Condition nécessaire pour un extremum

Théorème 23 - Soit f une fonction réelle définie sur un ouvert Ω d'un espace affine normé E, et dérivable. Une condition nécessaire pour qu'elle admette en un point a de Ω un maximum ou un minimum relatif, est que l'application dérivée $f'(a) \in \mathcal{L}(\vec{E}\,;\mathbb{R}) = \vec{E}'$ soit nulle.

Démonstration - Remarquons d'abord que si E est de dimension finie et si l'on en a choisi un référentiel 0, $(\vec{e}_i)_{i\in I}$, la condition s'exprime sous la forme bien connue :
Il est nécessaire, pour que a soit un maximum ou un minimum relatif pour f, que les dérivées partielles $\frac{\partial f}{\partial x_i}$ soient toutes nulles au point a, ou encore que la différentielle au point a : $\sum_{i\in I} \frac{\partial f}{\partial x_i}(a)\, dx_i$ soit identiquement nulle.

Démontrons maintenant le théorème dans le cas général.

Si f admet un maximum ou minimum relatif en a, la fonction $t \to f(a+t\vec{X})$, $\vec{X}$ fixé dans $\vec{E}$, est définie pour $t\in\mathbb{R}$, $|t|$ assez petit (parce que Ω est ouvert), et admet au point $t = 0$ un maximum ou minimum relatif. Alors la même démonstration que celle qui a été donnée pour le théorème de Rolle (théorème 4) montre que sa dérivée en t est nulle pour $t = 0$; or c'est $f'(a)\cdot\vec{X}$. Donc $f'(a)\cdot\vec{X} = 0$ quel que soit $\vec{X} \in \vec{E}$, donc $f'(a) = 0$.

Remarques. Le théorème 23, pour beaucoup de raisons, ne règle pas le problème de la recherche des extrema de f.

1°/ Généralement, f sera une fonction réelle définie, non sur un ouvert, mais sur un fermé F (par exemple, sur un segment $[a, b]$ de $\mathbb{R}$, ou dans une boule fermée de E); elle peut n'avoir aucun extremum, si son ensemble de définition F n'est pas compact (par exemple $f(x) = x$ n'a pas d'extremum sur $\mathbb{R}$); si elle a des extrema, ceux-ci peuvent échapper au théorème, qui ne s'applique qu'aux extrema dans un ouvert ; naturellement si f est dérivable dans l'intérieur $\Omega = \overset{\circ}{F}$ de F, tout extremum a situé dans Ω sera obtenu en écrivant $f'(a) = 0$, mais justement les extrema seront souvent sur la frontière de F (par exemple, si f est la fonction x^2 sur l'intervalle fermé $[-1, +1]$, le minimum $x = 0$ est obtenu par annulation de la dérivée, parce que 0 est intérieur à $[-1, +1]$; mais les maxima $x = \pm 1$

échappent, parce qu'ils sont réalisés aux extrémités de l'intervalle et que la dérivée ne s'y annule pas *

2°/ Inversement, l'équation $f'(a) = 0$ donne des points qui ne sont pas nécessairement des extrema, autrement dit la condition nécessaire $f'(a) = 0$ est loin d'être suffisante. Il faut encore étudier le développement de Taylor de f au voisinage du point a .

Nous étudierons maintenant ce problème. Mais il est certain que, bien souvent, le fait que $f'(a)$ soit nul est plus important que la propriété d'extremum elle-même. Par exemple, si on considère l'hypersurface d'équation $y = f(x)$ dans $\Omega \times \mathbb{R}$, $f'(a) = 0$ signifie que l'hyperplan tangent au point $A = (a, f(a))$ est horizontal; c'est une condition nécessaire d'extremum de l'ordonnée y, mais géométriquement plus importante que la propriété d'extremum. Qu'il y ait ou non extremum, nous dirons que f est stationnaire en a si $f'(a) = 0$.

Recherche de conditions nécessaires et suffisantes pour un extremum

Théorème 24 - Soit f une fonction à valeurs réelles définie sur un ouvert Ω d'un espace affine normé E, et m fois dérivable. Si les dérivées de f d'ordre $1, 2, \ldots m-1$, sont toutes nulles en $a \in \Omega$, si $f^{(m)}(a) \neq 0$, et si f admet en a un maximum relatif, alors m est pair, et l'on a:

(III,7;27) $$f^{(m)}(a) \cdot \vec{X}^m \leq 0 \quad \text{pour tout } \vec{X} \in \vec{E} .$$

Inversement, si f est m fois dérivable dans Ω, si les dérivées de f d'ordre $1, 2, \ldots m-1$, sont toutes nulles en $a \in \Omega$, et si $f^{(m)}(a) \cdot \vec{X}^m$ est borné supérieurement par un nombre $-\delta < 0$ lorsque $\vec{X}$ parcourt la sphère unité $\|\vec{X}\| = 1$ de $\vec{E}$, alors f admet en a un maximum relatif strict.

* Nous avons dit, page 184, au début du § 2, que la plupart des théorèmes énoncés pour un ouvert Ω de E s'appliquaient aussi à d'autres ensembles, notamment à des intervalles non ouverts de $\mathbb{R}$; le théorème 23 est un de ceux qui ne s'appliquent pas. Sa démonstration est basée sur celle qui a été donnée au théorème de Rolle (théorème 4); elle suppose essentiellement l'extremum réalisé en un point intérieur à l'intervalle.

Démonstration

1") Démonstrons d'abord la première partie du **théorème, et** supposons donc que f ait en $a \in \Omega$ un maximum relatif. Alors la fonction $g : t \longrightarrow f(a + t\vec{X})$, définie au voisinage de $t = 0$ sur la droite **réelle** $\mathbb{R}$, a un maximum relatif, pour $t = 0$. Sa dérivée d'ordre k pour $t = 0$ est $f^{k}(a).\vec{X}^{k}$.

Ses dérivées d'ordre $\leqslant m-1$ sont donc nulles, pour $t = 0$, sa **dérivée** d'ordre m est $f^{(m)}(a).\vec{X}^{m}$. On a donc, pour t voisin de 0 :

$$(\text{III},7;28) \qquad g(t) - g(0) = \left(f^{(m)}(a) . \vec{X}^{m}\right) \frac{t^{m}}{m!} + \varepsilon_t t^{m} ,$$

où ε_t tend vers 0 avec $|t|$. Donc ou bien $f^{(m)}(a).\vec{X}^{m} = 0$, **ou bien** $g(t) - g(0)$, au voisinage de $t = 0$, est du signe de $\left(f^{(m)}(a) . \vec{X}^{m}\right)$. On a donc finalement, quels que soient $\vec{X} \in \vec{E}$ et t réel :

$$(\text{III},7;29) \qquad \left(f^{(m)}(a) . \vec{X}^{m}\right) t^{m} \leqslant 0 .$$

Si m est pair, c'est équivalent **à** **(III,7;27)**. Pour démontrer la première partie, **il** reste donc **à** montrer quemne saurait être impair. Or, **si** m **était** impair, **(III,7;29)** équivaudrait (en prenant sucesslvement $t = +1$ et $t = -1$) â

$$(\text{III},7;30) \qquad f^{(m)}(a) . \vec{X}^{m} = 0 \text{ pour tout } \vec{X} \in E .$$

Cela entrainerait, d'après le lemme démontré au théorème 21 bis,

$$(\text{III},7;31) \qquad f^{(m)}(a) . (\vec{X}_1 , \vec{X}_2, \dots \vec{X}_m) = 0, \text{ quel que soit } (\vec{X}_1 , \vec{X}_2 , \dots \vec{X}_m) \in \vec{E}^{m}$$

ou $f^{(m)}(a) = 0$, ce **qui serait** contraire **à** l'hypothèse.

Remarque - Avant de passer **à** la deuxième partie, remarquons ceci. On pourrait croire que, **si** f admet en a un maximum relatif strict, alors la forme de degré m : $f^{(m)}(a)$ **soit** définie négative, autrement dit qu'on ait :

$$(\text{III},7;35) \qquad f^{(m)}(a) . \vec{X}^{m} < 0 \quad \text{pour} \quad \vec{X} \neq \vec{0}.$$

Il n'en est rien (sauf évidemment si E a la dimension 1) Considérons en effet la fonction réelle f de deux variables réelles x, y :

(III,7;36) $$f(x,y) = -(x^2+y^4).$$

Elle admet à l'origine un maximum absolu strict. Comme c'est un polynome, son développement suivant les puissances de x et y est son développement de Maclaurin, de sorte que, ici, si on appelle (X, Y) un vecteur de $\vec{E} = \mathbb{R}^2$, on a :

(III,7;37) $$\frac{1}{2!} f''(0,0) \cdot (\overrightarrow{X,Y})^2 = -X^2$$

est bien toujours ≤ 0, mais non défini négatif (c'est nul pour le vecteur $(\overrightarrow{0,Y}) \neq (\overrightarrow{0,0})$).

2°) Démontrons maintenant la 2ème partie. <u>Remarquons toutefois d'abord que les hypothèses de cette deuxième partie sont plus fortes que les conclusions de la Premiere</u>. Il n'en peut être autrement. Supposer que $f^{(k)}(a) = 0$ pour $k = 1, 2, \ldots m-1$, que $f^{(m)}(a) \neq 0$ et que l'on a (III,7;27), n'est pas suffisant pour affirmer que f ait en a un maximum relatif.

Considérons en effet la fonction f de 2 variables réelles x, y :

(III,7;38) $$f(x,y) = -x^2 + y^4 .$$

Comme nous l'avons vu pour (III,7;36), on a (III,7;37) c'est-à-dire (III,7;27); cependant $f(0,0) = 0$, et $f(0,y) > 0$ pour $y \neq 0$, donc f n'a pas à l'origine un maximum relatif. C'est pourquoi nous supposons les hypothèses plus fortes Indiquées dans l'énoncé : $f^{(m)}(a) \cdot \vec{X}^m$ est majoré, sur la sphère unité, par $-\delta < 0$. Cela entraine, par homogénéité :

(III,7;39) $$f^{(m)}(a) \cdot \vec{X}^m \leq -\delta \|\vec{X}\|^m .$$

Cela entraine donc (III,7;35) : $f^{(m)}(a)$ est définie négative. <u>C'est équivalent à</u> (III,7;35) <u>si</u> E <u>est de dimension finie</u>; car la sphere unité est alors compacte, la fonction $\vec{X} \longrightarrow f^{(m)}(a) \cdot \vec{X}^m$ est continue: si elle est partout < 0 sur la sphère unité, comme il résulte de (III,7;35), elle y admet un maximum, et celui-ci est bien un nombre $-\delta < 0$. Mais c'est une hypothèse plus forte que (III,7;35) si E est de dimension Infinie.

Ecrivons le développement de Taylor d'ordre m sous la forme (III,7;10) :

$$f(a+\vec{h}) = f(a) + \frac{f^{(m)}(a)}{m!} \cdot \vec{h}^{m} + \alpha(\vec{h}) \|\vec{h}\|^{m}, \tag{III,7;40}$$

où $\alpha(a)$ tend vers 0 avec $\vec{h}$. Alors (III,7;39) donne

$$f(a+\vec{h}) \leqslant f(a) - \delta \|\vec{h}\|^{m} + \alpha(\vec{h}) \|\vec{h}\|^{m} ; \tag{III,7;41}$$

en vertu de la propriété de $\vec{\alpha}$, $f(a+\vec{h}) - f(a)$ est < 0 pour $\vec{h} \neq \vec{0}$ de norme assez petite, et f admet bien en a un maximum relatif strict.

Naturellement, on a un théorème analogue pour le minimum. D'où la règle, permettant de recnnnaitre si f admet en a un maximum ou minimum relatif :

Règle - On doit d'abord avoir $f'(a) = 0$. On cherche alors le premier entier m tel que $f^{(m)}(a) \neq 0$. Si m est impair, a n'est ni un maximum ni un minimum relatif. Si m est pair, on regarde le signe de la forme de degré m : $\vec{X} \longrightarrow f^{(m)}(a) \cdot X^{m}$. Si elle peut prendre des valeurs des deux signes, a n'est ni un maximum ni un minimum relatif * ~~. Si elle~~ est toujours $\leqslant 0$ (resp. $\geqslant 0$), sans être définie négative (resp. définie positive), on ne peut pas conclure sans étudier les dérivées d'ordre $> m$. Si elle est définie négative (resp. définie > 0) et si E est de dimension finie, ou si l'on a la condition plus forte $\sup\limits_{\|\vec{X}\|=1} (f^{(m)}(a) \cdot \vec{X}^{m}) < 0$ (resp. $\inf\limits_{\|\vec{X}\|=1} (f^{(m)}(a) . \vec{X}^{m}) > 0$) pour E de dimension Infinie, alors a est un maximum (resp. minimum) relatif strict.

Il résulte de toute cette étude qu'il n'existe pas de condition à la fois nécessaire et suffisante simple pour qu'une fonction réelle f, définie sur $\Omega \subset E$, admette, en un point a, un maximum ou un minimum relatif.

* Le point a s'appelle alors un col pour la fonction f. L exemple typique est celui de la fonction réelle de 2 variables réelles $f(x,y) = xy$. Son graphique dans $\mathbb{R}^3$ est le paraboloïde hyperbolique $z = xy$. L'origine est un col pour f. La restriction de f à la droite $y = mx$ est $x \rightarrow mx^2$, qui admet un maximum strict absolu à l'origine si $m < 0$ et un minimum strict absolu si $m > 0$. On pourra aussi parler d'un col si m est impair.

2

Remarques 1°/ Si une fonction f admet toutes ses dérivées successives nulles au point a, sans être identiquement nulle, l'étude du développement de Taylor ne permet absolument plus de voir si elle admet un maximum ou un minimum relatif.

2°/ Supposons que f ait un col en a. On pourra chercher dans quelle région de Ω, au voisinage de a, on a $f(x) \geqslant f(a)$, et dans quelle région on a $f(x) \leqslant f(a)$. On appliquera les mêmes règles. Si sur l'ensemble B de la sphère unité $\|\vec{X}\| = 1$, la quantité $f^{(m)}(a) \cdot \vec{X}^m$ est majorée par $-\delta < 0$, alors, dans tout le cône $\{a + \lambda\vec{X}\ ;\ \vec{X} \in B,\ \lambda$ réel $> 0\}$, on a, au voisinage de a, $f(x) < f(a)$.

Cas particulier d'une fonction f réelle de 2 variables réelles x, y

. On cherchera d abord les points ou les dérivées partielles $p = \frac{\partial f}{\partial x}$, $q = \frac{\partial f}{\partial y}$, sont nulles. Soit (a, b) un tel point. On formera alors le développement de Taylor de f suivant les puissances de $x - a = X$, $y - b = Y$. Ce développement commencera par

(III,7 ;42) $$\frac{1}{2}(r X^2 + 2s X Y + t Y^2)$$

1er cas $rt - s^2 > 0$, r (et t) < 0.

Alors la forme quadratique (III,7;42) est définie négative, f admet un maximum relatif strict en (a, b).

2ème cas $rt - s^2 > 0$, r (et t) > 0. f admet en (a, b) un minimum relatif strict.

3ème cas $rt - s^2 < 0$. La forme quadratique (III,7;42) prend à la fois des valeurs des 2 signes, on est en présence d'un col. L'équation $r X^2 + s X Y + t Y^2 = 0$ représente deux droites D, D', de $\mathbb{R}^2$, qui définissent 4 angles, 2 à 2 opposés par le sommet, soient (1), (2), (3), (4), dans l'ordre de parcours. Par exemple, supposons la forme $\geqslant 0$ dans (1) et (3), $\leqslant 0$ dans (2) et (4).

Supposons l'angle (1) défini, en coordonnées polaires, par $\rho \geqslant 0$, $\alpha \leqslant \varphi \leqslant \beta$. Alors tout angle $\alpha' \leqslant \varphi \leqslant \beta'$, où $\alpha' > \alpha$ et $\beta' < \beta$, coupe la "sphère unité" de $\mathbb{R}^2$, c'est-à-dire le cercle trigométrique, suivant un arc compact, sur lequel par conséquent la forme quadratique admet un minimum $\delta > 0$.

Alors il existera un $\varepsilon > 0$ tel que, si l'on pose $x = a + \rho\cos\varphi$, $y = b + \rho\sin\varphi$, $\rho \geq 0$, $\alpha' \leq \varphi \leq \beta'$, on ait $f(x,y) > f(a,b)$ pour $0 < \rho \leq \varepsilon$.

Mais cet ε dépend de α' et β' ; on ne peut rien affirmer lorsque ρ tend vers 0, si en même temps φ tend vers α ou β, sans l'examen des dérivées d'ordre > 2 de f en (a,b). Mêmes conclusions dans l angle (3); dans les angles (2) et (4), conclusionsanalogues mais avec remplacement de $f(x,y) > f(a,b)$ par $f(x,y) < f(a,b)$.

Les régions de $\mathbb{R}^2$ où $f(x,y) > f(a,b)$ et $f(x,y) < f(a,b)$ sont séparées par la courbe $f(x,y) = f(a,b)$. Celle-ci a, au point (a,b), 2 branches, respectivement tangentes à D et D'. Mais on ne peut pas dire dans quels angles sont ces 2 branches, sans examen des dérivées d'ordre > 2 de f en (a,b).

4ème cas $rt - s^2 = 0$, r ou $t < 0$.

La forme quadratique est l'opposée du carré d'une forme linéaire $\neq 0$. Elle s'annule sur une droite D et elle est < 0 ailleurs. Soit $\varphi = \varphi_0 + k\pi$ l'angle polaire de cette droite. Si

$x = a + \rho\cos\varphi$, $y = b + \rho\sin\varphi$, $\rho \geq 0$, $\alpha' \leq \varphi \leq \beta'$, où α' et β' sont fixés, $\alpha' > \varphi_0$ et $\beta' < \varphi_0 + \pi$, il existe $\varepsilon > 0$ tel que l'on ait $f(x,y) < f(a,b)$ pour $0 < \rho \leq \varepsilon$. Résultat analogue pour $\alpha' > \varphi_0 + \pi$, $\beta' < \varphi_0 + 2\pi$.

Mais on ne peut rien dire de plus, et en particulier on ne peut pas dire si (a, b) est un maximum relatif, sans examen des dérivées d'ordre > 2 de f en (a,b).

5ème Cas $rt - s^2 = 0$, r ou $t > 0$.

Résultat analogue, en remplaçant $f(x,y) < f(a,b)$ par $f(x,y) > f(a,b)$.

6ème Cas $r = s = t = 0$. On ne peut rien conclure sans l'examen des dérivées d'ordre > 2 de f en (a,b).

Conclusion Cette dtude montre que, pour une fonction de plus d'une variable **réelle,** les circonstances sont très différentes de celles qui ont été vues en Mathématiques Spéciales pour les fonctions d'une variable réelle.

Application de la formule de Taylor à l'étude da la position d'une hypersurface par rapport à son hyperplan tangent

Soit# une fonction réelle dérivable définie sur $\Omega \subset E$. Alors $y = f(x)$ est l'équation d'une variété différentiable de $E \times \mathbb{R}$; pour des raisons de dimension, on dit que c'est une hypersurface différentiable. Son hyperplan tangent en $A = (a, f(a))$ a l'équation (III,3;19 ter). La position de l'hypersurface par rapport à son hyperplan tangent, au voisinage de A , est donnée par le signe de

(III,7;43) $$\overline{f(x) - f(a)} - f'(CL).\overrightarrow{x-a} ,$$

pour x voisin de a .

Or $x \longrightarrow f(x) - f'(a).\overrightarrow{x-a}$ est une fonction réelle dérivable, de dérivée nulle en a . Tout revient à voir si a est un extremum pour f et quel type d'extremum, ou un col. Nous eommes donc ramenés au théorème 24. Conformément à la conclusion de ce qui précède, le col est une situation courante et non exceptionnelle, et par conséquent, sauf si E a la dimension 1, l'hypersurface peut normalement traverser son hyperplan tangent, même si $f''(a) \neq 0$. Dans le cas particulier α une fonction f réelle de 2 variablesréelles x , y , la surface sera· localement d'un même côté du plan tangent si $rt - s^2 > 0$, et le traversera si $rt - s^2 < 0$.

§ 8 THÉORÈME DES FONCTIONS IMPLICITES. POSITION DU PROBLÈME

Soit f une application d'un ensemble E dans un ensemble F . Alors, si& est un point de F ,on peut se proposer de chercher l'ensemble des x de E tels que $f(x) = b$ c'est-à-dire l'image réciproque $f^{-1}(\{b\})$. Cela s'appelle résoudre l'équation : $f(x) = b$. Soit maintenant f une application d'un produit d'ensembles, $E \times F$, dans un ensemble G , et soit c un point de G .

Considérons l'équation :

(III,8;1) $$f(x, y) = c .$$

Il se peut que, lorsque x est donné, l'équation en y : $f(x,y) = c$, admette une solution et une seule et cela quelle que soit la valeur donnée de x. Dans ce cas l'équation définit y comme une fonction $g(x)$ de x, cette fonction est appelee la fonction implicite définie par l'équation précédente, elle est caractérisée par la propriété :

(III,8;2) $$f(x, g(x)) = c.$$

On peut encore dire que, pour $(x,y) \in E \times F$, la relation (III,8;1) est identique à la relai on $y = g(x)$.

Naturellement il sera bien rare que d'aussi belles circonstances se produisent. Il arrivera souvent que, pour certaines valeurs de x, il n'y ait pas de solution en y, et que, pour d'autres valeurs de x il y ait plusieurs solutions en y, voire une infinité: Le cas particulier que nous nous proposons d'étudier est le suivant : E, F, G, sont des espaces topologiques; on suppose que l'on ait une solution particulière de l'équation,= $= a$, $y = b$. On se propose de savoir si, pour les points x suffisamment voisins du point a, l'équation ne posséderait pas une solution et une seule en y, pourvu que l'on astreigne cette solution y à être suffisamment voisine de b. Lorsqu'il en sera ainsi, on aura bien défini une fonction implicite $y = g(x)$ à partir de l'équation, tout au moins dans un voisinage de (a,b). L'interprétation géométrique est simple, notamment si nous supposons que E, F, G, sont tous les trois le corps des réels. L'équation $f(x,y) = 0$ définit alors une courbe de $\mathbb{R}^2$, et nous nous proposons d'exprimer cette courbe sous la forme résolue habituelle, en calculant y en fonction de x, au moins au voisinage du point (a,b). *

Existence de la fonction implicite

Théorème 25 - Soient E un espace topologique, F et G des espaces affines normés, Ω un ouvert de $E \times F$, (a,b) un point de Ω. Soit f une application continue de Ω dans G, $f(a,b) = c$. On suppose que, pour tout x fixé, f admette une dérivée partielle $\frac{\partial f}{\partial y}(x,y) \in \mathcal{L}(\vec{F};\vec{G})$, et que la fonction dérivée partielle $\frac{\partial f}{\partial y}$ soit continue de Ω dans $\mathcal{L}(\vec{F};\vec{G})$. On suppose en outre que $Q = \frac{\partial f}{\partial y}(a,b)$ soit une application inversible

2

* On voit pourquoi de telles restrictions de voisinages sont inévitables ; une courbe $f(x,y) = o$ ne peut pas, en général, être exprimée tout entière sous la forme $y = g(x)$.

de $\vec{F}$ sur $\vec{G}$, c'est-à-dire une bijection ayant une bijection réciproque Q^{-1} linéaire et continue. On suppose enfin F complet. Alors il existe des ouverts A et B de E et F contenant a et b respectivement, tels que, pour tout x de A l'équation en y (III,8;1), ait une solution et une seule dans B. Cette solution y est fonction de x et la fonction $y = g(x)$ ainsi définie est continue de A dans B.

Démonstration - Remarquons d'abord rapidement qu'on se souvient facilement des diverses conditions restrictives apportées à la fonction f. On cherche à résoudre une équation en y, or on suppose des propriétés re'latives à la dérivée partielle par rapport à y, et c'est l'espace F de la variable y qu'on suppose complet. Nous allons faire subir à l'équation trois modifications successives.

1") Posons $y = b + \vec{Y}$, et appelons $\vec{f_1}$ l'application de $\vec{\Omega}_1$ dans $\vec{G}$ (où $\vec{\Omega}_1$ est l'ensemble des $(x, \overrightarrow{y-b}), (x,y) \in \Omega)$ définie par

$$\overrightarrow{f_1(x,\vec{Y})} = \overrightarrow{f(x, b+\vec{Y})} - \vec{c}\,. \quad (*) \tag{III,8;3}$$

L'équation (III,8;1) est équivalente à

$$\overrightarrow{f_1(x,\vec{Y})} = \vec{0}\,, \tag{III,8;4}$$

au voisinage de la solution particulière $(a, \vec{0})$.

2°) Considérons la fonction $\vec{f_2}$ définie par :

$$Q^{-1} \circ \vec{f_1} : (x,\vec{Y}) \longrightarrow Q^{-1}\left(\vec{f_1}(x,\vec{Y})\right). \tag{III,8;5}$$

C'est une application de $\vec{\Omega}_1$ dans $\vec{F}$ possédant des propriétés analogues à celles de f ; toutefois sa dérivée partielle par rapport à la deuxième variable, à l'origine, est cette fois, en vertu du corollaire 1 théorème 11, donnée par la

* Dans la pratique, F n'est pas nécessairement un espace vectoriel, mals G est presque toujours un espace vectoriel, e on a une équation $\vec{f}(x, y) = \vec{0}$. Nous avons pris G affine etc quelconque pour maintenir une symétrie entre F et G, et de toute façon on peut toujours, comme nous le voyons se ramener au cas où F et G sont vectoriels, avec $\vec{b} = \vec{0}$, $\vec{c} = \vec{0}$.

formule :

(III,8;6) $$\frac{\partial f_2}{\partial \vec{Y}}(a,\vec{o}) = Q^{-1} \circ Q = I \in \mathcal{L}(\vec{F};\vec{F}) :$$

c'est l'application identique. L'équation (III,8;1) est équivalente à la nouvelle équation

(III,8;7) $$f_2(x,\vec{Y}) = \vec{O} ,$$

car on passe de (III,8;4) a (III,8;7) en prenant l'image par Q^{-1} et de (III,8;7) à (III,8;4) en prenant l'image par Q .

3°) Considérons enfin la fonction $\vec{\Phi}$ définie par la formule :

(III,8;8) $$\vec{\Phi}(\lambda,\vec{Y}) = Y - f_2(A ,Y>.$$

$\vec{\Phi}$ est encore une application de $\vec{\Omega}_1$ dans $\vec{F}$ possédant des propriétés analogues à celles de f , sauf'toutefois sur un point particulier : son application derivee partielle par rapport à la deuxième variable, à l'origine, est nulle, puisqu'elle est donnée par r-1 = 0 . Par ailleurs l'équation (III,8;1) est maintenant équivalente à l'équation :

(III,8;9) $$\Phi(\lambda,\vec{Y}) = \vec{Y} , \quad \lambda = x , \vec{Y} = y - e ,$$

à résoudre au voisinage de $\lambda = a$, $\vec{Y} = \vec{o}$.

On voit que nous nous trouvons dans des conditions qui rappellent les théorèmes 45 et 46 bis du chapitre II (Théorèmes du point fixe). Soient A_1 et $\vec{B}_1$ des voisinages de a et $\vec{o}$ dans E etf respectivement, tels que $A_1 \times \vec{B}_1 \subset \vec{\Omega}_1$. Les espaces E et Λ du théorème 46 bis étant respectivement remplacés par $\vec{B}_1$ et A_1 , la fonction $\vec{\Phi}$ est oontinue sur $A_1 \times \vec{B}_1$ donc à fortiori séparément continue. Il existe cependant, pour pouvoir appliquer ces théorèmes, des conditions qui ne sont pas ici directement vérifiées et qui exigent encore une modification des données du problème.

1°/ L'application Φ devrait être une application de $A_1 \times \vec{B}_1$ dans $\vec{B}_1$, or elle applique $A_1 \times \vec{B}_1$ dans $\vec{F}$, mais non nécessairement dans $\vec{B}_1$. C'est ce qui va nous obliger à remplacer A_1 et $\vec{B}_1$ par deux voisinages plus petits A_2 et B_2 de a et 0 , de façon que la restriction de Φ a $A_2 \times \vec{B}_2$ applique cet ensemble dans $\vec{B}_2$.

2°/ Il faudra alors vérifier que, pour λ fixé dans A_1, l'application $\vec{\Phi}_\lambda : \vec{Y} \longrightarrow \vec{\Phi}(\lambda,\vec{Y})$, de $\vec{B}_2$ dans lui-même, est une contraction, avec un rapport de contraction $K < 1$ indépendant du choix de λ dans A_1.

3°) Il faut que l'espace métrique $\vec{B}_2$ soit complet.

Occupons nous d'abord de la condition 2°). Puisque la dérivée partielle $\frac{\partial\vec{\Phi}}{\partial\vec{Y}}$ est continue et qu'elle est nulle en $(a,\vec{0})$, on peut trouver un voisinage A'_2 de a dans A, et une boule <u>fermée</u> * $\vec{B}_2$ de centre origine et de rayon $\beta > 0$ dans B_1, tels que les relations : $\lambda \in A'_2$, $\vec{Y} \in \vec{B}_2$ entraînent l'inégalité :

$$\left\| \frac{\partial\vec{\Phi}}{\partial\vec{Y}}(\vec{\lambda},\vec{Y}) \right\| \leqslant \frac{1}{2} . \tag{III,8;10}$$

Si alors nous appliquons le théorème des accroissements finis, applicable puisque la boule $\vec{B}_2$ est convexe, et par conséquent, si $\vec{Y}'$ et $\vec{Y}''$ sont dans cette boule, il en est de même de tout le segment $[\vec{Y}',\vec{Y}'']$, on a l'inégalité :

$$\| \vec{\Phi}(\lambda,\vec{Y}') - \vec{\Phi}(\lambda,\vec{Y}'') \| \leqslant \frac{1}{2} \| \vec{Y}' - \vec{Y}'' \| . \tag{III,8;11}$$

Ainsi, avec A'_2 et B_2 ou des ensembles plus petits, la condition 2°) est réalisée, avec $k = \frac{1}{2}$.

Occupons-nous maintenant de la condition 1°). Nous voulons choisir A_1 et $\vec{B}_2$ de façon que $\vec{\Phi}$ soit une application de $A_2 \times \vec{B}_2$ dans $\vec{B}_2$. Les voisinages A'_2 et $\vec{B}_2$ ayant été choisis, et compte tenu de la continuité de $\vec{\Phi}$ et de ce que $\vec{\Phi}(a,\vec{0}) = \vec{0}$, nous pouvons trouver un voisinage $A_1 \subset A'_2$ de a, tel que l'on ait l'inégalité

$$\| \vec{\Phi}(\lambda,\vec{0}) \| \leqslant \frac{\beta}{2} , \text{ pour } \lambda \in A_2 . \tag{III,8;12}$$

Dans ces conditions, pour $\lambda \in A_2$ et $\|\vec{Y}\| \leqslant \beta$, on a l'inégalité :

$$\| \vec{\Phi}(\lambda,\vec{Y}) \| \leqslant \| \vec{\Phi}(\lambda,\vec{0}) \| + \| \vec{\Phi}(\lambda,\vec{Y}) - \vec{\Phi}(\lambda,\vec{0}) \| \leqslant \frac{\beta}{2} + \frac{1}{2}\|\vec{Y}\| \leqslant \beta . \tag{III,8;13}$$

Ceci nous prouve que l'application $\vec{\Phi}$ applique bien $A_2 \times \vec{B}_2$ dans $\vec{B}_2$. Enfin la condition 3°) est alors réalisée, puisque $\vec{B}_2$, partie fermée de l'espace métrique complet $\vec{F}$, est un

.* Nous verrons un peu plus bas pourquoi B_2 doit être fermée.

espace métrique complet, d'après le théorème 43 du chapitre II Nous nous trouvons donc exactement dans les conditions d'application des théorèmes 46 et 46 bis.
Donc pour $\vec{\lambda}$ donné dans A_1, il y a une solution et une seule en $\vec{Y}$ de (III,8;9) telle, que $\|\vec{Y}\| \leq \beta$, et $\vec{Y}$ est une fonction continue de λ, $Y = \Psi(\lambda)$.

L'équation (III,8;1), équivalente à l'équation (III,8;9), admet donc une solution et une seule en y, dans la boule B_2 de centre b et de rayon β de F, pour tout x donné dans A_2. De plus, si nous appelons g l'application qui, à tout points de A_2, fait correspondre l'unique solution y de B_2, $y = g(x)$, alors on a $g(x) = b + \Psi(x)$, donc g est une application continue de A_2 dans B_2.

Remarquons que la détermination des voisinages A, et B_2 se fait en deux fois; on détermine d'abord A'_2 et B_2 assez petits; naturellement on peut encore, si l'on veut, les rapetisser autant que 1 on voudra; par contre, ensuite, une fois que B_2 a été fixé, on doit encore choisir A_2 suffisamment petit en fonction de B_2 C'est d'ailleurs a posteriori, un résultat évident, à Cause de la continuité de la fonction g obtenue. Il résulte aussi de cette continuité que, si B est l'intérieur $\mathring{B}_2$ de B_2, il existe un ouvert A * contenant a dans E, tel que $x \in A$ entraîne $g(x) \in B$ et le théorème est démontré.

Remarque - Considdrons l'équation, en apparence plus générale :

$$f(x,y) = z \tag{III,8;14}$$

où l'on cherche à trouver y, voisin de b, en fonction de x et z, voisinsde a et c, avec $f(a,b) = c$.

Elle se ramène en fait au théorème 25, en considérant l'équation

$$\vec{f}_1((x,z),y) = \vec{0}, \quad \text{où} \quad \vec{f}_1((x,z),y) = \overrightarrow{f(x,y)-z} \; ; \tag{III,8;15}$$

f_1 est une application d'un ouvert de $(E\times G)\times F$ dans $\vec{G}$, et on cherche y, voisin de b, en fonction de (x,z) voisin de (a,c), avec $\vec{f}_1((a,c),b) = \vec{0}$.

* Pour faire la démonstration, on est d'abord obligé de prendre une boule fermée B_2, car nous avons supposé $\vec{B}_2$ fermée, pour qu'elle soit complète ! Mais dans la pratique, c'est l'énonce relatif à des ouverts A et B qui est le plus utilisé.

Dérivabilité de la fonction implicite

Nous allons maintenant supposer la fonction f dérivable et voir si la fonction implicite g définie par f est aussi dérivable. Nous nous supposerons dans des conditions qui peuvent être un peu plus générales que celles du théorème : nous supposerons l'existence d'une fonction implicite et sa continuité, même si cette existence et cette continuité peuvent ne pas résulter directement de l'application du théorème 25 (nous ne supposons pas l'existence de dérivées ailleurs qu'au point (a,b)). Par contre, E sera supposé affine, pour qu'on puisse parler de la dérivabilité de g.

Théorème 26 - Soient E, F, G, trois espaces affines normés, Ω un ouvert de $E\times F$, f une application de Ω dans G. Soient A et B des ouverts de E et F respectivement, $A\times B\subset\Omega$, et g application de A dans B, vérifiant identiquement (III,8;2). Si alors, au point (a,b), $b=g(a)$, l'application f est dérivable, si P et Q sont ses dérivées partielles au point considéré, applications linéaires continues respectivement de $\vec{E}$ et $\vec{F}$ dans $\vec{G}$, si Q est inversible, si enfin g est continue en a, alors g est dérivable en a, et son application dérivée est donnée par la formule :

$$g'(a) = -Q^{-1}\circ P = -\left(\frac{\partial f}{\partial y}(a,b)\right)^{-1}\circ\left(\frac{\partial f}{\partial x}(a,b)\right). \tag{III,8;16}$$

Démonstration. Si l'on supposait connue à l'avance l'existence d'une application dérivée de g, la formule (III,8;16) s'obtiendrait immédiatement. En effet, nous pouvons différencier l'identité (III,8;9), en appliquant le théorème des fonctions composées, et obtenir par conséquent l'équation :

$$\begin{cases}\dfrac{\partial f}{\partial x}(a,b) + \dfrac{\partial f}{\partial y}(a,b)\circ g'(a) = 0 \in \mathcal{L}(\vec{E};\vec{G}), \text{ ou} \\ P + Q\circ g'(a) = 0, \end{cases} \tag{III,8;17}$$

d'où l'on déduit (III,8 16) en composant avec Q^{-1} à gauche. Mais la démonstration est nécessairement un peu plus compliquée, car nous ne savons pas à l'avance que la fonction g est dérivable au point a. Soit donc $\overrightarrow{dx} = \overrightarrow{x-a}$ un accroissement de x, et soit $\overrightarrow{\Delta y}$ l'accroissement correspondant de $y=g(x)$. Du fait que g est toujours fonction implicite définie par l'équation considérée, et que f est dérivable en (a,b), on a la formule :

$$\vec{0} = \overrightarrow{\Delta f} = P.\overrightarrow{dx} + Q.\overrightarrow{\Delta y} + \vec{\alpha}\left(\|\overrightarrow{dx}\| + \|\overrightarrow{\Delta y}\|\right). \tag{III,8;18}$$

Le vecteur $\vec{\alpha}$ dépend de $\overrightarrow{dx}$ et de $\overrightarrow{\Delta y}$, donc finalement de $\overrightarrow{dx}$ seul; il tend vers $\vec{0}$ si $\overrightarrow{dx}$ et $\overrightarrow{\Delta y}$ tendent vers $\vec{0}$, donc simplement si $\overrightarrow{dx}$ tend vers $\vec{0}$, car $\overrightarrow{\Delta y}$ tend alors aussi vers $\vec{0}$ en vertu de la continuité supposée de g au point a . On en déduit le calcul suivant pour $\overrightarrow{\Delta y}$:

(III,8;19)
$$\begin{cases} Q\cdot\overrightarrow{\Delta y} = -P.\overrightarrow{dx} - \vec{\alpha}\left(\|\overrightarrow{dx}\| + \|\overrightarrow{\Delta y}\|\right) & \text{ou} \\ \overrightarrow{\Delta y} = -(Q^{-1}\circ P)\cdot\overrightarrow{dx} - (Q^{-1}\cdot\alpha)\left(\|\overrightarrow{dx}\| + \|\overrightarrow{\Delta y}\|\right). \end{cases}$$

Bien noter que $\overrightarrow{\Delta y}$ figure aussi au 2ème membre, il faut évidemment s'en débarraser. C'est ce que nous ferons par une majoration.

De ce calcul, déduisons d'abord une majoration :

(III,8;20)
$$\|\overrightarrow{\Delta y}\| \leqslant \|Q^{-1}\|\,\|P\|\,\|\overrightarrow{dx}\| + \|Q^{-1}\|\,\|\vec{\alpha}\|\,\|\overrightarrow{dx}\| + \|Q^{-1}\|\,\|\vec{\alpha}\|\,\|\overrightarrow{\Delta y}\|.$$

Mais, lorsque $\overrightarrow{dx}$ tend vers $\vec{0}$, nous avons vu que $\vec{\alpha}$ tend vers $\vec{0}$; on peut donc choisir $\overrightarrow{dx}$ assez petit, pour que le coefficient $\|Q^{-1}\|\,\|\vec{\alpha}\|$ soit majoré par $\frac{1}{2}$. En faisant alors passer dans le premier membre la dernière expression, on en déduit la majoration, valable pour $\|\overrightarrow{dx}\|$ assez petit

(III,8;21)
$$\begin{cases} \frac{1}{2}\|\overrightarrow{\Delta y}\| \leqslant \left(\|Q^{-1}\|\;\|P\| + \frac{1}{2}\right)\|\overrightarrow{dx}\| \quad , & \text{ou} \\ \|\overrightarrow{\Delta y}\| \leqslant (2\|Q^{-1}\|\,\|P\| + 1)\|\overrightarrow{dx}\| \quad , & \text{ou} \\ \left(\|\overrightarrow{dx}\| + \|\overrightarrow{\Delta y}\|\right) \leqslant \left(2\|Q^{-1}\|\;\|P\| + 2\right)\|\overrightarrow{dx}\| = k\,\|\overrightarrow{dx}\| \;. \end{cases}$$

où k est une constante fixe. En tenant alors compte de cette majoration dans le dernier terme du 2ème membre de (III,8;19) On obtient finalement :

(III,8;22)
$$\overrightarrow{\Delta y} = -\left(Q^{-1}\circ P\right).\overrightarrow{dx} \;\; \vec{\beta}\,\|\overrightarrow{dx}\|,\ \|\vec{\beta}\| \leqslant k\,\|Q^{-1}\|\;\;\|\vec{\alpha}\|,$$

de sorte que $\vec{\beta}$ tend vers $\vec{0}$ lorsque $\overrightarrow{dx}$ tend vers 0 , ce qui démontre le théorème.

<u>Remarque.</u> Avec la notation différentielle, on retient très facilement la règle :

<u>Règle</u> <u>On différentie l'équation</u> $f(x,y) = c$, <u>ce qui donne</u>

(III,8;23)
$$P.\overrightarrow{dx} \;\; t \;\; Q\cdot\overrightarrow{dy} = 0.$$

On trouve la différentielle de y, fonction implicite de x définie par l'équation, en résolvant (III,8;23) par rapport à $\overrightarrow{dy}$:

(III,8;24) $$\overrightarrow{dy} = -(Q^{-1} \circ P) \cdot \overrightarrow{dx} .$$

Dérivabilité de la fonction $u \to u^{-1}$ sur $\mathcal{L}(\vec{F};\vec{G})$

Pour pouvoir utiliser pleinement les théorèmes 25 et 26, il nous faut savoir comment varie $\left(\frac{\partial f}{\partial y}(x,y)\right)^{-1}$ quand (x,y) varie. De même que la dérivabilité des fonctions composées (théorème 19) utilisait la dérivabilité du produit (théorème 18), celle des fonctions implicites utilisera celle de l'inverse.

Théorème 27 - Soient $\vec{F}$, $\vec{G}$, des espaces vectoriels normés, et soit $\mathcal{U}$ (resp. $\mathcal{U}^{-1}$) l'ensemble des **elements** inversibles de $\mathcal{L}(\vec{F};\vec{G})$ (resp. $\mathcal{L}(\vec{G};\vec{F})$) . Alors :

1°) La bijection $u \longrightarrow u^{-1}$ de $\mathcal{U}$ sur $\mathcal{U}^{-1}$ est un homéomorphisme.

2°) Si $\vec{F}$ et $\vec{G}$ sont complets, $\mathcal{U}$ et $\mathcal{U}^{-1}$ sont ouverts, et la bijection $u \longrightarrow u^{-1}$ est dérivable ainsi que sa bijection réciproque, et sa dérivée au point $u \in \mathcal{L}(\vec{F},\vec{G})$ est l'application (III,8 25) $du \to -u^{-1} \circ du \circ u^{-1}$, de $\mathcal{L}(\vec{F};\vec{G})$ dans $\mathcal{L}(\vec{G};\vec{F})$.

Première démonstration

Pour simplifier, supposons dans tous les cas $\vec{F}$ et $\vec{G}$ complets. On a alors vu (théorème 62) que, si $u_0 \in \mathcal{U}$, tout élément de $\mathcal{L}(\vec{F};\vec{G})$, appartenant à la boule ouverte de centre u_0 et de rayon $\|u_0^{-1}\|^{-1}$, **était** encore dans $\mathcal{U}$; donc $\mathcal{U}$ est ouvert, et aussi $\mathcal{U}^{-1}$ par échange des rôles de $\vec{F}$ et $\vec{G}$.

Montrons maintenant que $u \longrightarrow u^{-1}$ est dérivable au point u_0 de $\mathcal{U}$; elle sera à fortiori continue, donc, comme l'application réciproque est la même à l'échange près de $\vec{F}$ et de $\vec{G}$, elle sera un homéomorphisme, dérivable ainsi que l'homéomorphisme réciproque. Le formule (II,14;26) nous dit que, pour $\|du\| < \|u_0^{-1}\|^{-1}$, on a :

(III,8;26) $$(u_0+du)^{-1} = u_0^{-1} - u_0^{-1}\,du\,u_0^{-1} + u_0^{-1}\,du\,u_0^{-1}\,du\,u_0^{-1} - \cdots$$

On en **déduit** la majoration :

(III,8;27)
$$\left\|\left((u_0+du)^{-1}-u_0^{-1}\right)+u_0^{-1}\,du\,u_0^{-1}\right\|$$
$$\leq \|u_0^{-1}\|^3\ \|du\|^2\left(1+\|u_0^{-1}\|\ \|du\|+\left(\|u_0^{-1}\|\ \|du\|\right)^2+\cdots\right)$$
$$=\frac{\|u_0^{-1}\|^3\ \|du\|^2}{1-\|u_0^{-1}\|\ \|du\|}=\frac{\|u_0^{-1}\|^2\ \|du\|^2}{\|u_0^{-1}\|^{-1}-\|du\|}\ .$$

Lorsque $\|du\|$ tend vers 0 , cette expression est majorée par constante $\times\|du\|^2$, donc infiniment petite devant $\|du\|$. Comme $du \longrightarrow -u_0^{-1}\,du\,u_0^{-1}$ est une application linéaire continue de $\mathcal{L}(\vec{F};\vec{G})$ dans $\mathcal{L}(\vec{G};\vec{F})$, c'est bien la dérivée de l'application $u \longrightarrow u^{-1}$ au point u_0 , en vertu de la définition (III,3;13) * .

Deuxième démonstration

Bornons-nous encore au cas où $\vec{F}$ et $\vec{G}$ sont complets, alors $\mathcal{L}(\vec{F};\vec{G})$ et $\mathcal{L}(\vec{G};\vec{F})$ le sont aussi (théorème 50 du chapitre II). Soit $u_0 \in \mathcal{U}$. On appelle Inverse à droite de $u \in \mathcal{L}(\vec{F};\vec{G})$ un élément v de $\mathcal{L}(\vec{G};\vec{F})$ tel que $u \circ v = I$. Un tel élément v n'existe pas toujours (par exemple si $u = 0$); il peut en exister une infinité; mais, si u est inversible, il n'en existe qu'un, et c'est u^{-1} , car de $uv = I$ on déduit, en composant à gauche avec u^{-1}, $v = u^{-1}$.

La recherche d'un Inverse à droite de u est la résolution d'une équation en v, $uv = I$. Nous pouvons donc lui appliquer le théorème des fonctions implicites. L'application

* Notons que, si $\vec{F}$ et $\vec{G}$ sont arbitraires, $\mathcal{U}$ et $\mathcal{U}^{-1}$ sont vides ! Par exemple, si $\vec{F}$ et $\vec{G}$ sont de dimension finie, il n'existe d'applications inversibles de $\vec{F}$ sur $\vec{G}$ que si $\vec{F}$ et $\vec{G}$ ont la même dimension.

Noter aussi que, si $\mathcal{U}$ n'est pas vide, alors $\vec{F}$ ne peut être complet sans que $\vec{G}$ le soit, et vice-versa; car un élément de $\mathcal{U}$ donne une correspondance biunivoque entre $\vec{F}$ et $\vec{G}$, conservant la convergence et les suites de Cauchy.

$(u,v) \longrightarrow uv$ est continuement dérivable, comme application-, **bilinéaire continue** (théorème 9 bis). Au point (u_0, v_0) , $v_0 = u_0$, **la dérivée partielle** en v **est l'application** linéaire continue $V \longrightarrow W = u_0 V$ * de $\mathcal{L}(\vec{G};\vec{F})$ **dans** $\mathcal{L}(\vec{G};\vec{G})$ (formule (**II,3;34**)); cette application est **inversible**, **et** sa réciproque est l' application continue $W \longrightarrow V = u_0^{-1} W$ **de** $\mathcal{L}(\vec{G};\vec{G})$ **dans** $\mathcal{L}(\vec{G};\vec{F})$. Le théorème 25 montre alors **qu'il** existe un voisinage **ouvert** $(\mathcal{U}_0)_d$ de u_0 dans $\mathcal{L}(\vec{F};\vec{G})$ et un voisinage ouvert $(\mathcal{U}_0)_d^{-1}$ de u_0^{-1} dans $\mathcal{L}(\vec{G};\vec{F})$, tel que toutu de $(\mathcal{U}_0)_d$ ait un inverse à droite u_d^{-1} et un seul qui soit **situé** dans $(\mathcal{U}_0)_d^{-1}$; en outre l'application $u \longrightarrow u_d^{-1}$ ainsi définie est continue de $(\mathcal{U}_0)_d$ dans $(\mathcal{U}_0)_d^{-1}$, et dérlvable au point u_0 , d'après le théorème 26. Mais on peut faire le même raisonnement avec l'inverse **à** gauche; un inverse **à** gauche w de $u \in \mathcal{L}(\vec{F};\vec{G})$ est un **élément** de $\mathcal{L}(\vec{G};\vec{F})$ vérifiant $wu = I$. On détermine **alors** des ouverts $(\mathcal{U}_0)_g$ et $(\mathcal{U}_0)_g^{-1}$ et une application $u \longrightarrow u_g^{-1}$ de $(\mathcal{U}_0)_g$ dans $(\mathcal{U}_0)_g^{-1}$. Soit alors $\mathcal{U}_0 = (\mathcal{U}_0)_d \cap (\mathcal{U}_0)_g$. **C'est un** voisinage ouvert **de** u_0 . **Si** $u \in \mathcal{U}_0$, il a **à** la **fois un** Inverse **à** gauche u_g^{-1} **dans** $(\mathcal{U}_0)_g^{-1}$ et un inverse **à** droite u_d^{-1} dans $(\mathcal{U}_0)_d^{-1}$. Mais si un **élément** admet **à** la fois un Inverse **à** gauche **et un** Inverse **à** droite, ceux-cl **coincident**, et cet élément est alors inversible; car de $u u_d^{-1} = I$ **et** $u_g^{-1} u = I$ on déduit $u_g^{-1} u\, u_d^{-1} = (u_g^{-1} u) u_d^{-1} = u_d^{-1}$ et aussi $= u_g^{-1}(u\, u_d^{-1}) = u_g^{-1}$.

De cela nous déduisons d'abord que tout élément de $\mathcal{U}_0$ est inversible; cela prouve que $\mathcal{U}$ est un voisinage de chacun de ses **points, donc** est ouvert, et par suite aussi $\mathcal{U}^{-1}$ **par** échange de $\vec{F}$ et $\vec{G}$. Ensuite l'application $u \longrightarrow u^{-1}$, qui, dans le voisinage $\mathcal{U}_0$ du point u_0 **coincide** avec $u \longrightarrow u_d^{-1}$, est continue au point u_0 , donc continue partout **sur** $\mathcal{U}$; son application réciproque a la **même** propriété par échange de $\vec{F}$ et $\vec{G}$, donc c'est un **homéomorphlsme de** $\mathcal{U}$ sur $\mathcal{U}^{-1}$. Enfin, pour la **même** raison, elle est dérlvable, ainsi **que** son application réciproque.

Le calcul de la **dérivée** se fait par la règle (**III,8;23**) En **différentiant** $uv = I$, on a

* En notation **différentielle**, $d(uv) = du\, v_0$ t $u_0\, dv$, et la **dérivée** partielle **en** v est $dv \longrightarrow u_0\, dv$, au point (u_0, v_0).

(III,8;31) $$du \circ v + u \circ dv = 0$$

ce qui redonne, en composant à gauche avec u^{-1} :

(III,8;32) $$dv = - u^{-1} \circ du \circ v = - u^{-1} \circ du \circ u^{-1} .$$

Troisième démonstration

Cette démonstration n'est valable que si $\vec{F}$ et $\vec{G}$ sont de dimension finie; auquel cas on peut les supposer de même dimension% * . Prenons un référentiel dans chacun d'eux; on les identifie ainsi à $\mathbb{K}^n$, et $\mathcal{L}(\vec{F};\vec{G})$ et $\mathcal{L}(\vec{G};\vec{F})$ sont identifiés à l'espace vectoriel $\mathbb{K}^{n^2}$ des matrices à n lignes et n colonnes. Soit $\mathfrak{M}$ une telle matrice; appelons m_{ij} ses termes, $i = 1,2,\dots n$, $j = 1,2,\dots n$; une fonction $\mathfrak{M} \to f(\mathfrak{M})$ est de classe C^m si et seulement si c'est une fonction de classe C^m des n^2 variables scalaires m_{ij}. Le déterminant $\mathfrak{M} \to \det \mathfrak{M}$ est un polynome par rapport aux m_{ij} , donc c'est une fonction scalaire indéfiniment dérivable de $\mathfrak{M}$; l'ensemble% des matrices inversibles est l'image réciproque de $\complement 0$ par cette application continue, car une matrice est inversible si et seulement si son déterminant est $\neq 0$; or $\complement 0$ est ouvert dans $\mathbb{K}$, donc $\mathcal{U}$ est ouvert.

Alors l'inverse $\mathfrak{M}^{-1}$ de $\mathfrak{M}$ est la matrice dont les termes sont les $\frac{M_{ji}}{\det \mathfrak{M}}$, M_{ji} étant le mineur d'indlces j, i, de $\mathfrak{M}$;

M_{ji} est un polynome par rapport aux m_{ij} , donc $\frac{M_{ji}}{\det \mathfrak{M}}$ est une fonction scalaire indéfiniment dérivable sur $\mathcal{U}$; alors $\mathfrak{M} \longrightarrow \mathfrak{M}^{-1}$ est une application de $\mathcal{U}$ dans $\mathbb{K}^{n^2}$ dont les composantes sont des fonctions scalaires de classe C^∞ , donc elle est de classe C^∞ . Comme elle est sa propre application réciproque, c'est un homéomorphisme de classe C^∞ , et nous avons ainsi démontré plus que ne comportait l'énoncé. Voir à ce sujet la remarque après le corollaire qui suit.

Pour le calcul de la dérivée de l'application $u \longrightarrow u^{-1}$ le mieux est d'utiliser la fin de la 2ème démonstration. Si nous revenons à cette démonstration, signalons que, <u>dans le cas de la dimension finie</u>, si u possède un inverse à droite u_d^{-1} , on a, en passant aux matrices, $\mathfrak{M}\mathfrak{M}_d^{-1} = I$, donc $(\det \mathfrak{M})\,(\det \mathfrak{M}_d^{-1}) = 1$, donc $\det \mathfrak{M} \neq 0$, et on sait qu'alors u est inversible (et d'inverse u_d^{-1}).

* Voir note (*) page **286**; s'ils n'ont pas même dimension, le problème est sans objet.

Corollaire - soit $x \longrightarrow u(x)$ une application dérivable d'un ouvert Ω d'un espace affine E dans $\mathcal{L}(\vec{F};\vec{G})$ ($\vec{F}$, $\vec{G}$, espaces de Banach.

Si, pour tout x de Ω, $u(x)$ est un élément inversible de $\mathcal{L}(\vec{F};\vec{G})$, alors $x \longrightarrow u^{-1}(x)$ est une application dérivable de Ω dans $\mathcal{L}(\vec{G};\vec{F})$, et sa ddrivée est donnée, pour $\vec{X} \in \vec{E}$, par :

(III,8;33) $$(u^{-1})'(x).\vec{X} = -u^{-1}(x) \circ (u'(x)\cdot\vec{X}) \circ u^{-1}(x) \in \mathcal{L}(\vec{G};\vec{F}).$$

En particulier, si $\mathcal{M}$ est une matrice carrée à n lignes et n colonnes, dépendant d'un paramètre scalaire t, et dérivable en t, si d'autre part elle est inversible pour tout t, la dérivée de la matrice inverse $\mathcal{M}^{-1}$ est donnée par la formule :

(III,8;34) $$\frac{d(\mathcal{M}^{-1})}{dt} = -\mathcal{M}^{-1}\frac{d\mathcal{M}}{dt}\mathcal{M}^{-1}.$$

Il suffit en effet d'appliquer le corollaire à $E = \mathbb{K}$, $\vec{F} = \vec{G} = \mathbb{K}^n$. Pour $n = 1$, on retrouve la dérivée de l'inverse d'une fonction scalaire f partout $\neq 0$:

(III,8;35) $$\left(\frac{1}{f}\right)' = -\frac{f'}{f^2}.$$

Nous pouvons maintenant appliquer ce résultat aux fonctions implicites.

Théorème 28 - (Théorème des fonctions implicites) Soient E, F, G, 3 espaces affines normés, f une application continuement dérivable d'un ouvert Ω de $E \times F$ dans G. 1°) Soient A et B des ouverts de E et F respectivement, $A \times B \subset \Omega$, soit g une application continue de A dans B vérifiant (III,8;2). Si, pour tout x de A, $\frac{\partial f}{\partial y}(x, g(x))$ est un élément inversible de $\mathcal{L}(\vec{F};\vec{G})$, alors g est continuement dérivable de A dans B.

2°) Soit (a, b) un point de Ω, $f(a, b) = c$. Si $\frac{\partial f}{\partial y}(a, b)$ est un élment inversible de $\mathcal{L}(\vec{F};\vec{G})$, et si $\vec{F}$ est complet, il existe des ouverts A et B de E et F contenant a et b respectivement, $A \times B \subset \Omega$, tels que, pour tout $x \in A$, l'équation en y (III,8;1) ait une solution et une seule dans B et que la fonction $y = g(x)$ ainer définie soit continuement derivable de A dans B. Sa dérivee est donnée par (III,8;16).

Démonstration de 1°) Il résulte du théorème 26 que g est dérivable en tout point de A , et que sa dérivée en $x \in A$ est l'élément de $\mathcal{L}(\vec{E};\vec{F})$ donné par :

$$g'(x) = -\left(\frac{\partial f}{\partial y}(x,g(x))\right)^{-1} \circ \frac{\partial f}{\partial x}(x,g(x)). \tag{III,8;37}$$

L'application $x \longrightarrow g(x)$ est supposée continue sur A , et $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont continues sur Ω ; donc $x \longrightarrow \frac{\partial f}{\partial x}(x,g(x))$ et $x \longrightarrow \frac{\partial f}{\partial y}(x,g(x))$ sont continues de **A** dans $\mathcal{L}(\vec{E};\vec{G})$ et $\mathcal{L}(\vec{F};\vec{G})$ respectivement. Comme alors $u \to u^{-1}$ est continue de $\mathcal{U}$ dans $\mathcal{U}^{-1}$ (théorème 27), la fonction $x \longrightarrow \left(\frac{\partial f}{\partial y}(x,g(x))\right)^{-1}$ est continue de **A** dans $\mathcal{L}(\vec{G};\vec{F})$.

Alors, la composition $(u,v) \to v \circ u$ étant continue de $\mathcal{L}(\vec{E};\vec{G}) \times \mathcal{L}(\vec{G};\vec{F})$ dans $\mathcal{L}(\vec{E};\vec{F})$ (théorème 54 du chapitre II), g' est bien continue' de **A** dans $\mathcal{L}(\vec{E};\vec{F})$, g est **continuement** dérivable sur **A** .

Démonstration de 2°

Comme cette fois F est complet, **ainsi** que G (voir **note** * page 286 : on connait un élément $\frac{\partial f}{\partial y}(a,b)$ de $\mathcal{L}(\vec{F};\vec{G})$ **qui est inversible**), l'ensemble $\mathcal{U}$ des éléments inversibles de $\mathcal{L}(\vec{F};\vec{G})$ est ouvert; **comme** $(x,y) \longrightarrow \frac{\partial f}{\partial y}(x,y)$ est **supposée** continue **de** Ω dans $\mathcal{L}(\vec{F};\vec{G})$, l'image **réciproque** de $\mathcal{U}$ par cette application est un **ouvert** Ω_1 de Ω, contenant (a,b) . Le **théorème** 25 est applicable **à** la restriction de f à Ω_1 (toujours parce que F est complet). On peut donc **déterminer** des ouverts A et B de E et F contenant a et b respectivement, $A \times B \subset \Omega_1$, **tels** que, pour touts de A , il existe un **élément** et un seul y de B solution de (III,8;1), et que la fonction $y = g(x)$ ainsi définie de A dans B soit continue, Comme alors $(x, g(x)) \in A \times B \subset \Omega_1$, **pour** tout x de A , on se trouve dans les conditions de 1°, donc g est **continuement** dérivable; et le théorème 26 donne (III;8;16)

<u>Corollaire - **Si** $\vec{F}$ et $\vec{G}$ sont des espaces de **Banach,** $\mathcal{U}$ l'ouvert des **éléments inversibles** de $\mathcal{L}(\vec{F};\vec{G})$, **l'application** $u \longrightarrow u^{-1}$ **de** $\mathcal{U}$ sur $\mathcal{U}^{-1}$ **est** continuement **dérivable** ainsi que son **application réciproque**</u>

Il suffit d'appliquer le théorème 28, 1°, à l'équation $uv = I$, (voir théorème 27, deuxième **démonstration**). On peut aussi démontrer directement la **continuité** de la **dérivée** au théorème 27.

Cas particulier où E = F = G = K , corps des scalaires

Rappelons que, si l'on appelle $q = \frac{\partial f}{\partial y}(a,b)$ la dérivée partielle usuelle de la fonction f scalaire de deux variables scalaires, son application dérivée partielle Q est la multiplication par q ; par suite elle est inversible, si et seulement si la dérivée partielle usuelle q est distincte de 0 . Il en résulte que nous nous trouvons dans les conditions d'application du théorème 28 pour l'existence, l'unicité et la dérivabilité de la fonction implicite, dès que l'on suppose que f admet des dérivées partielles usuelles continues $\partial f/\partial x$, $\partial f/\partial y$ (f est alors totalement dérivable d'après le théorème 15) avec $q = \frac{\partial f}{\partial y}(a,b) \neq 0$. C'est dans de telles conditions que le théorème des fonctions implicites a été énoncé en Mathématiques Spéciales. La formule (III,8; 16) devient:

(III,8;38)
$$g'(a) = - \frac{\frac{\partial f}{\partial x}(a,b)}{\frac{\partial f}{\partial y}(a,b)} .$$

Supposons maintenant que $\frac{\partial f}{\partial y}(a,b) = 0$ mais que $\frac{\partial f}{\partial x}(a,b) \neq 0$. On peut alors, au voisinage de (a,b), calculer x en fonction de y, $x = h(y)$; on trouve que $h'(b) = 0$, la courbe d'équation $f(x,y) = c$ admet en (a,b) une tangente verticale, et l'impossibilité de calculer y comme fonction dérivable de x est à posteriori évidente.

Supposons enfin que $\frac{\partial f}{\partial x}(a,b) = \frac{\partial f}{\partial y}(a,b) = 0$. La courbe présente en (a,b) un point singulier. C'est peut-être un point singulier isolé (par exemple, si $K = \mathbb{R}$, $(a,b) = (0,0)$, $f(x,y) = x^2 + y^2$) ; alors l'étude de la courbe au voisinage de (a,b) est sans objet. Mais si la courbe a, par exemple, un "point double à tangentes distinctes", on peut espérer séparer, au voisinage de (a,b) , 2 branches distinctes, et, pour chacune d'elles, calculer y en fonctions de x , ou l'inverse - Pour simplifier, supposons f de classe C^3 , et $a = b = c = 0$. Au lieu de prendre x et y comme variables, prenons x et $m = \frac{y}{x}$. L'équation devient $F(x,m) = f(x,mx) = 0$; voyons si elle permet de trouver m en fonction de x . La fonction F est nulle, quelle que soit m, pour $x = 0$, et cela même si $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ ne sont pas simultanément nuls à l'origine.

Mais **ici**, en outre, $\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} + m\frac{\partial f}{\partial y}$ est nulle pour $x = 0$, quel que soit m . Nous montrerons **ultérieurement** (thdorème 119 du chapitre IV) que cela implique que $G(x,m) = \frac{F(x,m)}{x^2}$ soit de classe C^1 (moyennant une définition de la **valeur** de cette fonction, aux points où $x = 0$ **, par** passage **à** la limite).'

Supposons que le développement de Taylor de f à l'origine soit :

(III,8;38bis) $$f(x,y) = \frac{1}{2}(r x^2 + 2s x y + t y^2) + \cdots .$$

Alors, au point $x = 0$, $m = m_0$,on aura

(III,8;38ter) $$G(0, m_0) = \lim_{x \to 0} \frac{F(x, m_0 x)}{x^2} = \frac{1}{2}(r + 2s\, m_0 + t\, m_0^2).$$

Si alors le **trinome** $m \to r + 2sm + tm^2$ a 2 racines distinctes dans $\mathbb{K}$, $m = m_1$ et $m = m_2$, on voit que $G(0,m_1) = 0$, et $G(0,m_2) = 0$. Alors nous possédons de l'dquation $G(x,m) = 0$, deux solutions particulières, $(0, m_1)$ et $(0, m_2)$. Si $\frac{\partial G}{\partial m}(0, m_1)$ e t $\frac{\partial G}{\partial m}(0, m_2)$ **sont** $\neq 0$, le **théorème** des fonctions implicites nous permettra de **résoudre** $G(x,m) = 0$, au voisinage de ces 2 points, par des fonctions $m = \mu_1(x)$, $m = \mu_2(x)$, de classe C^1 ; et alors les 2 branches de courbe cherchées seront **déterminées** par le8 **équations** résolues de classe C^1 : $y = x\mu_1(x)$, $y = x\mu_2(x)$, donnant des tangentes **à** l'origine de coefficients angulaires m_1 **et** m_2 **.**

Or **précisément** on a sûrement $\frac{\partial G}{\partial m_1} \neq 0$ et $\frac{\partial G}{\partial m_2} \# 0$ **. Car**

(III,8;38quart) $$\frac{\partial G}{\partial m}(0, m_0) = s + m_0 t .$$

Comme le **trinôme** est suppose avoir 2 racines <u>distinctes</u>, sa dérivée enm ne saurait **être** nulle en l'une quelconque de ces racines !

<u>Ainsi le théorème des fonctions implicites s'applique même dans des cas qui, a priori, paraissent lui échapper</u> !

Cas où E, F, G sont de dimension finie

Supposons maintenant que E, F, G, soient des espaces affines de dimension finie: supposons choisies dans $\vec{E}, \vec{F}, \vec{G}$, des bases, $(\vec{e}_k)_{k\in K}$; $(\vec{f}_j)_{j\in J}$; $(\vec{g}_i)_{i\in I}$; et des origines dans E, F, G.

Une application f de $\Omega \subset E\times F$ dans G, est alors définie par les formules- :

(III,8;39) $$z_i = F_i\left((x_k)_{k\in K}, (y_j)_{j\in J}\right) ; \quad i \in I .$$

La contlnuitd de f exprime que les fonctions F_i sont des fonctions scalaires continues des variables x_k et y_j. Dire que f admet une application dérivée partielle par rapport à la deuxième variable, qui soit continue, c'est dire que les F_i ont des dérivées partielles du premier ordre par rapport aux variables y_j, continues (théorème 15). Il faut maintenant écrire que, au point (a, b, c) tel que $f(a,b) = c$, l'application dérivée partielle Q est inverslble. Cela exige d'une part que E et F aient même dimension. On peut alors supposer $I = J$. Ceci étant réalisé, il en sera ainsi si et seulement si le déterminant jacoblen, determinant des $\frac{\partial F_i}{\partial y_j}(a,b)$, est $\neq 0$, puisque c'est le déterminant de la matrice définissant l'application linéaire Q. SI ces conditions sont réalisées, alors on pourra (théorème 25), au voisinage du point (a,b), déterminer y comme fonction implicite $g(x)$, c'est-a-dire exprimer les y_j par des formules :

(III,8;40) $$y_j = G_j\left((x_k)_{k\in K}\right) ; \quad j \in J = I .$$

Si maintenant les F_i admettent non seulement des dérivées partielles par rapport aux y_j, mais aussi par rapport aux x_k, continues dans Ω, alors cn peut affirmer que l'application f est de classe C'. Nous pourrons alors appliquer le théorème 28, et la formule : (III,8;16), et trouver la matrice dérivée de g c'est-a-dire des y_j par rapport aux x_k, sous la forme d'un quotient de 2 matrices. si $I = J = \{1,2,\dots m\}$, $K = \{1,2,\dots n\}$:

(III,8;41) $$\begin{pmatrix} \frac{\partial G_1}{\partial x_1} & \frac{\partial G_1}{\partial x_2} & \cdots & \frac{\partial G_1}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial G_m}{\partial x_1} & \frac{\partial G_m}{\partial x_2} & \cdots & \frac{\partial G_m}{\partial x_n} \end{pmatrix} = - \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial F_m}{\partial y_1} & \frac{\partial F_m}{\partial y_2} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix} \Bigg| \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial F_m}{\partial x_1} & \frac{\partial F_m}{\partial x_2} & \cdots & \frac{\partial F_m}{\partial x_n} \end{pmatrix}$$

Au lieu de faire des calculs sur des **matrices, on pourra** aussi trouver les dérivées **partielles** $\frac{\partial G_j}{\partial x_k}$, k fixe , $j \in J$, en donnant **à** x un accroissement $\vec{dx}$ défini par $dx_{k'} = dx_k$ si $k' = k$, et $= 0$ si $k' \neq k$, et en calculant les différentielles dy_j , $j \in J$, par résolution d'un système dem équations linéaires aminconnues, cramérien puisque le **jacobien est** $\neq 0$:

(III,8;42) $$\frac{\partial F_i}{\partial x_k} dx_k + \sum_{j \in J} \frac{\partial F_i}{\partial y_j} dy_j = 0 \; ; \qquad i \in I .$$

Fonction réciproque comme fonction implicite

Trouver la fonction réciproque $x = g(y)$ d'un **homéomorphisme** $y = f(x)$, c'est résoudre l'équation en x : $y = f(x)$, pour tout y donné: c'est un problème de fonction implicite.

Théorème 29 - Soient E, F , des espaces affines normés complets, f une application continuement dérivable d'un **ouvert** Ω de E **dans** F .

1°/ Si a est un point **de** Ω et si $f'(a)$ est un élément inversible de $\mathcal{L}(\vec{E},\vec{F})$, alors l'image **par** f de tout voisinage de a est un voisinage de $f(a) = b$, et l'image **par** f de toute boule ouverte de centre a **contient** une boule **ouverte** de **centre** b. Il existe des ouverts A et B de E et F contenant a et b respectivement, tels que f soit un homéomorphisme de A sur **B**, continuement dérivable ainsi que son homéomorphisme réciproque.

On a en outre :

(III,8;43) $$(f^{-1})'(b) = \left(f'(a)\right)^{-1} .$$

2°/ Si, pour tout x de Ω , $f'(x)$ est un élément inversible de $\mathcal{L}(\vec{E};\vec{F})$, alors l'image pour f de tout ouvert de Ω est un ouvert, et en particulier $f(\Omega)$ est un ouvert de F . Si en outre f est **injective,** c'est un homéomorphisme de Ω sur $f(\Omega)$ continuement **dérivable** ainsi que son **homéomorphisme** réciproque.

Démonstration - Nous avons vu que l'application **réciproque** $g = f^{-1}$ de f **est** une fonction Implicite-s $= g(y)$ définie par l'équation $\vec{f_1}(x,y)$ 5 $\overrightarrow{f(x) - y} = \vec{0}$, f_1 étant une application de $\Omega \times F$ **dans** $\vec{F}$ **.** On va donc appliquer les théorèmes précédents, mais avec un changement inévitable de notations : c'est x qu'on cherche en fonction de y et non le contraire. C'est pourquoi on fait des hypothèses relatives à la dérivée **partielle** en x , qui est $\frac{\partial f_1}{\partial x} = \frac{\partial f}{\partial x}$; par ailleurs E est complet. $\vec{f_1}$ est bien continuement dérivable sur $\Omega \times F$, parce que ses **dérivées** partielles $\frac{\partial f_1}{\partial x}(x,y) = f'(x)$ et $\frac{\partial f_1}{\partial y} = -I$ sont continues (théorème 15).

Démontrons 1°/ - Les conditions nécessaires à l'application du théorème 28,2°, sont réalisées en (a, b). On peut donc trouver des ouverts, que nous appellerons A_0 et B,

$a \in A_0 \subset E$, $b \in B \subset F$, tels que l'équation $f(x) = y$,

pour $y \in B$, ait une solution unique x dans A_0, et que la fonction $x = g(y)$ ainsi définie soit continuement dérivable de B dans A_0. Mais les ouverts A_0 et B ainsi trouvés ne répondent pas directement aux conditions de l'énoncé. L'image $f(A_0)$ "recouvre B une fois et une seule" mais elle est plus grande, $f(A_0) \supset B$. Nous devons prendre

$A = A_0 \cap f^{-1}(B)$. Alors sûrement f applique A dans B; par ailleurs, pour $y \in B$, $g(y)$ est un point de A_0 qui est forcément aussi dans l'image réciproque $f^{-1}(B)$ (puisque

$f(g(y)) = y \in B$), donc dans A, donc g applique B dans A; f

et g sont deux applications réciproques, de A dans B et de B dans A. B est ouvert par construction, A comme intersection de 2 ouverts; f est continuement dérivable par définition, g d'après le théorème des fonctions implicites, on a donc bien deux homéomorphismes réciproques, continuement dérivables. Soit maintenant $\mathcal{V}$ un voisinage de a dans Ω; alors, $\mathcal{V} \cap A$ est un voisinage de a dans A, et comme f est un homéomorphisme de A sur B, $f(\mathcal{V} \cap A)$ est un voisinage de b dans B, donc aussi dans F puisque B est ouvert dans F; à fortiori $f(\mathcal{V})$ est un voisinage de b dans F. $\mathcal{V}$ peut être en particulier une boule ouverte de centre a; $f(\mathcal{V})$, voisinage de b, contient alors une boule ouverte de centre b. Ainsi 1°/ est démontré. La dérivée de g en b se calcule par la règle (III,8;23) : la résolution en $\overrightarrow{dx}$ de l'équation

$$f'(a) \cdot \overrightarrow{dx} - \overrightarrow{dy} = \vec{0} \tag{III,8;44}$$

donne

$$\overrightarrow{dx} = \left(f'(a)\right)^{-1} \cdot \overrightarrow{dy}, \tag{III,8;45}$$

d'où (III,8;43).

Démontrons maintenant 2°/ - Soit $\mathcal{O}$ un ouvert de Ω. Alors c'est un voisinage de chacun de ses points x, donc, $f'(x)$ étant inversible, $f(\mathcal{O})$ est un voisinage de $f(x)$ d'après 1°, donc de chacun de ces points; donc $f(\mathcal{O})$ est ouvert. En particulier, $f(\Omega)$ est ouvert. Si en outre f est injective, elle est alors bijective et continue de Ω sur $f(\Omega)$ et l'image directe de tout ouvert est un ouvert; d'après le théorème 11 du chapitre II f est un homéomorphisme. Si a est un point quelconque de Ω, alors dans l'ouvert B, déterminé à 1°, f^{-1} coïncide avec g, donc est continuement dérivable dans B, donc dans $f(\Omega)$ tout entier.

Remarque Le théorème 29,1°, aboutit à la même formule que le corollaire 4 du théorème 11; mais les deux énoncés sont un peu réciproques l'un de l'autre, puisque l'un part de deux homéomorphismes réciproques dérivables pour aboutir à l'inversibilité de $f'(a)$, l autre part de l'inversibilité de $f'(a)$ pour trouver deux homéomorphismes réciproques dérivables. Il est bien évident que l'énoncé 29,2°, est d'un caractère beaucoup plus délicat.

Si E et F sont de dimension finie, et si l'on y a choisi des référentiels, l'inversibilité de $f'(a)$ signifie simplement: que E et F ont même dimension et que le déterminant jacobien de f en a, par rapport aux référentiels, est $\neq 0$.

Le théorème 29 nous amène à poser les définitions suivantes :

Définitions - Soient Ω et F deux espaces topologiques, f une application continue de Ω dans F. On dit que f est une application ouverte, si l'image directe par f de tout ouvert de Ω est un ouvert de F. On dit que f est un homéomorphisme local de Ω dans F si, pour tout point a de Ω, il existe un ouvert A de Ω contenant a et un ouvert B de F contenant $f(a)$, tels que f soit un homeomorphisme de A sur B.

On démontre alors immédiatement ce qui suit :

A) Pour que f soit ouverte, il faut et il suffit que, pour tout point a de Ω, l'image par f de tout voisinage de a soit un voisinage de $f(a)$; ou encore que, pour tout a, l'image par f de toute boule ouverte de centre a contienne une boule ouverte de centre $f(a)$ (Nous avons démontré le "il suffit dans la démonstration du théorème 29, 2°, et le "il faut" est évident).

B) Tout homéomorphisme local est une application ouverte. C'est ce que nous avons montré dans la démonstration de 29,1° : quand f est un homeomorphisme local, alors, pour tout x de Ω, et tout voisinage $\mathcal{V}$ de x, $f(\mathcal{V})$ est un voisinage de $f(x)$; donc f est ouverte d'après A.

C) Si f est une application continuement dérivable d'un ouvert Ω de l'espace affine normé complet E dans l'espace affine normé F, et si, pour tout x de Ω, $f'(x)$ est inversible, alors f est un homéomorphisme local de Ω dans F. C'est ce qu'indique le théorème 23. En particulier, f est une application ouverte, et $f(\Omega)$ est un ouvert de F.

D) Une application8 ouverte injective est un homéomorphisme de Ω sur son image $f(\Omega)$. C'est ce que nous avons démontré dans la démonstration de' 29,2° .

Le tableau suivant résume ces résultats : les flèches indiquent les Implications. On suppose que f est une application continue d'un espace topologique Ω dans un espace topologique F

E,F sont affines normés complets, Ω ouvert de E, f continuement dérivable, $f'(x)$ inversible pour tout x de Ω .	$\Longrightarrow$	f est un homéomorphisme local	$\Longrightarrow$	f est ouverte
				$\Updownarrow$
				L'image par f de tout voisinage de x est un voisinage de $f(x)$

Il est bon de savoir qu'on ne peut guère aller plus loin, et qu'en particulier les réciproques des implications énoncées sont fausses :

Prenons C) . Si f est un homéomorphiame local de Ω dans F, continuement dérivable, sa dérivée $f'(x)$ n'est pas nécessairement inversible pour tout x de Ω . Par exemple $x \longrightarrow x^3$ est un homdomorphisme dérivable de $\mathbb{R}$ sur $\mathbb{R}$, maie sa dérivée est nulle à l'origine. Son homéomorphlsme réciproque, $y \longrightarrow \sqrt[3]{y}$, n'est pas dérivable à l'origine.

Prenons B) . Une application ouverte n'est pas nécessairement un hombomorphisme local: Par exemple, toute projection d'un produit d'espaces topologiques sur un des espaces facteurs, est ouverte, comme le montre immédiatement la définition des voisinages; ce n'est pas en général un homéomorphisme local (la projection $(x,y) \rightarrow x$ de $\mathbb{R}^2$ sur $\mathbb{R}$ n'est pas un homéomorphisme local !).

Enfin dans D) l'hypothèse " f est Injective" est indispensable de même que dans le théoréme 29,2°.

Par exemple, l'application $z \longrightarrow z^2$ de $\mathbb{C}$ dans $\mathbb{C}$, restreinte à l'ouvert Ω complémentaire de l'origine, vérifie bien la condition " $f'(z) = 2z \neq 0$ est inversible pour tout z de Ω "; elle est donc un homéomorphisme local et une application ouverte, mais n'est pas un homéomorphisme parce que non injective.

Reprenons le théorème 29. En supposant $f'(a)$ inversible, nous avons démontré une propriété d homéomorphisme; on en déduisait comme conséquence une propriété d'application ouverte, moins forte d'après ce que nous venons de voir. En fait le caractère d'application ouverte est en effet lié à une propriété moins forte que l'inversibilité de $f'(a)$, son caractère surjectif :

Théorème 30 - Soit f une application continuement dérivable d'un ouvert Ω d'un espace affine normé E dans un espace affine F de dimension finie. 1°/ Si a est un point de Ω et si $f'(a)$ est une surjection de $\vec{E}$ sur $\vec{F}$, alors l'image pour f de tout voisinage de a est un voisinage de $f(a)$, l'image par f de toute boule ouverte de centre a contient une boule ouverte de centre $f(a)$.

2°/ Si, pour tout x de Ω, $f'(x)$ est une surjection de $\vec{E}$ sur $\vec{F}$, l'image par f de tout ouvert de Ω est un ouvert de F ; en particulier $f(\Omega)$ est ouvert (Autrement dit, f est une application ouverte).

L'hypothèse" F est de dimension finie" est évidemment très restrictive: on démontre qu'on peut la lever, en supposant E et F complets.

Démonstration. Démontrons 1°)

Soit $(\vec{f}_i)$, $i = 1, 2, \ldots m$, une base de $\vec{F}$. Puisque $f'(a)$ est surjective, il existe des vecteurs $\vec{X}_1, \vec{X}_2, \ldots \vec{X}_m$ ayant pour images les $\vec{f}_i$; les $\vec{X}_i$ sont indépendants puisque leurs images par $f'(a)$ le sont, et par-suite la restriction de $f'(a)$ au sous-espace vectoriel $\vec{E}_0$ de dimension m engendré parles $\vec{X}_i$ est une surjection de $\vec{E}_0$ sur $\vec{F}$. Mais, $\vec{E}_0$ et $\vec{F}$ ayant la même dimension, $f'(a)$ est une bijection de $\vec{E}_0$ sur $\vec{F}$, autrement dit, considérée comme application linéaire de $\vec{E}_0$ dans $\vec{F}$, elle est inversible; et $\vec{E}_0$, de dimension finie, est complet (théorème 41 du chapitre II). Alors la restriction de f au sous espace affine E_0, mené par a parallèlement à $\vec{E}_0$, a une dérivée en a qui est inversible, et E_0 est complet; on peut lui appliquer le théorème 29, et cela démontre 1°). Nous voyons que, si $\mathcal{V}$ est un voisinage de a, non seulement l'image $f(\mathcal{V})$ est un voisinage de $f(a)$, mais même $f(\mathcal{V} \cap E_0)$ est un voisinage de $f(a)$. Le passage à 2° se fait comme dans la démonstration du théorème 29, c'est-à-dire en utilisant le critère A pour les applications ouvertes. Naturellement il n'est pas question ici que f soit un homéomorphisme local, la dérivée $f'(x)$ n'étant pas supposée injective.

Calcul des dérivées d'ordre supérieur d'une fonction implicite

Théorème 31 - Si dans les conditions de l'énoncé du théorème 28, la fonction f est m fois continuement dérivable dans 4 , alors la fonction implicite g est m fois continuement dérivable dans A .

Démonstration - Le théorème étant démontré pour $m = 1$ (théoreme 28) faisons une récurrence : supposons le théorème démontré pour la dérivabilité d'ordre $m-1$, démontrons-le pour la dérivabilité d'ordre $m \geqslant 2$. Remarquons d'abord que l'équation $uv = I$ traitée dans la 2ème démonstration du théorème 27 rentre dans le cas étudié ici. car $(u,v) \longrightarrow uv$ est bilinéaire continue, donc indéfiniment dérivable; donc $u \longrightarrow u^{-1}$ est, en vertu de l'hypothèse de récurrence, de classe C^{m-1} sur $\mathcal{U}$. Alors considérons la formule (III,8 37). La fonction g est connue comme C^{m-1} en vertu de l'hypothèse de récurrence; $\frac{\partial f}{\partial x}$ est C^{m-1}, puisque f est C^m; alors $x \longrightarrow \frac{\partial f}{\partial x}(x, g(x))$ est C^{m-1} d'après le théorème 19. Il en est de même de $x \longrightarrow \frac{\partial f}{\partial y}(x, g(x))$.

Comme, d'après ce que nous venons de voir, $u \rightarrow u^{-1}$ est encore C^{m-1}, le théorème 19 montre encore que $x \longrightarrow \left(\frac{\partial f}{\partial y}(x, g(x))\right)^{-1}$ est C^{m-1} . Enfin $(u,v) \longrightarrow v \circ u$ eest bilinéaire continue de $\mathcal{L}(\vec{E};\vec{G}) \times \mathcal{L}(\vec{G};\vec{F})$ dans $\mathcal{L}(\vec{E};\vec{G})$ (théorème 54), donc le théorème 18 montre que g' est C^{m-1} : donc g est bien C^m .

Corollaire 1 - Si $\vec{F}$ et $\vec{G}$ sont des espaces de Banach, $\mathcal{U}$ (resp. $\mathcal{V}^{-1}$) l'ensemble ouvert des éléments inversibles de $\mathcal{L}(\vec{F};\vec{G})$ (resp. $\mathcal{L}(\vec{G};\vec{F})$), $u \longrightarrow u^{-1}$ est un homéomorphisme de $\mathcal{U}$ sur $\mathcal{U}^{-1}$, indéfiniment dérivable ainsi que son homéomorphisme réciproque.

Ceci complète le théorème 27. En fait nous avons, par récurrence, ddmontré ce corollaire au cours même de la démonstration par récurrence du théorème 30.

Définition. On appelle C^m- difféomorphisme une bijection, de classe C^m ainsi que sa bijection réciproque.

Alors $u \longrightarrow u^{-1}$ est un C^∞- difféomorphisme de $\mathcal{U}$ sur $\mathcal{U}^{-1}$.

Corollaire 1 bis - Si $\vec{u}$ est une application m fois dérivable (resp. de classe C^m) d'un ouvert Ω d'un espace affine normé E dans $\mathcal{L}(\vec{F};\vec{G})$, et si, pour tout x de Ω, $\vec{u}(x)$ est un élément inversible de $\mathcal{L}(\vec{F};\vec{G})$, alors $x \to u^{-1}(x)$ est m fois dérivable (resp. de classe C^m), de Ω dans $\mathcal{L}(\vec{G};\vec{F})$.

Cela résulte du corollaire 1, et du théorème 19, l'application étant composée de $x \to u(x)$ et de $u \to u^{-1}$.

Comme cas très particulier, l'inverse $\frac{1}{f}$ d'une fonction scalaire f ($\vec{F} = \vec{G} = \mathbb{K} = \mathcal{L}(\vec{F};\vec{G})$) partout $\neq 0$, m fois dérivable ou de classe C^m sur $\Omega \subset E$, est aussi m fois dérivable ou de classe C^m.

Si $t \longrightarrow \mathcal{M}(t)$ est une fonction 2 fois dérivable sur $\mathbb{K}$ à valeurs dans un espace de matrices carrées, et si $\mathcal{M}(t)$ est inversible pour tout t, la dérivée seconde $\frac{d^2}{dt^2}(\mathcal{M}^{-1})$ s'obtient immédiatement, en dérivant (III,8;34):

$$\frac{d^2}{dt^2}(\mathcal{M}^{-1}) = 2\mathcal{M}^{-1}\frac{d\mathcal{M}}{dt}\mathcal{M}^{-1}\frac{d\mathcal{M}}{dt}\mathcal{M}^{-1} - \mathcal{M}^{-1}\frac{d^2\mathcal{M}}{dt^2}\mathcal{M}^{-1}$$

Corollaire 2 - Si, dans les conditions du théorème 29, 1°, f est de classe C^m, alors la bijection réciproque f^{-1} de f est une application de classe C^m de B dans A: f est un C^m-difféomorphisme de A sur B.

Naturellement le calcul, dans le cas le plus général, de la dérivée d'ordre m de g, à partir des dérivées partielles d'ordre $\leq m$ de la fonction f, donne lieu à d'inextricables complications dans l'écriture algébrique.

Nous nous bornerons à donner un exemple particulier : soit f une fonction de trois variables réelles et considérons l'équation :

$$f(x, y, z) = 0. \tag{III,8;46}$$

On se propose, ayant calculé z en fonction de x et de y au voisinage d'un point où la dérivée partielle $\frac{\partial f}{\partial z}$ est $\neq 0$, d'exprimer, en ce point, les dérivées partielles p, q, r, s, t, de z, à partir des dérivées partielles de f. On a déjà vu (en (III,8;16)) les formules :

$$p = \frac{\partial z}{\partial x} = \frac{-\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial z}}, \quad q = \frac{\partial z}{\partial y} = \frac{-\frac{\partial f}{\partial y}}{\frac{\partial f}{\partial z}} \tag{III,8;47}$$

Il reste. alors à dériver les expressions précédentes par rapport à x et à y , en tenant compte naturellement du théorème des fonctions composées, et 1 on obtient les expressions suivantes :

(III,8;48)
$$\left\{\begin{aligned} r &= \frac{\partial^2 z}{\partial x^2} = \frac{-\dfrac{\partial^2 f}{\partial x^2}\left(\dfrac{\partial f}{\partial z}\right)^2 + 2\dfrac{\partial^2 f}{\partial x\,\partial z}\dfrac{\partial f}{\partial x}\dfrac{\partial f}{\partial z} - \dfrac{\partial^2 f}{\partial z^2}\left(\dfrac{\partial f}{\partial x}\right)^2}{\left(\dfrac{\partial f}{\partial z}\right)^3} \\ s &= \frac{\partial^2 z}{\partial x\,\partial y} = \frac{-\dfrac{\partial^2 f}{\partial x\,\partial y}\left(\dfrac{\partial f}{\partial z}\right)^2 + \dfrac{\partial^2 f}{\partial x\,\partial z}\dfrac{\partial f}{\partial y}\dfrac{\partial f}{\partial z} + \dfrac{\partial^2 f}{\partial y\,\partial z}\dfrac{\partial f}{\partial x}\dfrac{\partial f}{\partial z} - \dfrac{\partial^2 f}{\partial z^2}\dfrac{\partial f}{\partial x}\dfrac{\partial f}{\partial y}}{\left(\dfrac{\partial f}{\partial z}\right)^3} \\ t &= \frac{\partial^2 z}{\partial y^2} = \frac{-\dfrac{\partial^2 f}{\partial y^2}\left(\dfrac{\partial f}{\partial z}\right)^2 + 2\dfrac{\partial^2 f}{\partial y\,\partial z}\dfrac{\partial f}{\partial y}\dfrac{\partial f}{\partial z} - \dfrac{\partial^2 f}{\partial z^2}\left(\dfrac{\partial f}{\partial y}\right)^2}{\left(\dfrac{\partial f}{\partial z}\right)^3} \end{aligned}\right.$$

On peut encore opérer comme suit :

On écrit sous forme entière la relation qui donne p , c'est-à-dire :

(III,8;49)
$$\frac{\partial f}{\partial x} + p\frac{\partial f}{\partial z} = 0,$$

on peut alors différencier cette relation, par rapport à x et y , en remplaçant naturellement dz, dp, dq par $p\,dx + q\,dy$, $r\,dx + s\,dy$, $s\,dx + t\,dy$. On obtient alors la relation :

(III,8;50)
$$\frac{\partial^2 f}{\partial x^2}dx + \frac{\partial^2 f}{\partial x\,\partial y}dy + \frac{\partial^2 f}{\partial x\,\partial z}(p\,dx + q\,dy) + \frac{\partial f}{\partial z}(r\,dx + s\,dy) + p\frac{\partial^2 f}{\partial x\,\partial z}dx + p\frac{\partial^2 f}{\partial y\,\partial z}dy + p\frac{\partial^2 f}{\partial z^2}(p\,dx + q\,dy) = 0.$$

En égalant à zéro les coefficients de dx et de dy, on obtient aussitôt le calcul des dérivées partielles *h* et s . Un calcul analogue fait à partir de la relation : $\frac{\partial f}{\partial y} + q\frac{\partial f}{\partial z} = 0$, donnerait les dérivées partielles s et t

Il existe encore une autre méthode : c'est l'utilisation du développement de Taylor (voir théorème 21 bis et 21 ter).

Ecrivons le développement de Taylor d'ordre 2 de f, au voisinage du point x, y, z. On a la formule :

$$(\text{III},8;51)\quad \begin{cases} 0 = f(x+X, y+Y, z+Z) - f(x,y,z) = \dfrac{\partial f}{\partial x}X + \dfrac{\partial f}{\partial y}Y + \dfrac{\partial f}{\partial z}Z \\ + \dfrac{1}{2}\left[\dfrac{\partial^2 f}{\partial x^2}X^2 + \dfrac{\partial^2 f}{\partial y^2}Y^2 + \dfrac{\partial^2 f}{\partial z^2}Z^2 + 2\dfrac{\partial^2 f}{\partial x \partial z}YZ + 2\dfrac{\partial^2 f}{\partial z \partial x}ZX + 2\dfrac{\partial^2 f}{\partial x \partial y}XY\right] + \cdots \end{cases}$$

Si l'on remplace, dans cette formule, z par son développement de Taylor au voisinage du point x, y, c est-à-dire :

$$(\text{III},8;52)\qquad Z = pX + qY + \frac{1}{2}\left(rX^2 + 2sXY + tY^2\right) + \cdots$$

on doit obtenir identiquement0 . En égalant à 0 les coefficients de X, Y, X^2, XY, Y^2 on a 5 équations qui donnent p, q, r, s, t On trouve'd'abord p et q en égalant à 0 les coefficients de X et Y En égalant à 0 les coefficients de X^2, XY, Y^2, on'trouve :

$$(\text{III},8;53)\quad \begin{cases} \dfrac{\partial f}{\partial z}r + \dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial z^2}p^2 + 2\dfrac{\partial^2 f}{\partial x \partial z}p = 0 \\ \dfrac{\partial f}{\partial z}s + \dfrac{\partial^2 f}{\partial z^2}pq + \dfrac{\partial^2 f}{\partial y \partial z}p + \dfrac{\partial^2 f}{\partial x \partial z}q + \dfrac{\partial^2 f}{\partial x \partial y} = 0\,. \\ \dfrac{\partial f}{\partial z}t + \dfrac{\partial^2 f}{\partial y^2} + \dfrac{\partial^2 f}{\partial z^2}q^2 + 2\dfrac{\partial^2 f}{\partial y \partial z}q = 0\,,\ \text{d'où (III,8;48).} \end{cases}$$

On peut encore, dans cette méthode du développement de Taylor procéder avec une variante. Ayant calculé les dérivées premières$p$ et q, on calculera les dérivées secondes en égalant à o l'expression (III,8;51), et en y calculant Z en fonction des autres quantités, c'est-à-dire X, Y, X^2, XY, Y^2, XZ, YZ, Z^2 : On obtiendra alors la formule :

$$(\text{III},8;54)\quad \begin{cases} \dfrac{\partial f}{\partial z}Z = -\dfrac{\partial f}{\partial x}X - \dfrac{\partial f}{\partial y}Y \\ -\dfrac{1}{2}\left[\dfrac{\partial^2 f}{\partial x^2}X^2 + \dfrac{\partial^2 f}{\partial y^2}Y^2 + \dfrac{\partial^2 f}{\partial z^2}Z^2 + 2\dfrac{\partial^2 f}{\partial y \partial z}YZ + 2\dfrac{\partial^2 f}{\partial z \partial x}ZX + 2\dfrac{\partial^2 f}{\partial x \partial y}XY\right] + \cdots \end{cases}$$

Dans cette formule, Z est exprimé à l'aide de XZ, YZ et Z^2; mais il est possible, pour avoir une approximation du second ordre, de se contenter, dans ces expressions, de remplacer Z par $pX + qY$, p et q ayant été calculés auparavant.

On obtient ainsi l'expression :

$$(\text{III},8;55)\quad \left\{\begin{aligned} \frac{\partial f}{\partial z} Z = &-\frac{\partial f}{\partial x} X - \frac{\partial f}{\partial y} Y - \frac{1}{2}\Big[\frac{\partial^2 f}{\partial x^2} X^2 + 2\frac{\partial^2 f}{\partial x \partial y} XY + \frac{\partial^2 f}{\partial y^2} Y^2 \\ &+ 2\frac{\partial^2 f}{\partial x \partial z} X(pX+qY) + 2\frac{\partial^2 f}{\partial x \partial z} Y(pX+qY) \\ &+ \frac{\partial^2 f}{\partial z^2}(pX+qY)^2\Big] + \cdots, \end{aligned}\right.$$

ce qui donne encore $\frac{1}{2}r$, $\frac{1}{2}s$, $\frac{1}{2}t$, en calculant les coefficients de X^2, $2XY$, Y^2, dans le développement de Z.

Technique du changement de variable et du changement de fonction

Nous avons déjà vu ce qu'était le problème du changement de variable antérieurement au théorème des fonctions implicites, mais ce théorème nous permet maintenant des changements beaucoup plus importants.

Supposons qu'on ait une équation aux dérivées partielles du second ordre par exemple, satisfaite par une fonction scalaire z de 2 variables scalaires x et y, soit

$$(\text{III},8;56)\qquad F(x,y,z,p,q,r,s,t) = 0 .$$

On fait alors à la fois le changement de variables et le changement de fonction définis par les formules :

$$(\text{III},8;57)\qquad X = X(x,y,z)\,,\qquad Y = Y(x,y,z)\,,\qquad Z = Z(x,y,z)\ ^*$$

On voudrait savoir comment se transforme l'équation aux dérivées partielles, lorsqu'on fait le changement de variables et de fonction considéré. Un tel changement de variables n est légitime que si, d'une part (III,8;57) définit un C^2-difféomorphisme d'un ouvert de $\mathbb{K}^3$ sur un ouvert de $\mathbb{K}^3$, et si la surface $z = z(x,y)$ est ainsi transformée en une surface qui

* X, Y, Z n'ont évidemment pas la même signification que dans ce qui précède.

soit encore de la forme $Z = Z(X,Y)$ les fonctions écrktes étant de classe C^2. Suivant la'méthode habituelle, il faut donc calculer p, q, r, s, t, en fonction des dérivées partielles nouvelles P, Q, R, S, T. On pourra commencer par les dérivées du premier ordre, et continuer par les dérivées du second ordre. Pour les dérivées du premier ordre, par exemple, on partira de la relation $dZ = P\,dX$, $Q\,dY$, on remplacera alors dX, dY, dZ, par leurs expressions différentielles tirées de (III,8;57), et l'on obtiendra ainsi une relation entre dx, dy, dz; cette relation, mise sous la forme $dz = p\,dx, q\,dy$, permettra de calculer p, q, en fonction de x, y, z, P, Q. En différentiant la relation obtenue, on calculera dp en fonction de dx, dy, dz, dP, dQ ; remplaçant dz, dP, dQ, par $p\,dx + q\,dy$, $R\,dX, S\,dY$, $S\,dX$, $T\,dY$, puis p et q par leurs valeurs antérieurement calculées, dX et dY par leurs valeurs tirées de (III,8:57) (où dz est encore remplacée par $p\,dx+q\,dy$, p et q déjà calculés), on obtiendra dp comme une combinaison de dx, dy, qui, mise sous la forme $dp = r\,dx + s\,dy$, donnera r et A en fonction de x, y, z, P, Q, R, S, T.

On calculera de même t. On portera tous les résultats obtenus dans (III,8;56).

Naturellement 1 dquation qu'on obtiendra en P, Q, R, S, T, contiendra en fait les quantités x, y, z, et non X, Y, Z: il faudra bien, si elles interviennent effectivement, calculer (x, y, z) en fonction de (X, Y, Z) en résolvant les équations (III,8;57) *

* Au lieu de (III,8;57), il se peut qu'on donne les anciennes coordonnées en fonctions des nouvelles, ou 3 relations non résolues entre les anciennes et les nouvelles; il y a une grande richesse de cas possibles. Plutôt que d'utiliser une méthode standard, il faut savoir réfléchir ! En Mathématiques Spéciales, on a utilisé de tels changements de variables quand, par exemple, partant de la courbure exprimée en fonction de x', y'', pour une courbe $y = y(x)$ en cartésiennes, on l'a calculée en fonction de r, r', r'' pour une courbe en polaires $r = r(\varphi)$.

§ 9 VARIÉTÉS DIFFÉRENTIABLES

Nous avons vu, à propos du théorème 8 A, que l'ensemble& du produit d'espaces affines $E \times F$ défini par une équation $y = f(x)$, s'appelle une variété différentiable, si f est fonction dérivable. Il s'appelle une variété m fois différentiable si la fonction f est m fois dérivable, variété continuement différentiable, ou m fois continuement différentiable, si f est continuement dérivable ou m fois continuement dérivable, variété indéfiniment différentiable, si f est indéfiniment dérivable; on dit aussi variété de classe C^m au lieu de dire variété m fois continuement différentiable, et variété de classe C^∞ au lieu de variété indéfiniment différentiable. Ce sont surtout ces variétés de classe C^m et C^∞ qui sont utilisées. Si $m = 0$, on a affaire à ce qu'on appelle une variété topologique, définie seulement par une équation $y = f(x)$, où f est une fonction continue. Les variétés topologiques ont aussi un rôle très important, mais nous nous bornerons uniquement à étudier les variétés de classe C^m pour $m \geq 1$.

Il est bien évident que les exemples de variétés ainsi donnés ne sont pas les plus généraux; ainsi il est normal de considérer que, dans un espace euclidien affine de dimension N une sphère est une variété indéfiniment différentiable de dimension $N-1$. cependant cet espace affine n'est pas donné naturellement comme un produit, et, si l'on prend un système de coordonnées dans cet espace affine, on ne peut pas représenter la sphère toute entière en exprimant 1 une des coordonnées comme fonction indéfiniment dérivable des autres;* on sera obligé de partager la sphère en régions suffisamment petites, et, dans chacune de ces régions, l'une des coordonnées, convenablement choisie, pourra s'exprimer comme fonction indéfiniment dérivable des autres. Ceci va nous mener au concept général des variétés de classe C^m

Définition d'une variété par expression de certaines coordonnées comme fonctions des autres. Soit E un espace affine sur le corps $\mathbb{K}$ des reels ou des complexes, de dimension N; soit V un ensemble de E. On dit que V est une variété de dimension n, de classe C^m, si, quel que soit le point a de V, il existe un référentiel de E, soit : $O, e_1, e_2, \dots, e_N$, un ouvert $\mathcal{B}$ du sous-espace des n premiers axes de coordonnées, un système de $N-n$ fonctions G_k, $k = 1, 2, \dots N-n$, de classe C^m, définies sur $\mathcal{B}$, à valeurs scalaires et un ouvert $\mathcal{V}$ de E contenant a, dont la projection sur le sous-espace des n premiers axes de coordonnées soit $\mathcal{B}$, tels que l'intersection $V \cap \mathcal{V}$ soit exactement l'ensemble des points $x = (x_1, x_2, \dots, x_N)$ de E vérifiant les équations

* Une équation de la sphère de centre origine et de rayon R dans R^n est $x_n = \pm \sqrt{R^2 - x_1^2 - \cdots - x_{n-1}^2}$. A cause de $\pm$, ce n'est pas une fonction (et même si on choisit $+$ ce n'est pas une fonction dérivable aux points où $x_1 = 0$).

(III,9;1) $$x_{n+k} = G_k(x_1, x_2, \dots, x_n)^{*} \quad ; \quad k = 1, 2, \dots N-n .$$

Ainsi, dans $\mathcal{V}$, les $N-n$ dernières coordonnées x_{n+1}, $x_{n+2}, \dots, x_N$, sur la variété V s'expriment comme fonctions de classe C^m des n premières coordonnées x_1, $x_2, \dots, x_n$, qu'on peut choisir arbitrairement dans $\mathcal{B}$. (On pourrait s'étonner de ce rôle particulier joué par les n premières coordonnées; il n'est pas étonnant, puisque, pour le point a considéré de V, nous nous sommes permis de choisir un référentiel de E; si donc n des coordonnées jouent un rôle particulier, on peut toujours, par changement d'ordre des vecteurs du référentiel, se ramener au cas commode où ce sont les n premières).

Nous verrons plus loin (corollaire 1 du théorème 32) une définition équivalente, dans laquelle au contraire les n premières coordonnées ne jouent pas de rôle particulier, parce qu'on a choisi une fois pour toutes un même référentiel de E.

On peut encore dire que la séparation, dans le système des N coordonnées, d'un côté des n premières, de l'autre des $N-n$ suivantes, identifie E au produit $\mathbb{K}^n \times \mathbb{K}^{N-n}$, et que, si l'on appelle g l'application de l'ouvert $\mathcal{B}$ de $\mathbb{K}^n$ dans $\mathbb{K}^{N-n}$, définie par (III,9;1), alors g est de classe C^m, et l'ensemble $V \cap \mathcal{V}$ est exactement l'ensemble des points $x = (y, z) \in E = \mathbb{K}^n \times \mathbb{K}^{N-n}$, vérifiant l'équation $z = g(y)$.

Nous retrouvons ainsi la situation particulière étudiée au théorème 8 A, mais, au voisinage du point a, nous nous sommes permis de choisir un référentiel particulier, amenant à cette situation particulière.

Définition d'une variété par une représentation paramétrique

Soit V un ensemble d'un espace affine E de dimension N. On appelle représentation paramétrique vraie de V, de classe C^m, de dimension n, une application $\Phi : u \to \Phi(u)$, d'un ouvert $\mathcal{O}$ de $\mathbb{K}^n$ dans E, ayant les propriétés suivantes :

1°/ Φ est un homéomorphisme de $\mathcal{O}$ sur V;

2°/ Φ est une application de classe C^m de $\mathcal{O}$ dans E;

3°/ En tout point $a = (\alpha_1, \alpha_2, \dots, \alpha_n)$ de $\mathcal{O}$, l'application dérivée $\Phi'(\alpha) \in \mathcal{L}(\mathbb{K}^n; \vec{E})$ est exactement de rang n; autrement dit, l'image, par cette application, de

* D'après la propriété supposée de la projection de $\mathcal{V}$, $(x_1, x_2, \dots x_n)$ est dans $\mathcal{B}$, et G_k est définie sur $\mathcal{B}$.

l'espace vectoriel $\mathbb{K}^n$, est un sous-espace vectoriel de dimension n de E ; ou encore, les vecteurs dérivées partielles $\frac{\partial \Phi}{\partial u_j}(\alpha)$, j 1, 2, ... n, sont n vecteurs indépendants de E ; ou encore, $\exists$ l'un au moins des déterminants de rang n de la matrice dérivée (matrice des $\frac{\partial \Phi_i}{\partial u_j}(\alpha)$) est $\neq 0$.

On se doute évidemment que les variétés, de dimension n et de classe C^m de E, sont les seuls ensembles à pouvoir admettre de telles représentations paramétriques. Par contre, il est bien évident qu'une variété n'admettra pas, en général, de telles représentations paramétriques; cela impliquerait en effet qu'elle soit homéomorphe à un ouvert de $\mathbb{K}^n$, or l'exemple de la sphère, qui est compacte, montre qu'une variété ne peut pas être en général homéomorphe à un ouvert de $\mathbb{K}^n$; ce sont seulement, encore une fois, des réginns suffisamment petites de la variété, qui pourront admettre de telles représentations paramétrlques.

Bien entendu, la donnée de Φ contient celle de l'ouvert $\mathcal{O}$ de $\mathbb{K}^n$ et de son image $\Phi(\mathcal{O})$ dans V; mais on l'indique souvent par la notation $\Phi : \mathcal{O} \longrightarrow \Phi(\mathcal{O})$.

<u>Théorème 32</u> - <u>Pour qu'un ensemble V d'un espace affine E de dimension N soit une variété de classe C^m de dimension n, il faut et il suffit que, pour tout point a de V, il existe un voisinage ouvert $\mathcal{V}$ de a dans E, tel que l'intersection $V \cap \mathcal{V}$ admette une représentation paramétrique vraie, de dimension n et de classe C^m</u>.

<u>Démonstration</u> 1°/ Supposons que V soit une variété de dimension n et de classe C^m. Soit a un point de V. Si alors on détermine les ouverts $\mathcal{B}$ et $\mathcal{V}$, comme il a été fait dans la définition, il suffit de prendre $\mathcal{O} = \mathcal{B}$, et de définir l'application $\Phi : u \to x = \Phi(u)$, à partir des fonctions G_k de (III,9;1), par :

(III,9;2) $$x_1 = u_1, \dots, x_n = u_n, \; x_{n+1} = G_1(u_1, u_2, \dots, u_n), \dots x_N = G_{N-n}(u_1, u_2, \dots u_n),$$

pour avoir une représentation paramétrique du type cherché; c'est ici le déterminant jacobien $\frac{D(x_1, \dots, x_n)}{D(u_1, u_2, \dots, u_n)}$, qui est sûrement $\neq 0$ *.

2°/ Inversement, supposons que V soit un ensemble de E vérifiant les propriétés de l'énoncé. Montrons que V est une variété de dimension n et de classe C^m.

* Ce que nous venons de faire est l'astuce bien connue du tôpin qui, pour étudier une courbe $y = f(x)$, préfère poser $x = t$, $y = f(t)$.

Soient a un point de V, et Φ une représentation paramétrique vraie, homéomorphisme d'un ouvert@ de $\mathbb{K}^n$ sur un voisinage de a dans V. Choisissons arbitrairement un référentiel de E. Alors l'application Φ possède, au point $\alpha = \Phi^{-1}(a)$, une matrice dérivée, par rapport aux référentiels de $\mathbb{K}^n$ et de E ; cette matrice dérivée, par hypothèse, est de rang n, et l'un au moins de ses déterminants à n lignes et n colonnes est $\neq 0$. En changeant au besoin l'ordre des vecteurs de la base de E, nous pouvons supposer que c'est le déterminant jacobien $\frac{D(x_1, x_2, \cdots, x_n)}{D(u_1, u_2, \cdots, u_n)}$. Le référentiel ainsi formé dans E l'identifie à $\mathbb{K}^N$ et la distinction du système des n premières coordonnées et des $N-n$ dernières l'identifie au produit $\mathbb{K}^n \times \mathbb{K}^{N-n}$. C'est pourquoi nous représenterons tout point x de E comme un couple (y, z), $y \in \mathbb{K}^n$, $z \in \mathbb{K}^{N-n}$; soit $a = (b, c)$.

L'application Φ devient alors une application de l'ouvert@ dans le produit $\mathbb{K}^n \times \mathbb{K}^{N-n}$ et se décompose par suite en un système de deux applications X, Ψ, de $\mathcal{O}$ dans $\mathbb{K}^n$ et $\mathbb{K}^{N-n}$ respectivement, à savoir $y = X(u)$, $z = \Psi(u)$. Mais alors le déterminant jacobien de l'application X de $\mathcal{O} \subset \mathbb{K}^n$ dans $\mathbb{K}^n$ au point α est $\neq 0$. D'après le théorème 29, il existe donc un ouvert& contenant α dans $\mathcal{O}$, et un ouvert $\mathcal{B}$ contenant& dans $\mathbb{K}^n$, tel- que X soit un homéomorphisme de $\mathcal{A}$ sur $\mathcal{B}$,

continuement dérivable, ainsi que son homéomorphisme réciproque ; en outre, X est de classe C^m, et le corollaire 2 du théorème 31 indique que son homéomorphisme réciproque est aussi de classe C^m. Désignons par X^{-1} cet homéomorphisme réciproque, de $\mathcal{B}$ sur $\mathcal{A}$. L'image de $\mathcal{A}$ par@ est alors un ouvert de V, puisque Φ est supposée être un homéomorphisme de 6 sur son image, et que celle-ci est ouverte dans V. IL existe donc un voisinage ouvert $\mathcal{V}$ de a dans E, tel que $\Phi(\mathcal{A}) = V \cap \mathcal{V}$. Comme $\Phi(\mathcal{A})$ se projette sur $\mathbb{K}^n$ suivant $\mathcal{B}$, on peut supposer que $\mathcal{V}$ aussi se projette suivant $\mathcal{B}$ (sans quoi on le remplacerait par son intersection avec l'ensemble ouvert $\{x ; (x_1, x_2, \ldots, x_n) \in \mathcal{B}\}$ de $\mathbb{K}^N$). Alors, pour qu'un point $x = (y, z)$ appartienne à $V \cap \mathcal{V}$, il faut et il suffit qu'il soit l'image par Φ d'un point de $\mathcal{A}$, c'est-a-dire que y appartienne à $\mathcal{B}$, et que z soit reliée à y par la formule $z = \Psi(X^{-1}(y))$. Cette application composée $g = \Psi \circ X^{-1}$ de deux applications de classe C^m est elle-même de classe C^m, d'après le théorème 19; nous voyons donc bien que nous avons pu trouver un voisinage $\mathcal{V}$ de a dans E, et un référentiel de E, tels que l'intersection $V \cap \mathcal{V}$ puisse se représenter par une équation $z = g(y)$, où g est une application de classe C^m de $\mathcal{B}$ dans E. Ceci étant vrai pour tout point a de V, V est bien une variété de dimension n et de classe C^m.

Nous avons en même temps démontré que, si l'on a choisi au préalable un référentiel quelconque de E, alors, pour tout point a de V, pour obtenir un référentiel vérifiant les propriétés de la définition d'une variété, il suffit de changer convenablement l'ordre des vecteurs de base; ce changement dépendant évidemment de a.

On peut donc énoncer :

Corollaire 1 - Soit V un ensemble d'un espace affine E de dimension N. Pour que V soit une variété de dimension n et de classe C^m, il faut et il suffit que, un référentiel de E ayant été choisi arbitrairement une fois pour toutes, alors, pour tout point a de V, il existe une permutation σ de l'ensemble $1, 2, \ldots, N$, un ouvert $\mathcal{B}$ du sous-espace engendré par l'origine et les vecteurs de base $\vec{e}_{\sigma_1}, \vec{e}_{\sigma_2}, \ldots, \vec{e}_{\sigma_n}$, et un ouvert $\mathcal{V}$ de E, contenant a, dont la projection sur le sous-espace précédent soit $\mathcal{B}$, enfin un système de $N-n$ fonctions scalaires G_k, $k = 1,2,\ldots,N-n$, de classe C^m, définies sur $\mathcal{B}$, de telle manière que l'intersection $V \cap \mathcal{V}$ soit définie par les équations

$$x_{\sigma_{n+k}} = G_k(x_{\sigma_1}, x_{\sigma_2}, \ldots, x_{\sigma_n}) \,; \qquad k = 1, 2, \ldots, N-n \,. \tag{III,9;3}$$

Corollaire 2 - Soient E, F, deux espaces affines de même dimension finie, H un difféomorphisme de classe C^m d'un ouvert $\mathcal{U}$ de E sur un ouvert $H(\mathcal{U})$ de F. Alors, si V est une variété de dimension n et de classe C^m de $\mathcal{U}$, son image H(V) est une variété de $H(\mathcal{U})$, de dimension n et de classe C^m.

Ce corollaire exprime que les variétés de classe C^m se conservent par les C^m- difféomorphismes.

Démonstration - Soit $\Phi : \mathcal{O} \to \Phi(\mathcal{O})$, une représentation paramétrique vraie d'un ouvert $\Phi(\mathcal{O})$ de V. Alors $H \circ \Phi$ est un homéomorphisme de $\mathcal{O}$ sur $H(\Phi(\mathcal{O}))$; c'est une application de classe C^m de $\mathcal{O}$ dans F ; si $a \in \mathcal{O}$, $(H \circ \Phi)'(\alpha)$

$= H'(\Phi(\alpha)) \circ \Phi'(\alpha)$ est de rang n, parce que $\Phi'(\alpha)$

est de rang n et que $H'(\Phi(\alpha))$ est une bijection linéaire (corollaire 4 du théorème 11 appliqué à H). Donc $H \circ \Phi$ est une représentation paramétrique vraie, de dimension n et de classe C^m. Il existe des représentations telles que Φ, dont les images forment un recouvrement de V, donc les images des $H \circ \Phi$ forment un recouvrement de H(V), et le corollaire est démontré.

Corollaire 2 bis - Pour qu'un ensemble V d'un espace affine E de dimension N soit une variété de dimension n et de classe C^m, il faut et il suffit que, pour tout point a de V, il existe un C^m- difféomorphisme $\tilde{\Phi}$ d'un ouvert $\tilde{\mathcal{O}}$ de $\mathbb{K}^N$ sur un voisinage $\mathcal{V}$ de a dans E, tel que $V \cap \mathcal{V}$ soit l'image par $\tilde{\Phi}$ de l'intersection de $\tilde{\mathcal{O}}$ et du sous-espace vectoriel $\mathbb{K}^n$ engendré par les n premiers vecteurs de base de $\mathbb{K}^N$. Ce corollaire exprime que, localement, on peut amener par un C^m-difféomorphisme, une variété de classe C^m et de dimension n de E à devenir un sous-espace vectoriel de dimension n de $\mathbb{K}^N$.

Démonstration 1°/ Supposons d'abord que V possède la propriété énoncée. En tout point α de $\tilde{\Phi}^{-1}(V)$, la dérivée $\tilde{\Phi}'(\alpha)$ est une bijection linéaire de $\mathbb{K}^N$ sur $\vec{E}$ (corollaire 4 du théorème 11); donc elle est injective, ainsi que sa restriction au sous-espace $\mathbb{K}^n$ considéré, qui est par conséquent de rang n. Alors la restriction Φ de $\tilde{\Phi}$ à $\mathcal{O} = \tilde{\mathcal{O}} \cap \mathbb{K}^n$ est une représentation paramétrique vraie de dimension n et de classe C^m, et comme chaque point de V a un voisinage ayant une telle représentation paramétrique, V est bien une variété de dimension n et de classe C^m.

2°/ Réciproquement, soit V une variété de dimension= et de classe C^m. D'après la définition, si $a \in V$, il existe un référentiel de E et un voisinage $\mathcal{V}$ de a dans E, dans lequel V peut se définir par (III,9;1), avec $(x_1, x_2, \ldots, x_n) \in \mathcal{B}$. Appelons $\mathcal{B}^*$ l'ouvert de $\mathbb{K}^N$, formé des points dont les n premières coordonnées définissent un point de $\mathcal{B}$: $\mathcal{B}^* = \{x \,;\, (x_1, x_2, \cdots, x_n) \in \mathcal{B}\}$.

Alors les formules :

$$(\mathrm{III},9;3bis)\quad \begin{cases} x_j = u_j\,, & j = 1,2,\ldots,n\,; \\ x_{n+k} = u_{n+k} + G_k(u_1, u_2, \ldots, u_n)\,, & k = 1,2,\ldots,N-n \end{cases}$$

définissent une application Φ de $\mathcal{B}^*$ dans $\mathbb{K}^N$ identifié à E; on a $\tilde{\Phi}(\mathcal{B}^*) = \mathcal{B}^*$, et $\tilde{\Phi}$ est une bijection, de bijection réciproque $\tilde{\Phi}^{-1}$ définie par :

$$(\mathrm{III},9;3ter)\quad \begin{cases} u_j = x_j\,, & j = 1,2,\ldots,n\,; \\ u_{n+k} = x_{n+k} - G_k(x_1, x_2, \ldots, x_n)\,, & k = 1,2,\ldots,N-n \end{cases}$$

$\tilde{\Phi}$ et $\tilde{\Phi}^{-1}$ sont tous deux de classe C^m puisque les G_k sont de classe C^m; donc $\tilde{\Phi}$ est bien un'homéomorphisme, de classe C^m ainsi que son homéomorphisme réciproque . Si on appelle $\tilde{\mathcal{O}}$ l'image réciproque $\tilde{\Phi}^{-1}(\mathcal{V})$, c'est un ouvert de $\mathbb{K}^N$, et Φ, restreinte à $\tilde{\mathcal{O}}$, a les mêmes propriétés; mais alors $V \cap \mathcal{V}$ est défini par (III,9:1), et par suite est l'image par $\tilde{\Phi}$ du sous-espace $u_{n+k} = 0$, $k = 1,2,\ldots N-n$, ce qui démontre le corollaire.

Si V est une variété de dimension n et de classe C"', toute représentation paramétrique propre Φ de classe C^m d'un ouvert $V \cap \mathcal{V}$ de cette variété s'appelle aussi <u>une carte locale (ou carte) de classe C^m · ou simplement une carte</u> si aucune confusion n est à craindre. L'ensemble $\Phi(\mathcal{O}) = V \cap \mathcal{V}$ s'appelle aussi <u>l'image de la carte,</u> et Φ est aussi appelée <u>une carte de $V \cap \mathcal{V}$</u>. Si $\Phi(\mathcal{O})$ contient le point a de V, on dit aussi que <u>la carte recouvre ce polnt.</u> On appelle <u>atlas de la variété V, tout ensemble de cartes d'ouverts de V, dont les images forment un recouvrement de V</u>. <u>L'atlas universel de V</u> est l'ensemble de toutes les cartes (sous-entendu : de dimension-n et de classe C^m) d'ouverts de V. * Ces mots de carte et d'atlas sont évidemment tirés des représentations de la terre. La surface de la terre est sensiblement une sphère et on peut la considérer comme une variété de dimension 2 . Il n'existe pas de représentation paramétrique globale de cette variété compacte à partir d'un ouvert de $\mathbb{R}^2$; mais il existe un recouvrement de cette terre V par un système d'ouverts, tels que chacun d'eux soit exactement l'image d'une carte. Naturellement, comme V est compacte, il existe <u>un atlas fini de V</u>, c'est-a-dire comprenant un nombre fini de cartes ** . Chaque "page" d'un atlas est un rectangle, que nous considérons comme étant un rectangle <u>ouvert</u> de $\mathbb{R}^2$; en face de différents points de ce rectangle on a marqué le nom d'un lieu de la terre, c'est-à-dire d'un point particulier de l'ensemble V . Une carte parfaite du point de vue mathématique devrait avoir placé, en présence de <u>tout</u> point du rectangle, le point correspondant de la terre, c'est-à-dire exactement avoir défini une application Φ du rectangle ouvert sur un ouvert de la variété V. Du fait que les cartes sont des ensembles ouverts, il résulte inévitablement que certains points de V sont recouverts par plusieurs cartes de l'atlas (nous exigeons en effet que chaque ville du monde soit au moins représentée par un point intérieur à l'un des rectangles-cartes); si en effet chaque point n'était recouvert qu'une fois, V serait réunion d'un

* <u>L'atlas universel est un ensemble de cartes ayant la puissance du contlnu.</u>

** Les atlas usuels ont bien un nombre fini de cartes. L'atlas universel, pour la raison indiquée dans la note * , serait encombrant et coûteux. Ha Ha! d'où gag!

nombre fini >1 d'ouverts disjoints, donc ne serait pas connexe, or la terre, comme toute sphère, est connexe, qu'on le regrette ou non.

Exemples de variétés

1°/ Toute variété linéaire de dimension n est une variété indéfiniment différentiable; elle comprend en effet un atlas à une seule carte, obtenue en choisissant un référentiel de cette variété, référentiel qui définit une bijection Φ linéaire de $\mathbb{K}^n$ sur la variété, ayant toutes les propriétés voulues. En particulier l'espace E lui-même est une variété.

2°/ Nous avons implicitement supposé $n \geq 1$. En fait, on est amené à considérer comme variété de dimension 0 de E tout ensemble V de points isolés. Une carte d'un voisinage de $a \in V$ est alors l'application $0 \rightarrow a$. de $\mathbb{K}^0 = \{0\}$ dans $\{a\}$.

On est aussi amené à considérer que la partie vide $\emptyset$ est une variété sans cartes, de dimension -1.

Une variété de dimension 1 s'appelle une courbe, une variété de dimension 2 une surface, une variété de dimension $N-1$ une hypersurface. Toutefois ces mots ont été et sont encore employés dans tellement de sens différents qu'il faut être prudent dans leur usage. Il est habituel de considérer la lemniscate de Bernoulli comme une "courbe", mais ce n'est pas une variété à cause de son point singulier; en disant "courbe de classe C^m", on précisera bien qu'on veut dire "variété de dimension 1, de classe C^m", et la lemniscate sera exclue.

3°/ Une variété V de dimension N d'un espace affine E de dimension N est simplement un ouvert de E (en effet, ici $N-n=0$; si l'on revient à la définition, il n'y a pas d'équations (III,9;1), et $V \cap \mathcal{V} = \mathcal{V} = \mathcal{B}$, ouvert de E contenant a ; donc la définition exprime simplement que tout point a de V a un voisinage $\mathcal{V}$ dans E qui appartient à V, donc que V est ouverte). Un tel ouvert est alors une variété de classe C^∞, et il admet un atlas a une seule carte, définie par l application identique de $V \subset E$, identifié à $\mathbb{K}^N$ par un référentiel, dans V.

Toute partie d'une variété V de E, qui est ouverte dans V, est une variété de même classe, et de même dimension.

Inversement, si W et V sont des variétés de E, de même dimension n, $W \subset V$, alors W est simplement un ouvert de V. Soit en effet $a \in W$. D'après le corollaire 2 bis du théorème 32, il existe un C^m- difféomorphisme $\tilde{\Phi}$ d'un ouvert $\mathcal{O}$ de $\mathbb{K}^N$ sur un ouvert de E contenant a, tel que $\tilde{\Phi}^{-1}(V)$ soit l'intersection de 6 avec un sous-espace vectoriel $\mathbb{K}^n$ de $\mathbb{K}^N$. Mais le corollaire 2 montre que $\Phi^{-1}(W)$ est aussi une variété de dimension* de $\mathbb{K}^N$, nécessairement contenue dans $\mathbb{K}^n$; ce que nous venons de voir au début de ce 3°/ montre que c'est un ouvert de $\mathbb{K}^n$, et alors $\tilde{\Phi}^{-1}(W)$ est un voisinage de $\tilde{\Phi}^{-1}(a)$ dans $\tilde{\Phi}^{-1}(V)$, donc W est un voisinage de a dans V, ceci étant vrai pour tout point a de W, W est bien un ouvert de V.

4°/ Montrons que, dans un espace affine euclidien de dimension N sur $\mathbb{R}$, une sphère est une variété indéfiniment différentiable de dimension $N-1$. Pour simplifier, prenons un référentiel et pour S la sphère unité de centre origine. Appelons $\mathcal{V}_{j,+}$ l'ouvert de E défini par l'inégalité $x_j > 0$. Appelons $\mathcal{B}_j$ l'ouvert de K^{N-1} défini par $x_1^2 + x_2^2 + \cdots + x_{j-1}^2 + x_{j+1}^2 + \cdots + x_N^2 < 1$. Alors, dans $\mathcal{V}_{j,+}$, la sphère S est définie par l'équation

$$(x_1, x_2, \ldots, x_{j-1}, x_{j+1}, \ldots x_N) \in \mathcal{B}_j, \quad x_j = \sqrt{1 - x_1^2 - \cdots - x_{j-1}^2 - x_{j+1}^2 - \cdots - x_N^2}\,. \tag{III,9;4}$$

On opérera de même dans l'ouvert $\mathcal{V}_{j,-}$ défini par $x_j < 0$, et, comme ceci peut se faire pour $j = 1, 2, \cdots, N$, on voit que notre affirmation est démontrée, et qu'on a recouvert S par un atlas de $2N$ cartes de classe C^∞ et de dimension $N-1$.

Considérons la représentation paramétrique classique de la sphère à deux dimensions dans $\mathbb{R}^3$ par les formules

$$\left\{\begin{aligned} x &= \sin\theta\,\cos\varphi\,, \\ y &= \sin\theta\,\sin\varphi\,, \\ z &= \cos\theta \end{aligned}\right. \tag{III,9;5}$$

Ici $\mathcal{O}$ sera, par exemple, l'ouvert $0 < \varphi < 2\pi$, $0 < \theta < \pi$, de $\mathbb{R}^2$, et l'application (III,9;5) donne une carte de l'ouvert de la sphère défini par les inégalités : $y \neq 0$ ou $y = 0$, $x < 0$ (complémentaire du demi méridien fermé $\varphi = 0$, $0 \leq \theta \leq \pi$). On se sert habituellement de la représentation précédente pour toute la sphère en admettant les valeurs $0 \leq \varphi \leq 2\pi$, $0 \leq \theta \leq \pi$, mais alors ce n'est pas une représentation parametrique vraie. Autrement dit, cette carte à elle seule ne peut pas, comme il a déjà été vu, constituer un atlas de toute la sphère.

Par contre, à l'aide de la projection stéréographique, il est facile de montrer qu'il existe un atlas de la sphère constitué par deux cartes seulement. Considérons en effet l'application réciproque de la projection stéréographique, c'est-à-dire l'application de $\mathbb{R}^{N-1}$ dans la sphère, définie par les formules $u = (u_1, u_2, \dots, u_{N-1}) \longrightarrow x = (x_1, x_2, \dots, x_N)$, avec :

$$(\text{III}.9;10) \quad \begin{cases} x_i = \dfrac{2u_i}{\sum_{j=1}^{N-1} u_j^2 + 1}, & i = 1, 2, \dots, N-1; \\[2ex] x_N = \varepsilon \dfrac{\sum_{j=1}^{N-1} u_j^2 - 1}{\sum_{j=1}^{N-1} u_j^2 + 1} \end{cases}$$

On définit bien là, pour $\varepsilon = +1$ (resp -1), une carte de classe C^∞, dont l'image est 1 ouvert de la sphère, complémentaire du pôle Nord (resp. Sud) (c'est-à-dire du point 0, 0, ..., ε (*))

Nous avons donc bien un atlas de la sphère formé de 2 cartes.

Il résulte de la démonstration du théorème 32 que, si nous considérons l'application $\Theta : (y, z) \longrightarrow X^{-1}(y)$, c'est une application de $\mathcal{V}$ dans $\mathcal{A}$, dont la restriction à la variété $V \cap \mathcal{V}$ est aussi la restriction de l'homéomorphisme réciproque Φ^{-1} de Φ. On peut donc énoncer :

Théorème 33 - Si $\Phi : \mathcal{O} \longrightarrow \Phi(\mathcal{O})$, est une représentation paramétrique vraie, de dimensionn et de classe C^m, d'un ouvert d'une variété V de E, et si $\Phi(\alpha) = a$, alors il existe un ouvert $\mathcal{A}$ de $\mathcal{O}$ contenant α, un ouvert $\mathcal{V}$ de E contenant a, $V \cap \mathcal{V} = \Phi(\mathcal{A})$, et enfin une application Θ de classe C^m de $\mathcal{V}$ dans $\mathcal{A}$, dont la restriction à $V \cap \mathcal{V}$ coincide avec la restriction de l'homéomorphisme récioroaue de Φ ; en particulier, on a dans $\mathcal{A}$: $\Theta \circ \Phi = I$.

Mais, naturellement, on n'a pas $\Phi \circ \Theta = I$, puisque, si l'on part d'un point de $\mathcal{V}$ qui n'appartienne pas à V, son image par Θ est dans $\mathcal{A}$, et par suite l'image par Φ est nécessairement dans V, donc distincte du point initial. On

* On pourra, à titre d'exercice, vérifier que toutes les propriétés exigées d'une carte sont bien réalisées.

peut dire au contraire que l'application $\Phi\circ\Theta$, qui est l'application $(y,z)\to(y,g(y))$ de $\mathcal{V}$ sur $V\cap\mathcal{V}$, est une sorte de "projection de $\mathcal{V}$ sur $V\cap\mathcal{V}$".

Corollaire 1 - Soient Φ_1 et Φ_2 deux représentations paramétriques vraies de classe C^m du même ouvert $V\cap\mathcal{V}$ d'une variété V de E. Soient $\mathcal{O}_1$ et $\mathcal{O}_2$ les ouverts de définition de Φ_1 et Φ_2 , de sorte que. $\Phi_1(\mathcal{O}_1)=\Phi_2(\mathcal{O}_2)$. Si on appelle Φ_2^{-1} l'homéomorphisme réciproque de Φ_2 , application de $V\cap\mathcal{V}$ dans $\mathcal{O}_2$, alors l'application $\Phi_2^{-1}\circ\Phi_1$, de $\mathcal{O}_1$ dans $\mathcal{O}_2$, est un homéomorphisme de classe C^m , ainsi que son homéomorphisme réciproque $\Phi_1^{-1}\circ\Phi_2$.

Démonstration - Soit $a_1\in\mathcal{O}_1$, $\Phi_1(a_1)=a$. Considérons les applications Θ_1 Θ_2 définies par le théorème, sur deux voisinages $\mathcal{V}_1,\mathcal{V}_2$ de a dans E . Soit $\mathcal{V}_0=\mathcal{V}_1\cap\mathcal{V}_2$. Appelons d'autre part $\mathcal{U}'_1$ et $\mathcal{U}'_2$ les images réciproques de $\mathcal{V}_0$ par Φ_1 et Φ_2 . Alors, dans $\mathcal{U}'_1$, l'application $\Phi_2^{-1}\circ\Phi_1$ coïncide avec l'application $\Theta_2\circ\Phi_1$ puisque, pour ut $\mathcal{U}'_1$, $\Phi_1(u)$ est dans $V\cap\mathcal{V}_0$, et que, sur $V\cap\mathcal{V}_0\subset V\cap\mathcal{V}_2$; Φ_2^{-1} coïncide avec Θ_2 ; par conséquent, comme composée de deux applications de classe C^m , elle est de classe C^m d'après le théorème 19 * . Ainsi tout point a_1 de $\mathcal{O}_1$ possède un voisinage ouvert où $\Phi_2^{-1}\circ\Phi_1$ est de classe C^m , donc elle est de classe C^m sur 6, .

Corollaire 2 - Si V est une variété de dimension n , elle ne peut pas être une variété de dimension $n'\neq n$.

Démonstration - Supposons en effet que, pour un point a de V, on ait une carte de dimension n d'un premier voisinage ouvert de a, et une carte de dimension n' d'un deuxième voisinage ouvert; alors on a, de leur intersection, 2 cartes différentes, de dimensions respectives n et n' . Mais le raisonnement que nous venons de faire au corollaire précédent, avec deux représentations paramétriques vraies, ne suppose nullement que les dimensions n , n' soient égales. Comme alors $\Phi_2^{-1}\circ\Phi_1$ doit être

* On ne peut pas faire de raisonnement directement sur $\Phi_2^{-1}\circ\Phi_1$, car Φ_2^{-1} est une application d'un ouvert de **V** , mais non d'un ouvert d'un espace affine, dans $\mathbb{K}^n$, et cela n'a donc aucun sens de dire que Φ_2^{-1} est de classe C". Cela aura un sens plus tard (page 321)

un C^m- dif'féomorphisme de $\mathcal{O}_1$ sur $\mathcal{O}_2$, le corollaire 4 du théorème 11 montre que nécessairement les dimensionssont égales.

Par contre, bien entendu, une variété de classe C^m est à fortiori de classe C^k, pour tout $k \leq m$.

Définition d'une variété par des équations implicites

Théorème 33 bis - pour qu'un ensemble V d'un espace affine E de dimension N soit une variété de dimension n et de classe C^m, il faut et il suffit que, pour tout a de V, il existe un voisinage $\mathcal{V}$ de a dans E, et un système de $N-n$ fonctions scalaires F_k, $k = 0,1,2,\ldots,N-n$ définies sur $\mathcal{V}$, de classe C^m, avec les propriétés suivantes:

1°/ Les dérivées $\overleftarrow{F'_k(a)}$ * sont $N-n$ formes linéaires indépendantes sur $\vec{E}$.

2°/ L'intersection $V \cap \mathcal{V}$ est exactement définie par les équations $F_k(x) = 0$.

On dit qu'un tel système d'équations est un système normal de N-n équations de V au voisinage de a.

Remarquons qu'on peut prendre $n = N$, alors il n'y a pas d équations, et V est un ouvert de E. On peut aussi prendre $n = 0$; alors il y a N Bquations, et, au voisinage de a, il n'y a pas d'autre solution que a lui-même; a est un point isole de V, V est un ensemble de points isolée de E, ou variété de dimension 0.

Démonstration 1°) Supposons que V soit une variété. Si alors nous utilisons la définition, nous voyons bien que cette variété est définie par les équations $F_k(x) = 0$, où F_k est définie par

$$F_k(x_1, x_2, \ldots, x_N) = x_{n+k} - G_k(x_1, x_2, \cdots, x_n) . \qquad (\text{III}, 9; 11)$$

Comme, au point a (ou en un point quelconque de $\mathcal{V}$), dF_k contient dx_{n+k} avec le coefficient 1, et aucun autre dx_{n+j}, $j = 1,2,\ldots,N-n$, les dF_k sont bien independantes.

* - $F'_k(a) \in \mathcal{L}(\vec{E}\,;\mathbb{K}) = \overleftarrow{E}'$.

2°/ Supposons rdclproquement que V soit un ensemble vérifiant les conditions de l'énoncé. Dire que le système des formes linéaires $\overrightarrow{F'_k}(a)$ est indépendant, c' est dire que, si nous choisissons un système de coordonnees dans E , l'un au moins des mineurs de rang n de la matrice des dérivées partielles est $\neq 0$. Supposons, pour fixer les idées (et on peut toujours s'y ramener par un changement éventuel de l'ordre des vecteurs de base de E), que ce soit le déterminant jacobien $\frac{D(F_1, F_2, \ldots, F_{N-n})}{D(x_{n+1}, x_{n+2}, \ldots, x_N)}(a)$.

Posons alors, comme précédemment $y = (x_1, x_2, \ldots, x_n)$ et $z = (x_{n+1}, x_{n+2}, \ldots, x_N)$, de sorte qu'un point x de E peut être Identifié à un couple (y, z) de $\mathbb{K}^n \times \mathbb{K}^{N-n}$. L'ensemble des fonctions F_k peut alors être considéré comme définissant une fonction $\vec{f}$ sur $\mathbb{K}^n \times \mathbb{K}^{N-n}$ à valeurs dans $\mathbb{K}^{N-n}$, avec $\vec{f}(a) = \vec{f}(b, c) = \vec{0}$.

L'hypothèse relative au déterminant jacobien revient exactement à dire que la dérivée partielle $\frac{\partial f}{\partial z}(a) = \frac{\partial f}{\partial z}(b, c) \in \mathcal{L}(\mathbb{K}^{N-n}; \mathbb{K}^{N-n})$ de cette fonction est inversible.
Le théorème des fonctions implicites (théorème 29) nous dit alors qu'il existe un voisinage& de b dans $\mathbb{K}^n$ et un voisinage $\mathcal{C}$ de c dans $\mathbb{K}^{N-n}$ tels que $\mathcal{B} \times \mathcal{C} \subset \mathcal{V}$, et que l'équation $\vec{f}(y, z) = \vec{0}$ admette une solution et une seule en z dans $\mathcal{C}$, lorsque y est donné dans $\mathcal{B}$. En outre, on définit ainsi une fonction implicite $z = g(y)$, qui est une application continue de $\mathcal{B}$ dans $\mathcal{C}$. Le théorème 31 nous indique en outre que, $\vec{f}$ étant de classe C^m , g est aussi de classe C^m. Nous voyons bien que l'intersection de V' avec $\mathcal{B} \times \mathcal{C}$. est exactement définie par l'équation $z = g(y)$, et que, ceci étant valable pour tout point a de V , V est bien une variété de dimension n et de classe C^m .

Donnons une autre démonstration de ce 2°) - Appelons Ψ l'application de $\mathcal{V}$ dans $\mathbb{K}^N$ définie par :

$$(x_1, x_2, \ldots, x_N) \longrightarrow (u_1, u_2, \ldots, u_N) \quad \text{avec :}$$

(III, 9; 11 bis) $$\begin{cases} u_1 = x_1 , \quad u_2 = x_2 , \ldots, u_n = x_n \\ u_{n+k} = F_k(x_1, x_2, \ldots, x_N), \quad k = 1, 2, \ldots, N-n , \end{cases}$$

en nous plaçant, comme dans la lère démonstration, dans le cas où

$$\frac{D(F_1, F_2, \ldots, F_{N-n})}{D(x_{n+1}, \ldots, x_N)}(a) \neq 0 .$$

Le déterminant jacobien de Ψ en a est $\neq 0$, parce qu'il est égal au précédent. Alors il existe un ouvert $\mathcal{V}_1 \subset \mathcal{V}$, contenant a, tel que la restriction de Ψ à $\mathcal{V}_1$ soit un homéomorphisme, de classe C^m ainsi que son homéomorphisme réciproque Φ, de $\mathcal{V}_1$ sur un ouvert $\mathcal{U}_1$ contenant $\Psi(a) = \alpha$ (théorèmes29 et 31). Mais alors Φ est un C^m- difféomorphisme de $\mathcal{U}_1$ sur $\mathcal{V}_1$, et $V \cap \mathcal{V}_1$ est l'image par Φ du sous-espace vectoriel $\mathbb{K}^n$ de $\mathbb{K}^N$: $u_{n+k} = 0$, $k = 1, 2, \ldots, N-n$;
le corollaire 2 bis du théorème 32 montre alors que V est une variété, de dimension n et de classe C^m.

Corollaire 1 - <u>Pour qu'un ensemble V de E soit une hypersurface (c'est-à-dire une variété de dimension $N-1$) de classe C^m de E, il faut et il suffit que, pour tout point a de V, il existe un voisinage $\mathcal{V}$ de a dans E, et une fonction scalaire F, définie dans $\mathcal{V}$ et de classe C^m, dont la dérivée $\overleftarrow{F'(a)}$ soit $\neq \vec{0}$ *, tels que l'ensemble $V \cap \mathcal{V}$ soit exactement défini par l'équation $F(x) = 0$.</u>

Il devient donc bien évidentqu'une sphère d'un espace euclidien est une variété, puisque la sphère de centre a et de rayon R est définie, dans E tout entier,par l'équation $(\overrightarrow{x-a} \mid \overrightarrow{x-a}) = R^2$.

Cet exemple nous montre d'ailleurs que, s'il est vrai qu'il était en général impossible de représenter une variété toute entière par des équations résolues correspondant à la définition,ou par une représentation paramétrique vraie, il est beaucoup plus possible de la définir toute entière par des équations implicites du type indiqué dans le théorème.

Naturellement la condition que les $F_k'(a)$ soient indépendantes ou, dans le cas d'une seule équation, que la dérivée $\overleftarrow{F'(a)}$ soit $\neq 0$, est absolument essentielle. Par exemple, le cône du second degré défini dans l'espace $\mathbb{R}^3$ par l'équation $x^2 + y^2 - z^2 = 0$ n'est pas une variété; il ne satisfait à aucune des définitions, à cause d'un point singulier, l'origine.

Variétés réelles et variétés complexes

Jusqu'à présent, le corps $\mathbb{K}$ pouvait être indifféremment le corps des réels ou le corps des complexes. Suivant qu'il s'agit du premier ou du deuxième, on dit que V est une variété réelle ou une variété complexe. Comme tout espace affine sur le corps des complexes est à fortiori un espace

* Rappelons que $\overleftarrow{F'(a)} \neq \vec{0}$ signifie, en prenant un référentiel de E, que les $\frac{\partial F}{\partial x_i}(a)$, $i = 1, 2, \ldots, N$, ne sont pas simultanément nulles.

affine sur le corps des réels, mais avec une dimension double, et comme de même $\mathbb{C}^n$ peut-être identifié, en tant qu'espace vectoriel sur le corps des réels, à $\mathbb{R}^{2n}$, comme d'autre part toute application dérivable par rapport au corps des complexes est à fortiori dérivable par rapport au corps des réels, nous voyons que toute variété V, de dimension n et de classe C^m par rapport au corps des complexes, peut être considéré comme une variété de dimension $2n$ et de même classe C^m par rapport au corps des réels. Quand rien de spécial n'est indiqué, la dimension d'une variété est toujours sa dimension par rapport au corps des réels.

Variétés abstraites (*)

On conçoit qu'il soit possible de définir une variété sans qu'elle soit nécessairement plongée dans un espace affine: Par exemple, l'ensemble des pages d'un atlas de géographie donne une description parfaite de la surface de la terre, sans qu'il soit nécessaire d'imaginer que cette terre, variété à deux dimensions, est auparavant plongee dans une espace affine à trois dimensions. D'où la notion de variété abstraite.

Définition - On appelle variété abstraite V, de dimension n, de classe C^m (non nécessairement plongée dans un espace affine), un espace topologique V, muni d'un atlas ou système de cartes, ayant les propriétés suivantes :

1°) Chaque carte est un homéomorphisme Φ d'un ouvert $\mathcal{O}$ de $\mathbb{K}^n$ sur un ouvert $\Phi(\mathcal{O})$ de V (Il n'est évidemment plus question de dire que cet homéomorphisme est de classe C^m, puisqu'il s'agit d une application d'un ouvert d'un espace affine dans un espace topologique, qui n'est pas plongé dans un espace affine).

2°) Soient Φ_1 et Φ_2 deux des cartes quelconques de l'atlas de V.

Si alors les images $\Phi_1(\mathcal{O}_1)$ et $\Phi_2(\mathcal{O}_2)$ ont une intersection non vide Ω, les images réciproques Ω_1 et Ω_2 de Ω par Φ_1 et Φ_2 sont des ouverts contenus dans $\mathcal{O}_1$ et $\mathcal{O}_2$; alors les restrictions de Φ_1 et Φ_2 à Ω_1 et Ω_2 sont deux applications ayant la même image; on leur impose la propriété suivante, qui est l'analogue du corollaire 1 du théorème 33 :

L'application $\Phi_2^{-1} \circ \Phi_1$ de l'ouvert Ω_1 de $\mathbb{K}^n$ sur l'ouvert Ω_2 de $\mathbb{K}^n$ est de classe C^m. Cette phrase possède un sens, puisqu'il s'agit d'une application d'un ouvert d'un espace affine dans un espace affine. C'est par cette condition qu'on retrouve la classe C^m de la variété V.

* Il ne s'agit ici que d'une esquisse. Les variétés abstraites sont étudiées en détail dans le Cours de Géométrie.

3°) L'ensemble des images des cartes doit constituer un recouvrement de V *.

L 'étude des variétés abstraites est évidemment très importante en mathématiques, et même dans beaucoup de parties de la physique. Considérons par exemple un système mécanique ayant un "nombre fini de degrés de liberté", par exemple un gyroscope dont un point de l'axe de révolution est fixe. La position de ce système peut "être définie par les valeurs d'un certain nombre de paramètres réels $q_1, q_2, \cdots, q_n$" ; en réalité ce système de paramètres est très arbitraire il est bien rare qu'on puisse, sans singularité, représenter effectivement toutes les positions du système mécanique par les valeurs d'un nombre fini de paramètres. C'est ainsi que, si l'on fixe la position du gyroscope par des angles d'Euler ψ, θ, φ, on ne définit pas là une représentation paramétrique vraie de l'ensemble des positions du gyroscope.

En réalité l'ensemble des positions du système mécanique admet une bonne definition comme variété abstraite, qui, dans le cas du gyroscope, est une variétd V de classe C^∞ à trois dimensions; mals cette variété est abstraite et n'est pas naturellement plongée dans un espace affine. Dire qu'on prend les 3 angles d Euler pour représenter la position du gyroscope, c'est dire qu'on considère une carte particulière, représentant seulement un ouvert de la variété qu'estl'ensemble des positions du gyroscope ** .

Cependant la théorie des variétés abstraites offre des complications notablement plus grandes, dans certaines questions que la théorie des variétés plongées dans les espaces affines. Lorsque cela n'introduira aucune complication , nous pnurrons prendre les variétés abstraites; toutes les fois que cela apportera une simplification. nous supposerons qu'il s'agit de variete plongees dans des espaces affines.

* On se permet généralement de compléter l'atlas initialement donné. Pour cela, on ajoute toutes les cartes qui ne perturbent pas la classe C^m, c'est-à-dlrc tous les homéomorphismes Φ d'ouverts $\mathcal{O}$ de $\mathbb{K}^m$ sur des ouverts de V, ayant la propriété suivante : pour toute carte Φ_i de l'atlas donné, telle que $\Phi(\mathcal{O}) \cap \Phi_i(\mathcal{O}_i) = \Omega \neq \emptyset$, $\Phi^{-1} \circ \Phi_i$ est un C^m-difféomorphisme de $\Phi_i^{-1}(\Omega)$ sur $\Phi^{-1}(\Omega)$.

** Pour avoir une représentation paramétrique vraie, on devra par exemple se borner à l'ouvert $0 < \psi < 2\pi$, $0 < \theta < \pi$, $0 < \varphi < 2\pi$, qui ne donne pas l'ensemble des positions du gyroscope. La variété est d'ailleurs ici compacte, il faut un nombre fini > 1 de cartes pour la représenter.

Définition - Soient V et W des variétés de classe C^m de dimensions respectives p et q, sur le même corps $\mathbb{K}$, et soit H une application continue de V dans W. On dit que H est de classe C^k, si, quels que soient le point a de V, $H(a) = b$, la carte Φ d'un voisinage $\mathcal{V}$ de a dans V, et la carte Ψ d'un voisinage $\mathcal{W}$ de b dans W, l'application composée $\Psi^{-1} \circ H \circ \Phi$ définie sur $\Phi^{-1}(\mathcal{V}) \cap \Phi^{-1}(H^{-1}(\mathcal{W}))$, ouvert de $\mathbb{K}^p$, à valeurs dans $\mathbb{K}^q$, est de classe C^k au sens habituel. On voit que, si les variétés sont de classe C^m, on ne peut pas définir d'applications de classe $C^{m'}$ pour $m' > m$. On voit aussi qu'une application de classe C^k est toujours de classe $C^{k'}$ pour $k' \leq k$. On peut toujours se borner au cas $k = m$, parce que, si V et W sont de classe C^m, $m \geq k$, elles sont à fortiori de classe C^k.

Si U, V, W sont 3 variétés de classe C^m, H une application C^m de U dans V, K une application C^m de V dans W, $K \circ H$ est évidemment C^m de U dans W. Si en effet Φ (resp. Ψ) est une carte d'un voisinage $\mathcal{U}$ de a (resp. $\mathcal{W}$ de $c = (K \circ H)(a)$) dans U (resp. W), et si nous appelons X une carte quelconque d'un voisinage $\mathcal{V}$ de $b = H(a)$ dans V, on a, dans

$$\Phi^{-1}(\mathcal{U}) \cap \Phi^{-1}(H^{-1}(\mathcal{V})) \cap \Phi^{-1}(H^{-1}(K^{-1}(\mathcal{W}))) :$$

$$\Psi^{-1} \circ K \circ H \circ \Phi = (\Psi^{-1} \circ K \circ X) \circ (X^{-1} \circ H \circ \Phi), \tag{III,9;12}$$

et chacune des 2 applications du second membre est C^m par hypothèse, donc leur composée d'après le théorème 19. On appelle C^m-difféomorphisme de V sur W un homéomorphisme, de classe C^m ainsi que son homéomorphisme réciproque. Si un tel difféomorphisme existe V et W ont même dimension, d'après le corollaire du théorème il.

Théorème 33 ter -

Soient V et W des variétés de classe C^m, H une application continue de V dans W. Pour que H soit de classe C^m, il faut et il suffit que, pour tout point a de V, il existe au moins une carte particulière Φ_a d'un voisinage de a dans V, et au moins une carte particulière Ψ_b d'un voisinage de $b = H(a)$ dans W, telles que $\Psi_b^{-1} \circ H \circ \Phi_a$ soit de classe C^m

Démonstration - La condition est trivialement nécessaire, nous devons montrer qu'elle est suffisante. Pour cela nous devons montrer que, si Φ et Ψ sont des cartes quelconques de voisinages de a. et b dans V et W, $\Psi^{-1} \circ H \circ \Phi$ est encore de classe C^m. Or on a, dans un voisinage de $\alpha = \Phi^{-1}(a)$:

$$\Psi^{-1} \circ H \circ \Phi = (\Psi^{-1} \circ \Psi_a) \circ (\Psi_a^{-1} \circ H \circ \Phi_a) \circ (\Phi_a^{-1} \circ \Phi) ; \tag{III,9;13}$$

or, dans le second membre, le terme du milieu est suppose C^m, et les termes extrêmes sont C^m d'après l'axiome 2 page 319 des cartes d'une variété, ce qui démontre le théorème.

Ce théorème a de nombreuses conséquences, donnons seulement les plus marquantes :

Corollaire 1 - Si V et W sont des ouverts d'espaces affines E, F, l'application H de V dans W est de classe C^m si et seulement si elle est de classe C^m au sens antérieur: défini page 249.

En effet prenons de V et de W une carte particulière, la carte Identique définie au 3° page : alors $\Psi^{-1} \circ H \circ \Phi$ se réduit à H.

Corollaire 2 - Soit V une variété de classe C^m d'un espace affine E. Alors l'injection canonique de V dans $W = E$ est de classe C^m.

Il suffit de prendre une carte quelconque Φ pour V, et, pour $W = E$, la carte Identique; alors $\Psi^{-1} \circ H \circ \Phi = \Phi$, et on sait que Φ est de classe C^m.

Corollaire 3 - Si V est contenue dans un espace affine E, toute application H de V dans W, restriction à V d'une application?4 de classe C^m d'un ouvert Ω de E, dans W', est de classe C^m de V dans W.

En effet, H est la composée $\tilde{H} \circ J$, où J est l'injection canonique de V dans E, de classe C^m d'après le corollaire 2.

Corollaire 4 - Soit V une variété de dimension n, de classe C^m, Φ une carte d'un ouvert $\mathcal{V}$ de V. Alors Φ est un C^m difféomorphisme de $\Phi^{-1}(\mathcal{V})$ sur $\mathcal{V}$.

Il suffit en effet, pour $\Phi^{-1}(\mathcal{V}) \subset \mathbb{K}^n$, de prendre la carte identique, et pour $\mathcal{V}$ de prendre la carte Φ ; alors, pour $H = \Phi$ ou Φ^{-1}, le $\Psi_1^{-1} \circ H \circ \Phi_1$ qu'on doit étudier est simplement l'application identique de $\Phi^{-1}(\mathcal{V})$, elle est bien de classe C^m. A Partir du corollaire 4, le corollaire 1 du théorème 33 est maintenant évident; mais nous avons dû déjà l'utiliser plusieurs fois pour obtenir les présents résultats !

Corollaire 5 - Supposons W de dimension q, plongée dans un espace affine F. Pour que H soit de classe C^m, il faut et il suffit que, pour toute carte Φ relative à V, $H \circ \Phi$ soit de classe C^m ; ou que, pour tout a de V, il existe une carte particulière Φ_a d'un voisinage de a dans V, telle que $H \circ \Phi_a$ soit de classe C^m.

Soit en effet Ψ une carte quelconque d'un voisinage de $b = H(a) \in W$. En rappetissant au besoin ce voisinage, on peut faire en sorte que Ψ^{-1} se prolonge en une application Θ de classe C^m d'un voisinage de b dans F dans $\mathbb{K}^q$. Alors, pour toute carte Φ (ou une carte particulière), nous devons montrer que $\Psi^{-1} \circ H \circ \Phi$ est de classe C^m, si et seulement si $H \circ \Phi$ est de classe C^m ; or $H \circ \Phi = \Psi \circ (\Psi^{-1} \circ H \circ \Phi)$, et $\Psi^{-1} \circ H \circ \Phi = \Theta \circ (H \circ \Phi)$, et Ψ et Θ sont de classe C^m sur des ouverts d'espaces affines, le théorème 19 donne donc le résultat.

Ce corollaire peut encore s'exprimer en disant que; si $W \subset F$, H est de classe C^m de V dans W, si et seulement si, considérée comme application de V dans F, elle est de classe C^m.

Espace vectoriel tangent en un point d'une variété d'un espace affine E de dimension N

Théorème 33 quarto Soit V une variété de dimension n, de classe C^1 dans un espace affine E. En tout point a de V le contingent vectoriel (resp. affine), au sens du théorème 8 A, est un sous-espace vectoriel de $\vec{E}$ (resp affine de E), de dimension n. Ce sous-espace est appelé espace vectoriel tangent (resp. espace affine tangent) au point a à la variété V ; il est noté $\vec{T}(a;V)$ (resp. $T(a;V)$).

Si $n = N$, l'espace vectoriel (resp. affine) tangent est $\vec{E}$ (resp. E) lui-même. Si $n = 0$, c'est $\{\vec{0}\}$ (resp $\{a\}$).

Démonstration - Il suffit de se reporter à la définition même de la variété. Il est possible, grâce au choix d'un système de coordonnées dans E d'identifier celui-ci à un produit $\mathbb{K}^n \times \mathbb{K}^{N-n}$, et de définir, au voisinage du point a, la variété par une équation $z = g(y)$. On est alors ramené au théorème 8 A, et on voit en outre, d'après la formule (III, 3; 19 bis), que le sous-espace vectoriel tangent en $a = (b, c)$, est défini, dans le système de coordonnées choisi, par l'équation

(III,9;14)
$$\vec{Z} = g'(b) \cdot \vec{Y} \,,$$

ou que la variété linéaire tangente en a, dans le même système l'équation

(III,9;15)
$$\overrightarrow{z - c} = g'(b) \cdot \overrightarrow{y - b}$$

Corollaire 1 - Soit V une variété de dimension n et de classe C^1 dans un espace affine E ; soit Φ une carte, application d'un ouvert $\mathcal{O}$ de $\mathbb{K}^n$ sur un ouvert de V, et soit α un point de $\mathcal{O}$ et $a = \Phi(\alpha)$; l'application dérivée $\Phi'(\alpha)$ est une bijection linéaire de $\mathbb{K}^n$ sur l'espace vectoriel tangent au point a à la variété V

Démonstration - supposons d'abord que Φ soit simplement une application C^1 de $\mathcal{O}$ dans V, sans être nécessairement une représentation paramétrique vraie ($\Phi'(\alpha)$ n'est pas nécessairement de rang n). Soit $\vec{X}$ un vecteur de $\mathbb{K}^n$. Considérons la suite des points $x_n = \Phi(\alpha + t_n \vec{X})$, où les t_n forment une suite de nombres réels >0, tendant vers 0 pour n tendant vers l'infini. Alors les points x_n appartiennent à V, tendent vers a, et d'autre part les vecteurs $\frac{1}{t_n}(\overrightarrow{x - a})$, qui peuvent encore s'écrire $\overline{\frac{\Phi(\alpha + t_n \vec{X}) - \Phi(\alpha)}{t_n}}$, convergent vers la dérivée de Φ suivant $\vec{X}$ au point α, c'est-à-dire vers $\Phi'(\alpha) \cdot \vec{X}$. Ainsi ce vecteur appartient nécessairement à l'espace vectoriel tangent au point a ; il en résulte que l'image par $\Phi'(\alpha)$ de $\mathbb{K}^n$ est contenue dans l'espace vectoriel tangent, si maintenant nous tenons compte de l'hypothèse faite sur $\Phi'(\alpha)$, à savoir qu'elle est de rang n, cela entraîne que cette image soit un sous-espace vectoriel de dimension n, c'est nécessairement l'espace vectoriel tangent en a tout entier; et l'application $\Phi'(\alpha)$ est bien une bijection linéaire de $\mathbb{K}^n$ sur cet espace vectoriel tangent.

Il en résulte que les vecteurs $\overline{\frac{\partial \Phi}{\partial u_j}(\alpha)}$ forment une base de l'espace vectoriel tangent en a à V, et que la variété linéaire tangente en a à V est représentée par l'équation paramétrique

(III, 9;16) $$x = a + \sum_{j=1}^{n} t_j \overrightarrow{\frac{\partial \Phi}{\partial u_j}(\alpha)}, \qquad t_j \in \mathbb{K} \quad (*)$$

Corollaire 1 bis - Soient V, W, deux variétés de classe C^1 de deux espaces affines E, F. Soit H une application de classe C^1 de E dans F, telle-que $H(V) \subset W$. Alors, pour tout $a \in V$, si $b = H(a) \in W$, l'image par $H'(a) \in \mathcal{L}(\vec{E}; \vec{F})$ de l'espace vectoriel tangent $\vec{T}(a; V)$ est contenue dans l'espace vectoriel tangent $\vec{T}(b; W)$: $H'(a) \cdot \vec{T}(a;V) \subset \vec{T}(b;W)$

Démonstration - Soit Φ une carte $(\mathcal{O} \subset \mathbb{K}^n) \to \Phi(\mathcal{O})$ d'un voisinage $\Phi(\mathcal{O})$ de a dans V, $\Phi(\alpha) = a$. Alors $H \circ \Phi$ est une application C^1 de $\mathcal{O}$ dans W ; le début de la démonstration du corollaire précédent montre que

* Comme cas très particulier, si V est une surface de R^3, d'équation paramétrique vraie $(u, v) \to M(u, v)$, le plan tangent en un point est engendré par les dérivées partielles $\frac{\partial \vec{M}}{\partial u}$, $\frac{\partial \vec{M}}{\partial v}$ en ce point.

$(H\circ\Phi)'(\alpha).\mathbb{K}^n \subset \vec{T}(b;W)$. Mais $(H\circ\Phi)'(a) = H'(a)\circ\Phi'(\alpha)$ d'après le théorème des fonctions composées, et $\Phi'(\alpha).\mathbb{K}^n = \vec{T}(a;V)$ d'après le corollaire 1, d'où le résultat.

Corollaire 2 - Si on se place dans les conditions du corollaire 1. et si $\mathcal{O}$ est un ouvert de $\mathbb{K}^n$ contenant α, tel que la restriction de Φ à $\mathcal{O}$ admette une application réciproque à gauche Θ dans les conditions du théorème 33, alors la bijection linéaire $\Phi'(\alpha)$ de $\mathbb{K}^n$ sur $\vec{T}(a;V)$ admet pour bijection réciproque la restriction à $\vec{T}(a;V)$ de l'application dérivée $\Theta'(a)$.

Démonstration - De l'identité $\Theta\circ\Phi = I$, on déduit $\Theta'(a)\circ\Phi'(\alpha) = I$ (théorème 11, fonctions composées), qui prouve bien que la restriction de $\Theta'(a)$ à l'image par $\Phi'(\alpha)$ de $\mathbb{K}^n$ c'est-à-dire à $\vec{T}(a;V)$, est la bijection réciproque de $\Phi'(\alpha)$.

Corollaire 3 - Soient Φ_1 et Φ_2 des applications d'ouverts $\mathcal{O}_1$ et $\mathcal{O}_2$ de $\mathbb{K}^n$ dans $\mathcal{V}$, formant des cartes de même image dans la variété V plongée dans E ; soit $\vec{X}$ un vecteur tangent en a à V, $a = \Phi_1(\alpha_1) = \Phi_2(\alpha_2)$; si $\vec{\xi}_1$ et $\vec{\xi}_2$ sont les vecteurs de $\mathbb{K}^n$ dont il est l'image par les applications $\Phi_1'(\alpha_1)$ et $\Phi_2'(\alpha_2)$, alors $\vec{\xi}_2$ est l'image de $\vec{\xi}_1$ par la dérivée au point α_1 de l'application de classe C^1 : $\Phi_2^{-1}\circ\Phi_1$

Il suffit en effet de remarquer que l'on peut écrire

(III,9;17) $$\vec{\xi}_2 = \left(\left(\Phi_2'(\alpha_2)\right)^{-1}\circ\Phi_1'(\alpha_1)\right)\cdot\vec{\xi}_1$$

Mais précisément, d'après le corollaire précédent, nous avons vu que l'application réciproque de $\Phi_2'(\alpha_2)$ est aussi la restriction, au sous-espace vectoriel en a, de $\Theta_2'(a)$; on peut donc aussi écrire que $\vec{\xi}_2$ est relié à $\vec{\xi}_1$ par

(III,9;18) $$\vec{\xi}_2 = \left(\Theta_2'(a)\circ\Phi_1'(\alpha_1)\right)\cdot\vec{\xi}_1 .$$

Mais alors Φ_1 et Θ_2 sont cette fois-ci des applications d'un ouvert d'un espace affine dans un espace affine.

Nous pouvons donc appliquer le théorème des fonctions composées (théorème 11), et remplacer l'égalité précédente par

(III,9;19) $$\vec{\xi}_2 = \left(\Theta_2\circ\Phi_1\right)'(\alpha_1)\cdot\vec{\xi}_1 .$$

Mais enfin, d'après les propriétés de Θ, on voit que $\Theta_2\circ\Phi_1$ coïncide avec $\Phi_2^{-1}\circ\Phi^1$ (c'est précisément comme cela que nous avons vu au théorème 33 que $\Phi_2^{-1}\circ\Phi_1$ était de classe C^1), on peut donc encore écrire

(III,9;20) $$\vec{\xi}_2 = (\Phi_2^{-1} \circ \Phi_1)'(\alpha_1) \cdot \vec{\xi}_1 ,$$

et ceci démontre le corollaire.

Corollaire 4 - Soit V une variété, de dimension n, de classe C^1, dans un espace affine E. Soit a un point de V, et soit $F_k(x) = 0$, $k = 1, 2, \ldots, N-n$, un système normal d'équations de V dans un voisinage de a dans E. Alors le sous-espace vectoriel tangent en a à V est défini dans $\vec{E}$ par les équations

(III,9;21) $$F_k'(a) \cdot \vec{X} = 0 , \qquad k = 1, 2, \ldots, N-n ,$$

et la variété linéaire tangente en a à V est définie dans E par les équations

(III,9;22) $$F_k'(a) \cdot \overrightarrow{x-a} = 0 , \qquad k = 1, 2, \ldots, N-n ,$$

Démonstration - Soit x_n une suite de points de V tendant vers a, et λ_n une suite de scalaires réels > 0 telle que la suite des vecteurs $\lambda_n(\overrightarrow{x_n - a})$ tende vers le vecteur $\vec{X}$ de $\vec{E}$, pour n tendant vers l'infini. Comme alors on a à la fois $F_k(x_n) = 0$ et $F_k(a) = 0$, la définition même de la dérivée (formule (III,3;13)) donne

(III,9;23) $$F_k'(a) \cdot \overrightarrow{x_n - a} + \alpha_n \|\overrightarrow{x_n - a}\| = 0 ,$$

$\|\vec{\alpha}_n\|$ tendant vers 0, pour n infini; d'où, par multiplication par λ_n :

(III,9;24) $$F_k'(a) \cdot \lambda_n(\overrightarrow{x_n - a}) + \vec{\alpha}_n \lambda_n \|\overrightarrow{x_n - a}\| = 0 .$$

Comme $\lambda_n(x_n - a)$ converge vers $\vec{X}$ et comme $F_k'(a)$ est une forme linéaire continue, on voit que nécessairement chaque vecteur tangent% en a à V vérifie les $N-n$ équations linéaires (III,9;21).

Mais, comme ces $N-n$ équations linéaires sont supposées indépendantes elles définissent précisément un sous-espace vectoriel de dimension n de $\vec{E}$; c'est donc exactement l'espace vectoriel tangent en a lui-même.

Le résultat relatif à la variété linéaire tangente s'en déduit immédiatement.

Corollaire 5 - Si E est un espace affine euclidien de dimension N sur le corps des réels et si V est une hypersurface définie par une équation $F(x) = 0$, où F est une fonction réelle de classe C^1 dont le gradient au point a est $\neq 0$, ce gradient est normal en a à V, et le vecteur $\dfrac{\mathrm{grad}\, F(a)}{\|\mathrm{grad}\, F(a)\|}$ est un vecteur unitaire de la normale en a à V *

*) Ce renvoi se trouve à la page suivante.

Corollaire 6 - Soit V une variété de dimension n et de classe C^1 dans un espace affine E, soit Φ une carte appliquant un ouvert $\mathcal{O}$ de $\mathbb{K}^n$ sur un ouvert de V. Alors l'application $(u, \vec{U}) \longrightarrow (\Phi(u), \Phi'(u)\cdot\vec{U})$, qui fait correspondre à tout couple d'un point u et d'un vecteur $\vec{U}$ de $\mathbb{K}^n$, le couple du point image $\Phi(u)$, point de V, et du vecteur $\Phi'(u)\cdot\vec{U}$ tangent en ce point à V, est un homéomorphisme de $\mathcal{O}\times\mathbb{K}^n$ sur son image.

Démonstration - C'est une application continue (théorème 10) et bijective. Pour démontrer que c'est un homéomorphisme, il suffit donc de savoir que son application réciproque est continue. Comme, par ailleurs, la continuité est une propriété locale, on peut se borner à restreindre l'application réciproque à l'ensemble des couples $(x, \vec{X})$, pour lesquels x parcourt un voisinage de a dans V. Utilisant alors le théorème 33, et choisissant ce voisinage comme il est indiqué dans ce théorème, on voit (corollaire 2) que l'application réciproque n'est autre que la restriction de l'application

$(x, \vec{X}) \rightarrow (\Theta(x), \Theta'(x)\cdot\vec{X})$ de $\mathcal{V}\times\vec{E}$ dans $\mathbb{K}^n\times\mathbb{K}^n$.

Cette application est continue, encore en vertu du théorème 10.

Espace vectoriel tangent en un point d'une variété abstraite (**)

Soit V une variété abstraite de classe C^1, de dimension n, et soit $a \in V$. Comment peut-on définir un vecteur tangent en a à V ? Soit $\Phi : \mathcal{O} \longrightarrow \Phi(\mathcal{O})$ une carte d'un ouvert $\Phi(\mathcal{O})$ de V contenant a, et soit $\alpha \in \mathbb{K}^n$ tel que $\Phi(\alpha) = a$. Il est intuitif que, si $\vec{\xi}$ est un vecteur quelconque de $\mathbb{K}^n$, on doit pouvoir, par Φ, lui faire correspondre un vecteur tangent $\vec{X}$ au point a à la variété V. Mais alors soient $\Phi_1 : \mathcal{O}_1 \longrightarrow \Phi_1(\mathcal{O}_1)$ et $\Phi_2 : \mathcal{O}_2 \longrightarrow \Phi_2(\mathcal{O}_2)$ deux cartes d'ouverts de la variété V contenant a. Soient $\vec{\xi}_1$ et $\vec{\xi}_2$ deux vecteurs de $\mathbb{K}$. Quand dira-t-on qu'il leur correspond, par Φ_1 et Φ_2 le même vecteur tangent $\vec{X}$ en a à V ? Soit Ω l'intersection $\Phi_1(\mathcal{O}_1) \cap \Phi_2(\mathcal{O}_2)$, et soient $\Omega_1 = \Phi_1^{-1}(\Omega)$ et $\Omega_2 = \Phi_2^{-1}(\Omega)$.

* - Renvoi de la page **précédente**.

Si on choisit une base orthonormale de E, on a :

$$\frac{\overrightarrow{\text{grad}}\, F(a)}{\|\overrightarrow{\text{grad}}\, F(a)\|} = \left(\frac{\frac{\partial F}{\partial x_1}(a)}{\sqrt{\sum_{i=1}^{N}\left(\frac{\partial F}{\partial x_i}(a)\right)^2}}, \frac{\frac{\partial F}{\partial x_2}(a)}{\sqrt{\sum_{i=1}^{N}\left(\frac{\partial F}{\partial x_i}(a)\right)^2}}, \ldots, \frac{\frac{\partial F}{\partial x_N}(a)}{\sqrt{\sum_{i=1}^{N}\left(\frac{\partial F}{\partial x_i}(a)\right)^2}} \right)$$

** Il ne s'agit ici que d'une esquisse de la théorie. Cette notion est étudiée en détail dans le Cours de Géométrie.

Alors on sait que $\Phi_2 \circ \Phi_1$ est un C^m-difféomorphisme de Ω_1 sur Ω_2. Il admet donc au point $a_1 \in \Omega_1$ une application dérivée $(\Phi_2^{-1} \circ \Phi_1)'(\alpha_1)$. Nous conviendrons que $\vec{\xi}_1$ et $\vec{\xi}_2$ représentent, par Φ_1 et Φ_2 le même vecteur $\vec{X}$ tangent en a à la variété V si l' on a la relation

$$(\text{III},9;24\text{ bis}) \qquad \vec{\xi}_2 = (\Phi_2^{-1} \cdot \Phi_1)'(\alpha_1) \cdot \vec{\xi}_1 \; .$$

Cette définition est justifiée par le corollaire 3 du théorème 33 quarto.

Cela revient à donner la définition suivante : On appelle vecteur $\vec{X}$ tangent en a à la variété V une famille $(\vec{\xi}_i)_{i \in I}$ de vecteurs de $\mathbb{K}^n$ (où I, ensemble d'indices, est l'ensemble de toutes les cartes Φ_i de la variété V pour lesquelles $\Phi_i(\mathcal{O}_i)$ contient a), telle que l'on ait les relation: $(\Phi_j^{-1} \circ \Phi_i)'(\alpha_i) \cdot \vec{\xi}_i = \vec{\xi}_j$, $i, j \in I$. On définit sans ambiguïté la somme de deux vecteurs tangents $(\vec{\xi}_i)_{i \in I}$, $(\vec{\eta}_i)_{i \in I}$, par la famille $(\vec{\xi}_i + \vec{\eta}_i)_{i \in I}$. On définit de même la multiplication du vecteur tangent $(\vec{\xi}_i)_{i \in I}$ par $\lambda \in \mathbb{K}$ par $(\lambda \vec{\xi}_i)_{i \in I}$. Les vecteurs tangents en a à la variété V forment alors un espace vectoriel, dont on vérifie immédiatement qu'il a la dimension n, dimension de la variété V elle-même. On appelle $\vec{T}(a ; V)$ l'espace vec toriel tangent en a à la variété V. Il est bon de remarquer que, si a et b sont deux points distincts de V les espaces vectoriels tangents $\vec{T}(a ; V)$ et $\vec{T}(b ; V)$ n'ont aucun rapport simple l'un avec l'autre. Ils ne sont pas, comme dans le cas d une variété contenue dans un espace affine, sous-espaces vectoriels d'un même espace vectoriel donné à l'avance.

soient V et W deux variété de classe C^1, de dimensions respectives p et q. Soit H une application de slasse C^1 de V dans W, $b = H(a)$. Reprenons les natations de la définition donnée page 321 : Φ est une carte d'un voisinage de a dans V, $\Phi(\alpha) = a$, et Ψ une carte d'un voisinage de b dans W. Soit $\vec{\xi}$ un vecteur de $\mathbb{K}^n$, définissant par Φ un vecteur $\vec{X}$ tangent en a à V.

Son image par l'application dérivée $(\Psi^{-1} \circ H \circ \Phi)'(\alpha)$ est un vecteur $\vec{\eta}$ de $\mathbb{K}^q$ définissant donc, par Ψ, un vecteur $\vec{Y}$ tangent en b à W. On démontre alors aisément que $\vec{Y}$ ne dépend que de $\vec{X}$ et de a, et non des cartes Φ et Y choisies; on peut donc dire que la donnée de H définit une application, manifestement linéaire, de $\vec{T}(a ; V)$ dans $\vec{T}(b ; W)$ On dit que cette application est l'application dérivé; de H au point a et on la note $H'(a)$.

Cette application dérivée est la généralisation de celles que nous avions vues jusqu'à présent pour une application de classe C^1 d'un espace affine dans un autre. En effet, si V et W sont des ouverts des espaces $\mathbb{K}^p$ et $\mathbb{K}^q$, on peut prendre pour Φ et Y l'application identique; l'espace vectoriel tangent en a à V est alors identifié à $\mathbb{K}^p$, et l'espace vectoriel tangent en b à W est identifié à $\mathbb{K}^q$, avec $\vec{X} = \vec{\xi}$ et $\vec{Y} = \vec{\eta}$; l'application dérivée que nous venons de définir n'est autre que l'application $H'(a)$ de $\mathbb{K}^p$ dans $\mathbb{K}^q$. On vérifie sans peine que les applications dérivées généralisées ainsi définies satisfont au théorème des fonctions composées (théorème 11). Par ailleurs, cette définition de $H'(a) \in \mathcal{L}(\vec{T}(a;V) ; \vec{T}(b;W))$ est rendue naturelle par le corollaire 1 bis du théorème 33 quarto.

Si d'autre part V est une variété contenue dans un espace affine E, on peut prendre pour H l'injection de V dans E; l'application dérivée $H'(a)$ définit alors une application linéaire de $\vec{T}(a;V)$ dans $\vec{T}(a;E) = \vec{E}$; on vérifie que cette application linéaire est une injection; elle permet donc d'identifier $\vec{T}(a;V)$ à son Image, c'est-à-dire à un sous-espace vectoriel de $\vec{E}$, et ce sous-espace n'est autre que celui que nous avions trouvé antérieurement comme étant le sous-espace vectoriel tangent en a à la variété V contenue dans l'espace affine E.

Nous n'insistons pas sur cette notion assez compliquée d'espace vectoriel tangent à une variété abstraite; elle est ici simplement destinée à montrer que ce que nous avons vu jusqu'à présent, relativement aux espaces affines, admet des généralisations aux variétés différentiables quelconques.

Théorème du rang constant

__Théorème 33 quinto__ - Soit $\mathcal{O}$ un ouvert de $\mathbb{K}^n$, et soit $\Phi : u \to u = \Phi(u)$, __une application de classe C^m de $\mathcal{O}$ dans un espace affine E, telle qu'en tout point α de $\mathcal{O}$ l'application dérivée $\Phi'(\alpha)$ soit du même rang $\ell \leq n$. Alors, pour tout point α de $\mathcal{O}$ il existe un ouvert $\mathcal{U}$ contenu dans $\mathcal{O}$ et contenant α, dont l'image par Φ est une variété V de E, de dimension ℓ et de classe C^m.__

__Démonstration__ - Prenons dans E un système de coordonnées. Dire que le rang de $\Phi'(\alpha)$ est ℓ, c'est dire d'abord que l'un au moins des déterminants à ℓ lignes et ℓ colonnes de la matrice dérivée de Φ au point α est $\neq 0$. En changeant au besoin l'ordre des vecteurs de la base de E, nous pouvons supposer que c'est le déterminant jacobien $\dfrac{D(x_1, x_2, \ldots, x_\ell)}{D(u_1, u_2, \ldots, u_\ell)}(\alpha)$.

Comme alors Φ est supposée de classe C', le déterminant jacobien $\frac{D(x_1, x_2, \ldots, x_\ell)}{D(u_1, u_2, \ldots, u_\ell)}$ est une fonction scalaire continue sur 6, et il existe un voisinage $\mathcal{O}_0$ de α, dans lequel ce déterminant est partout $\neq 0$. Comme par ailleurs, dans $\mathcal{O}$ et par conséquent dans $\mathcal{O}_0$, le rang de l'application dérivée est partout ℓ , nous voyons que, dans $\mathcal{O}_0$, tout mineur de la matrice dérivée contenant le déterminant précédent est nécessairement nul. Décomposons $\mathbb{K}^n$ en un produit $\mathbb{K}^\ell \times \mathbb{K}^{n-\ell}$; tout pointu de $\mathbb{K}^n$ sera alors appelé (v, w), avec $v = (u_1, u_2, \ldots, u_\ell)$, $w = (u_{\ell+1}, \ldots, u_n)$; soit $a = (\beta, \gamma)$. De la même manière, décomposons E, grâce à son système de coordonnées, en un produit $\mathbb{K}^\ell \times \mathbb{K}^{N-\ell}$, en posant $x = (y, z)$, avec $y = (x_1, x_2, \ldots, x_\ell)$, $z = (x_{\ell+1}, \ldots, x_N)$; soit $a = \Phi(\alpha) = (b, c)$, et soit $y = X(v,w)$, $z = \Psi(v, w)$ l'expression de $x = \Phi(u)$ dans cette décomposition. Alors, d'après l'hypothèse faite sur notre déterminant jacobien, il résulte du théorème des fonctions Implicites (théorèmes 28 et 31) * qu'il existe des ouverts $\mathcal{O}'$ et $\mathcal{O}''$ contenant respectivement β et γ dans $\mathbb{K}^\ell$ et $\mathbb{K}^{n-\ell}$, $\mathcal{O}' \times \mathcal{O}'' \subset \mathcal{O}_0$, et un ouvert $\mathcal{B}$ contenant b dans $\mathbb{K}^\ell$, tels que, pour w et y donnés arbitrairement dans $\mathcal{O}''$ et $\mathcal{B}$, il existe un élément v et un seul de $\mathcal{O}'$, pour lequel on ait $X(v,w) = y$, de plus, la fonction ainsi déterminée $v = \Lambda(y, w)$ est de classe C^m de $\mathcal{B} \times \mathcal{O}''$ dans $\mathcal{O}'$.
Comme nous pouvons toujours remplacer& et $\mathcal{O}''$ par des voisinages plus petits, sans changer $\mathcal{O}'$, à cause de la continuité de Λ, nous prendrons pour $\mathcal{O}''$ une boule de centre γ.
Appelons alors $\mathcal{O}_1$ le produit $\mathcal{O}' \times \mathcal{O}''$, et W son Image par Φ.
si $x = (y,z)$ appartient à W, et si en outre y appartient à $\mathcal{B}$, alors on a nécessairement $v = \Lambda(y, w)$, et par suite $z = \Psi(\Lambda(y, w), w) = g(y,w)$. Montrons qu'en fait g ne dépend pas de w.

Pour cela calculons la différentielle dg en un point (y, w) déterminé.

Pour faire ce calcul, d'après la règle (III, ;), on doit commencer par différencier X sous la forme

(III,9;25) $$\overrightarrow{dy} = \frac{\partial X}{\partial v}(v, w) \cdot \overrightarrow{dv} + \frac{\partial X}{\partial w}(v, w) \cdot \overrightarrow{dw} .$$

* On a ici à résoudre $X(v,w) = y$ en v, pour w et y donnés. Voir à ce sujet la remarque page

Le fait que la déterminant jacobien $\frac{D(x_1,x_2,\dots,x_\ell)}{D(u_1,u_2,\dots,u_\ell)}$ soit $\neq 0$,signifie que la dérivée partielle $\frac{\partial X}{\partial v}(v,w)\in\mathcal{L}(\mathbb{K}^\ell;\mathbb{K}^\ell)$ est inversible; on peut alors résoudre l'équation précédente sous la forme

$$(\text{III},9;26)\qquad \overrightarrow{dv} = \left(\frac{\partial X}{\partial v}(v,w)\right)^{-1}\cdot\overrightarrow{dy} - \left(\left(\frac{\partial X}{\partial v}(v,w)\right)^{-1}\circ\frac{\partial X}{\partial w}(v,w)\right)\cdot\overrightarrow{dw}$$

$$= \frac{\partial\Lambda}{\partial y}(y,w)\cdot\overrightarrow{dy} + \frac{\partial\Lambda}{\partial w}(y,w)\cdot\overrightarrow{dw}\ .$$

En portant le résultat trouvé dans la différentielle de Ψ :

$$(\text{III},9;27)\qquad \overrightarrow{dz} = \frac{\partial\Psi}{\partial v}(v,w)\cdot\overrightarrow{dv} + \frac{\partial\Psi}{\partial w}(v,w)\cdot\overrightarrow{dw}\ ,$$

on obtient la différentielle cherchée.

Mais alors, le fait que tous les mineurs de rang $>\ell$ de la matrice dérivée de Φ soient nuls dans $\mathcal{O}'\times\mathcal{O}''$, signifie précisément que les $dx_{\ell+k}$, $k=1,2,\dots,N-\ell$, ne dépendent que de $dx_1, dx_2,\dots, dx_\ell$, et non de $du_{\ell+1},\dots,du_n$, autrement dit que, dans le résultat précédent, qui s'exprime sous la forme développée

$$(\text{III}\ 9;28)\qquad \overrightarrow{dz} = \frac{\partial\Psi}{\partial v}\cdot\left[\left(\frac{\partial X}{\partial v}\right)^{-1}\cdot\overrightarrow{dy} - \left(\frac{\partial X}{\partial v}\right)^{-1}\cdot\frac{\partial X}{\partial w}\cdot\overrightarrow{dw}\right] + \frac{\partial\Psi}{\partial w}\cdot\overrightarrow{dw}\ ,$$

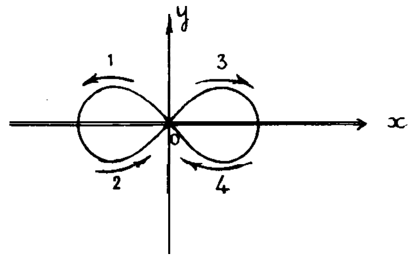
<u>$\overrightarrow{dz}$ ne dépend en fait que de $\overrightarrow{dy}$ et non de $\overrightarrow{dw}$</u> . Autrement dit la l'onction g définie sur $\mathcal{B}\times\mathcal{O}''$ à valeurs dans $\mathcal{O}'$, a sa dérivée partielle $\partial g/\partial w$ identiquement nulle. Comme nous avons choisi pour $\mathcal{O}''$ une boule, c'est-à-dire un ensemble connexe, on peut appliquer, pour y fixé, le théorème 22 à la fonction $w\to g(y,w)$ sur $\mathcal{O}''$, et voir par conséquent que, pour y fixé, $g(y,w)$ est une constante, autrement dit que g est indépendant de w. C'est une fonction de y seul, définie sur $\mathcal{B}$. Il en résulte que l'ensemble des points (y,z) appartenant à W, et pour lesquels y est dans $\mathcal{B}$, est une variété V de dimension ℓ et de classe C^m , définie par l'équation $z=g(y)$; cet ensemble V de points de E est l'image par Φ de l'intersection de $\mathcal{O}_1$ et de l'ensemble $X^{-1}(\mathcal{B})$ des points u tels que $X(u)\in\mathcal{B}$, c'est-à-dire de l'intersection de deux ouverts donc d'un ouvert $\mathcal{O}$ de $\mathcal{O}_1$ contenant α • ainsi l'image $\Phi(\mathcal{O})$, d'après ce que nous venons de voir: est la variété V de classe C^m , de dimension ℓ , et le théorème est démontré.

Z

Remarques 1°) Même si l'application Φ est une bijection, et même si $\ell = n$, on aurait tort de croire que l'image par Φ de l'ouvert $\mathcal{O}$ tout entier soit une variété de E.

Considérons par exemple une lemniscate de Bernoulli dans le plan $\mathbb{R}^2$. Prenons, sur cette lemniscate, le sens de parcours 1,2,3,4 indiqué par les flèches. On peut trouver une bijection de classe C^∞, $\Phi : u \to \Phi(u)$, de la droite réelle $\mathbb{R}$ dans le plan $\mathbb{R}^2$, telle que $\Phi(\mathbb{R})$ soit exactement la lemniscate, $\Phi(0)$ étant le point double de la lemniscate, et de manière que, lorsque u tend vers $-\infty$ (resp. $+\infty$), $\Phi(u)$ tende vers le point double sur la branche 1 (resp. 4).

$$(\text{III},9;29) \quad \begin{cases} x = \dfrac{u(1+u^2)}{1+u^4} \\[2ex] y = \dfrac{u(1-u^2)}{1+u^4} \end{cases}$$



La dérivée $\Phi'(\alpha)$ est de rang 1 en tout point α de $\mathbb{R}$, autrement dit $x'(u)$ et $y'(u)$ ne sont jamais simultanément nuls; en outre Φ est une bijection *. Cependant Φ n'est évidemment pas un homéomorphisme, car on peut trouver des points de la lemniscate convergeant vers le point double, et pour lesquels u, au lieu de converger vers 0, converge vers $+\infty$ ou vers $-\infty$; d'ailleurs la lemniscate est compacte et $\mathbb{R}$ ne l'est pas. L'image $\Phi(\mathbb{R})$ n'est pas une variété, à cause de son point singulier à l'origine; cependant, si l'on considère le point $\alpha = 0$, il est possible de trouver un voisinage de ce point, par exemple l'intervalle $]-A, +A[$, où A est un nombre > 0 quelconque, tel que l'image de cet intervalle par Φ soit une variété.

2°) Supposons $n \geq N$ et $\ell = N$. L'hypothèse que $\Phi'(\alpha)$ est de rang $\ell = N$, pour tout α de $\mathcal{O}$, signifie simplement que c'est une surjection de $\mathbb{K}^n$ sur $\vec{E}$. Alors $\Phi(\mathcal{U})$, variété de dimension N d'un espace affine de dimension N, est simplement un ouvert. Donc, si Ω est un ouvert de $\mathcal{O}$, $\Phi(\Omega)$ est un voisinage de chacun de ses points, donc un ouvert, et nous retrouvons le fait que Φ est une application ouverte (théorème 30). La première partie du théorème 30 en résulte aussi; si

* Le point double n'est obtenu que pour $u = 0$.

on sait seulement que $\Phi'(\alpha_0)$ est de rang N, pour un point α_0 particulier, il y a au moins un mineur de rang N de la matrice dérivée (pour un référentiel quelconque de E) qui est $\neq 0$ en α_0, mais, ce mineur étant continu, il est $\neq 0$ en tous les points α d'un ouvert $\mathcal{O}_1 \subset \mathcal{O}$ contenant α_0, Φ est une application ouverte de $\mathcal{O}_1$ dans E, et par suite l'image par Φ de tout voisinage de α_0 est un voisinage de $a_0 = \Phi(\alpha_0)$ dans E.

3°) Dans le cas où $\ell < n$, l'application Φ n'est naturellement pas injective et chacun des points de la variété $V = \Phi(\mathcal{U})$ est l'image d'une infinité de points de $\mathcal{U}$. Si en effet $x_0 = (y_0, z_0)$ est un point de V, on pourra encore choisir w arbitrairement dans $\mathcal{U}''$, calculer alors $v = \Lambda(y_0, w)$, et le point x_0 considéré sera l'image du point $u = (v, w) = (\Lambda(y_0, w), w)$. Ainsi, lorsqu'on restreint Φ à $\mathcal{U}$, l'image réciproque de x_0 est l'ensemble des points (v, w) de $\mathcal{U}$ pour lesquels on a $w \in \mathcal{U}''$, $v = \Lambda(y_0, w)$, c'est donc une variété de $\mathbb{K}^n$, de dimension $n-\ell$, et de classe C^m, définie par la relation explicite précédente. Au contraire, Φ est injective (pour $\mathcal{U}$ assez petit) si $\ell = n$.

4°) Naturellement le présent théorème redonne aussi le théorème 32. Si en effet pour tout a de V, il existe une représentation paramétrique vraie Φ, $\mathcal{O} \to \Phi(\mathcal{O})$, $\Phi(\alpha) = a$, Φ partout de rang $\ell = n$, alors le théorème 33 quinto dit qu'il existe un voisinage ouvert $\mathcal{U}$ de α dans $\mathcal{O}$ tel que $\Phi(\mathcal{U}) = W$ soit une variété. Mais Φ est en outre supposé être un homéomorphisme de $\mathcal{O}$ sur un ouvert de V ; alors $\Phi(\mathcal{U})$ est un ouvert de V, donc il existe un voisinage ouvert $\mathcal{V}_1$ de a dans E tel que $W = V \cap \mathcal{V}_1$; mais alors, si tout point a de V a un voisinage ouvert $\mathcal{V}_1$ dans E tel que $V \cap \mathcal{V}_1$ soit une variété, on déduit aussitôt de la définition que V elle-même est une variété.

5°) Lorsque le rang de l'application Φ n'est pas constant, il n'existe plus aucun théorème permettant d'affirmer que l'image d'un voisinage de α par Φ ait une structure simple. Considérons, par exemple, l'application Φ de $\mathbb{R}^2$ dans $\mathbb{R}^3$ définie par les formules $(r, \varphi) \to (x, y, z)$ avec :

$$(\mathrm{III},9;30) \qquad \left\{ \begin{aligned} x &= r \cos \varphi \\ y &= r \sin \varphi \\ z &= r \end{aligned} \right.$$

En un point $\alpha = (0, \varphi_0)$, le rang de l'application dérivée de Φ est égal à 1, parce que $\partial\Phi/\partial\varphi$ est nulle; mais, aux points voisins, le rang est égal à 1 ou à 2. L'image par Φ de $\mathbb{R}^2$ n'est autre que le cône de $\mathbb{R}^3$ d'équation $x^2 + y^2 - z^2 = 0$, dont Φ définit une représentation paramdtrlque non vraie. L'image par Φ de tout voisinage du pointu précédent est un voisinage du sommet du cône, qui n'est pas une variété à cause du point singulier au sommet.

Signalons à ce sujet la propriété suivante :

Si Φ est une application de classe C' d'un ouvert $\mathcal{O}$ de $\mathbb{K}^n$ dans E le rang $\ell(\alpha)$ de l'application dérivée $\Phi'(\alpha)$ est une fonction semi-continue inférieurement de α.

Nous allons voir en effet que, si $\alpha_0 \in \mathcal{O}$, il existe un ouvert $\mathcal{U}_0$ de $\mathcal{O}$ contenant a, , tel que, pour tout α de $\mathcal{U}_0$, $\ell(\alpha) \geq \ell(\alpha_0)$. Prenons en effet un référentiel dans E. Il existe au moins un mineur de rang $\ell(\alpha_0)$ de la matrice dérivée de Φ en α_0 qui soit $\neq 0$, d'après la définition même du rang; mais, Φ étant de classe C', ce mineur est une fonction continue de a, donc il existe bien un voisinage $\mathcal{U}_0$ de a, , où ce mineur est partout #o ; alors, en tout pointa de $\mathcal{U}_0$, le rang de $\Phi'(\alpha)$ est bien $\geq \ell(\alpha_0)$.

Fonctions dépendantes et fonctions indépendantes

Soient $\Phi_1, \Phi_2, \dots, \Phi_N$, N fonctions scalaires de classe C^1 de n variables $u_1, u_2, \dots, u_n$. Elles définissent une application $\Phi : u \longrightarrow x = \Phi(u)$, de classe C', d'un ouvert $\mathcal{O}$ de $\mathbb{K}^n$ dans $\mathbb{K}^N$.

Quand dira-t-on que ces fonctions sont dépendantes ou Indépendantes, au voisinage d'un point a de $\mathcal{O}$? Soit a $= \Phi(\alpha)$.

1er Cas - Indépendance.

Supposons que le rang de $\Phi'(\alpha)$ soit N, le nombre des fonctions. Autrement dit, l'un des mineurs de rang N de la matrice dérivée, matrice des $\frac{\partial\Phi_i}{\partial u_j}(\alpha)$, $i = 1,2,\dots,N$, $j = 1,2,\dots n$, est $\neq 0$; cela exige évidemment $n \geq N$ le nombre des variables doit être au moins égal au nombre'des fonctions. Alors le théorème 30 (ou 33 quinto) nous dit que l'image par Φ de tout voisinage de a est un voisinage de a . Alors $\Phi_1(u), \Phi_2(u), \dots, \Phi_N(u)$, peuvent prendre des valeurs arbitraires, pourvu qu'elles soient assez voisines de $a_1, a_2, \dots, a_N$.

Si donc il existe entre les Φ_i une relation de la forme

(III,9;31) $$R\left(\Phi_1(u_1,u_2,\ldots,u_n),\ \Phi_2(u_1,u_2,\ldots,u_n),\ldots,\Phi_N(u_1,u_2,\ldots,u_n)\right) \equiv 0\ ,$$

où R est une fonction de N variables, R doit être identiquement nulle au voisinage de $(a_1, a_2,\ldots,a_N)$, et nous n'avons pas là une 'vraie relation'".

<u>Il sera naturel de dire que les N fonctions Φ_i sont indépendantes au voisinage de α .</u>

<u>2ème Cas</u> - <u>Dépendance</u> -

Supposons au contraire que $\Phi'(\alpha)$ soit, pour tout α du même rang $\ell < N$ Alors le théorème 33 quinto nous dit qu'il existe un voisinage de α, dont l'image par Φ est une variété V, de dimension ℓ et de classe C^1, de $\mathbb{K}^N$ En changeant au besoin l'ordre des fonctions Φ_i données: on peut se ramener au cas où cette variété a des équations de la forme

(III,9;32) $$x_{\ell+k} = G_k(x_1, x_2,\ldots,x_\ell)\ , \qquad k = 1,2,\ldots, N-n.$$

Alors, au voisinage de α , les fonctions Φ_i satisfont aux $N-\ell$ relations non triviales

(III,9;33) $$\begin{cases} \Phi_{\ell+k}(u_1, u_2,\ldots,u_n) \equiv \\ G_k\left(\Phi_1(u_1,u_2,\ldots,u_n),\ \Phi_2(u_1,u_2,\ldots,u_n),\ldots,\ \Phi_\ell(u_1,u_2,\ldots,u_n)\right), \\ k = 1,2,\ldots,\ N-\ell. \end{cases}$$

Ces relations sont bien cette fois des vraies relations, puisque, quand les valeurs de $\Phi_1, \Phi_2,\ldots,\Phi_\ell$ sont connues, elles **déterminent** celles des $\Phi_{\ell+k}$. Ce sont $N-\ell$ <u>relations **indépendantes**</u>, en ce sens que les fonctions

$x \to x_{\ell+k} - G_k(x_1,x_2,\ldots,x_\ell)$, sont Indépendantes <u>au sens du **1er cas**.</u>

En effet, leurs différentielles sont les $dx_{\ell+k} - dG_k$, la $k$ième est seule à contenir $dx_{\ell+k}$ avec un coefficient $\neq 0$, donc il n'y a pas, pour x fixé, de relation linéaire à coefficients constants entre elles, le rang du système de ces $N-\ell$ différentielles est $N-\ell$, pour tout x voisin de a .

Il sera donc naturel de dire, dans ce cas, aue les fonctions $\Phi_1, \Phi_2, \ldots, \Phi_N$, sont dépendantes, et satisfont à $N - \ell$ relations indépendantes.

Il existe naturellement d'autres cas que nous n'avons pas traités : les cas où le rang de la dérivée n'est pas constant au voisinage du point α. Dans ces cas les conclusions sont bien plus compliquées que celles que nous venons d'écrire; c'est pourquoi, lorsque l'on dit, d'une façon assez courante, que N fonctions $\Phi_1, \Phi_2, \ldots, \Phi_N$, de n variables $u_1, u_2, \ldots, u_n$, à dérivées partielles du premier ordre continues, sont indépendantes, si et seulement si la matrice dérivée des Φ_i est de rang N, et sont dépendantes, si et seulement si cette matrice dérivée est de rang $< N$ on exprime quelque chose d'assez vague, et qui finalement n'est pas vrai.

Considérons par exemple, la représentation paramétrique (III,9;30) du cône $x^2 + y^2 - z^2 = 0$.

Les 3 fonctions écrites F, G, H de deux variables réelles r, φ satisfont bien à une relation non triviale, à savoir la relation $F^2 + G^2 - H^2 = 0$; mais, au voisinage d'un point tel que $(0, \varphi_0)$, représentant l'origine $x = y = z = 0$, ce n'est pas une relation dans laquelle une de ces 3 fonctions puisse être calculée comme une fonction continuement dérivable des deux autres, comme cela s'est produit dans le 2ème cas.

On peut aboutir à des figures bien plus compliqudes qu'un cône, et telles qu'on ne puisse plus donner aucun sens "utilisable" à la notion de dépendance ou d'indépendance. Il est sage de se borner aux deux cas précis que nous venons d'étudier.

Variétés singulières ou paramétriques

La nouvelle notion introduite ici généralise celle des chemins, introduits au chapitre II page 90.

On appelle variété singulière ou paramétrique, de dimension n et de classe C^m, dans un espace affine E de dimension N, une application6 de classe C^m, d'une variété V (abstraite ou plongée dans un espace affine) de dimension n et de classe C^m, dans E. L'image $\Phi(V) = W$ s'appelle l'image de la variété paramétrique; elle ne doit pas être confondue avec la variété paramétrique elle-même, qui est l'application Φ de V dans W. Par exemple, si Φ est constante, $\Phi(V) = \{a\}, a \in E$, l'image est un point, mais la variété paramétrique est l'application constante Φ. La lemniscate de Bernoulli dans $\mathbb{R}^2$

n'est pas une variété. Mais elle est l'image d'une variété paramétrique, où $V = \mathbb{R}$ (voir formule (III,9;29)); dans ce cas, Φ est injective; elle est aussi l'image d'une variété paramétrique où V est une circonférence, de telle manière que, si V est parcourue dans un sens déterminé, la lemniscate soit parcourue dans le sens 1, 2, 3, 4 de la figure 1; elle est encore l'image d'une variété paramétrique, avec toujours pour V une circonférence, mais de manière cettefois que, quand V est parcourue dans un sens déterminé, elle soit parcourue dans le sens 1, 2, 3, 4 de la figure 2; enfin elle est l'image d'une variété paramétrique où V est un système de 2 circonférences séparées, dont les images par Φ sont les 2 boucles de la lemniscate; autant de variétés paramétriques différentes .

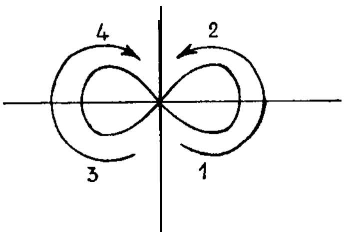


Fig. 1.

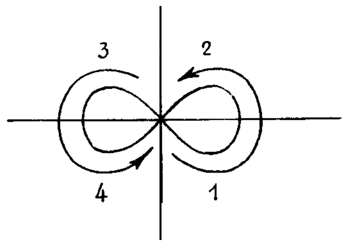


Fig. 2.

Si, au lieu de V , on prend un segment fermé $[a,b]$ de $\mathbb{R}$ une application Φ de classe C^m de $[a,b]$ dans E s'appelle' un <u>arc de courbe de classe C^m de E</u>, d'origine $\Phi(a)$ et d'extrêmité $\Phi(b)$. Si $\Phi(a) = \Phi(b)$, on dit que c'est un arc de courbe <u>fermé</u> (bien que cela n'ait aucun rapport avec la notion topologique d'ensemble fermé).

On dit que <u>2 variétés paramétriques</u> $\Phi_1 : V_1 \to E$, $\Phi_2 : V_2 \to E$, <u>sont C^m- équivalentes</u>, s'il existe un C^m-difféomorphisme H de V_1 sur V_2, tel que $\Phi_1 = \Phi_2 \circ H$ (alors $\Phi_2 = \Phi_1 \circ H^{-1}$) .

Alors leurs images W sont les mêmes. Cela correspond à ce qu'on considérait, en Mathématiques Spéciales, comme un <u>changement de représentation paramétrique</u> d'une courbe ou d'une surface. Les différentes représentations paramétriques de la lemniscate signalées ci-dessus <u>ne donnent pas</u> des variétés paramétriques équivalentes.

§ 10 MAXIMA ET MINIMA LIÉS

Soit Ω un ouvert d'un espace affine normé E sur $\mathbb{R}$, et soient $f, g_1, g_2, \ldots, g_m$, $m+1$ fonctions réelles continuement dérivables définies sur Ω. Quels sont les extrema, c'est-à-dire les maxima et les minima relatifs de f sur le sous-ensemble A de Ω, défini par les équations :

(III,10;1) $$g_1(x) = 0 \;,\; g_2(x) = 0 \;,\; \ldots \;,\; g_m(x) = 0 \;?$$

On cherche donc, non pas les maxima et les minima relatifs de f sur Ω, mais seulement sur un sous'ensemble fermé, donc non ouvert, de Ω ; on a affaire à ce qu'on appelle un maximum ou un minimum lié.

<u>Théorème 34 - Soient $f, g_1, g_2, \ldots, g_m$, des fonctions réelles continuement dérivables sur un ouvert Ω d'un espace affine normé E. Soit a un point de Ω vérifiant les équations</u>

$g_1(a) = g_2(a) = \ldots = g_m(a) = 0$. <u>Si les formes linéaires dérivées $g'_1(a), g'_2(a), \ldots, g'_m(a) \in \vec{E}'$, sont Indépendantes, une condition nécessaire pour que a soit un maximum ou minimum relatif de f sur le sous-ensemble A de Ω défini par les équations $g_1(x) = g_2(x) = \cdots = g_m(x) = 0$, est qu'il existem constantes réelles $\lambda_1, \lambda_2, \ldots, \lambda_m$, telles que l'on ait la relation :</u>

(III,10;2) $$f'(a) = \lambda_1 g'_1(a) + \lambda_2 g'_2(a) + \cdots + \lambda_m g'_m(a) \;.$$

<u>Les λ_i sont appelés les multiplicateurs de Lagrange relatifs à l'extrémum a.</u>

Avant de démontrer ce théorème, indiquons sa signification géométrique. Dans Ω, les équations $g_i(x) = 0$ définissent, au voisinage de a, une variété différentiable. La variété linéaire tangente a les équations $g'_i(a)\cdot$ x-a $= 0$, $i = 1, 2, \ldots, m$; la relation (III,10;2) exprime que l'hyperplan tangent $f'(a) . \overrightarrow{x-a}$ à l'hypersurface $f(x) = f(a)$ au point a contient cette variété linéaire.

Démonstration - Donnons d'abord la démonstration lorsque E est de dimension finie. On peut alors en choisir un référentiel : $O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, et représenter un point x de **E** par le système de ses coordonnées $x_1, x_2, \ldots, x_n$.

Supposer qu'au point a, les différentielles dg_i sont des formes linéaires indépendantes, c'est supposer qu'il existe m des coordonnées, par exemple $x_1, x_2, \ldots, x_m$, telles que le déterminant jacobien : $\dfrac{D(g_1, g_2, \ldots, g_m)}{D(x_1, x_2, \ldots, x_m)}(a)$ soit $\neq 0$.

Cela exige évidemment $n \geq m$. Alors, au voisinage de ce point a , il est possible, d'après le théorème des fonctions implicites 28, de calculer $x_1, x_2, \dots, x_m$ en fonction de $g_1, g_2, \dots, g_m, x_{m+1}, \dots, x_n$. On peut porter les valeurs trouvées dans f , et on obtient alors une nouvelle fonction continuement dérivable $F(g_1, g_2, \dots, g_m, x_{m+1}, \dots, x_n)$. On doit alors écrire que le point a est un extrémum de la fonction F , lorsque l'on fait varier, au voisinage de a , les coordonnées $x_{m+1}, \dots, x_n$, en laissant au contraire constamment nulles $g_1, g_2, \dots, g_m$. Une condition nécessaire pour qu'il en soit ainsi est que les dérivées partielles de F par rapport à $x_{m+1}, \dots, x_n$ soient nulles au point a . Cla exprime exactement que la différentielle dF au point a s'exprime sous la forme :

(III,10;3) $$dF = \lambda_1 dg_1 + \lambda_2 dg_2 + \cdots + \lambda_m dg_m ,$$

où $\lambda_1, \lambda_2, \dots, \lambda_m$, sont des constantes réelles; mais comme la différentielle df , exprimée en fonction des différentielles initiales, $dx_1, dx_2, \dots, dx_n$, s'obtient exactement en remplaçant, dans la différentielle dF par rapport aux variables $dg_1, dg_2, \dots, dg_m, dx_{m+1}, \dots, dx_n$ les dg_i par leurs différentielles en fonction des variables $dx_1, dx_2, \dots, dx_n$, on voit que cette différentielle df vérifie la formule

(III,10;4) $$df = \lambda_1 dg_1 + \dots + \lambda_m dg_m ,$$

ce qui revient à (III,10;2) et le théorème est démontré.

Supposons maintenant E de dimension infinie. Puisque les $g_i'(a)$ sont supposées être Indépendantes, elles peuvent prendre des valeurs arbitraires données-a l'avance. Pour tout i , on peut donc trouver un vecteur X_i de $\vec{E}$ tel que $g_j'(a) . \vec{X}_i = \delta_{i,j} = 0$ pour $j \neq i$, 1 pour $j = i$.

Les vecteurs $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m$ ainsi construits sont évidemment indépendants, car si on a une relation $\sum_{i=1}^{m} \lambda_i \vec{X}_i = \vec{0}$, on a $g_j'(a) \cdot \sum_{i=1}^{m} \lambda_i \vec{X}_i = \lambda_j = 0$. Posons alors $A_i = f'(a) . \vec{X}_i$; la forme linéaire $f'(a) - \sum_{i=1}^{m} \lambda_i g_i'(a)$ est nulle sur $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m$, donc sur le sous-espace vectoriel qu'ils engendrent. Soit alors $\vec{X}$ un vecteur quelconque de E . Appelons+ $\vec{E}_0$ le sous espace de E engendré par $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m, \vec{X}$. Sur le sous-espace affine E_0 mené para parallèlement à $\vec{E}_0$, f doit avoir un extremum en a, sur l'ensemble $A_1 = A \cap E_0$ défini par les équations $g_i(x) = 0$. Mais E_0 est de dimension finie, m ou $m+1$. Donc il-existe des constantes $\mu_1, \mu_2, \ldots, \mu_m$, telles que $f'(a) - \sum_{i=1} \mu_i\, g_i'(a)$ soit nul sur $\vec{E}_0$. En écrivant qu'elle est nulle sur $\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_m$, on trouve qu'on a **nécessairement** $\mu_i = \lambda_i$. Cela prouve que $f'(a) - \sum_{i=1}^{m} \lambda_i\, g_i'(a)$ est aussi nulle sur $\vec{X}$, **vecteur** arbitraire de **$\vec{E}$**, **donc elle est nulle sur** $\vec{E}$, et le **théorème** est démontrd.

Il résulte de cette démonstration que <u>les multiplicateurs de Lagrange λ_i sont déterminés d'une manière unique, ce qui était évident a priori, puisque les $g_i'(a)$ sont supposées independantes.</u>

<u>Remarque - Soit g une fonction de classe **C'** sur $\Omega \subset E$ à valeurs dans un espace affine **G** de dimension **finie** m, telle que sa dérivee $g'(a) \in \mathcal{L}(\vec{E}\,;\vec{G})$ soit exactement de **rang** m. Pour que f ait en a un extremum, sur le sous-ensemble de Ω d'équation $g(x) = g(a)$, il est nécessaire qu'il existe un élément λ de $\vec{G}' = \mathcal{L}(\vec{G}\,;\mathbb{R})$ tel que l'on **ait** :</u>

(III,10;4bis) $$f'(a) = \lambda \circ g'(a) \in \vec{E}' = \mathcal{L}(\vec{E}\,;\mathbb{R}) \,.$$

Prenons en effet un référentiel de **G**, ayant $g(a)$ comme **origine**: G est ainsi Identifié **à** $\mathbb{R}^m$; alors les composantes de la fonction g sont-m fonctions scalaires g_i, $i = 1,2,\ldots,m$, et **A** est **défini** par les **équations** : $g_i(x) = 0$. **Dire que** $g'(a) \in \mathcal{L}(\vec{E}\,;\vec{G})$ **est de rang** m, c'est dire que les $g_i'(a) \in \mathcal{L}(\vec{E}\,;\mathbb{R})$ sont indépendantes; nous nous trouvons dans les **conditions** d'application du théorème. L'existence des multiplicateurs de Lagrange λ_i, $i = 1,2,\ldots,m$, **tels** que l'on ait (**III,10;2**), est alors équivalente à l'existence de **l'élément** λ de $\vec{G}'$ tel que l'on **ait (III,10; 4bis), mals**, sous la forme (III,10; **4bis**), le résultat est indépendant de tout référentiel de G Noter que, dans (**III,10;4 bis**), $g'(a)$ est un élément de' $\mathcal{L}(\vec{E}\,;\vec{G})$, λ un **élément** de $\mathcal{L}(\vec{G}\,;\mathbb{R})$, donc $\lambda \circ g'(a)$ est bien un élément de $\mathcal{L}(\vec{E}\,;\mathbb{R})$.

Manière pratique de procéder pour trouver un maximum ou un minimum relatif lié

Supposons, pour fixer les idées, que E soit de **dimension** n, que 1 on en ait choisi un **référentiel**, et que chaque point x de E soit donné par ses coordonnees $x_1, x_2, \ldots, x_n$. Cherchons un extremum **lié** a = $(a_1, a_2, \ldots, a_n)$. Il doit d'abord vérifier l'équation (**III,10;2**), qui est **équivalente** au système de n équations :

$$(\text{III},10;5)\qquad \frac{\partial f}{\partial x_j}(a) = \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j}(a) \quad ; \qquad j = 1,2,\ldots,n .$$

Si on leur ajoute le système de m équations (III,10;1) (où x est remplacé par a), on voit que l on a $m+n$ équations par rapport à m + n inconnues, à savoir les coordonnées a_j de a, et les multiplicateurs de Lagrange λ_i . Il n'est pas en fait indispensable de résoudre complètement ce système; on peut, si l'on veut, éliminer les multiplicateurs λ_i pour se ramener à un système de n équations par rapport aux n inconnues a_j .

Remarque - Bien que, dans la pratique, on écrive les équations précédentes, sans vérifier nécessairement qu'au point a considéré, les dérivées des g_i sont indépendantes, ceci est une condition essentielle à la validité du theorème.

Supposons par exemple que f soit une fonction réelle continuement dérivable de n variables réelles $x_1, x_2, \cdots, x_n$, et que g soit la fonction : $x_1^2 + x_2^2 + \ldots + x_{n-1}^2$.

Que peut-on dire d'un point a , où la fonction f soit maxima ou minima, sur le sous-ensemble A défini par l'équation $g(x) = 0$ dans $\mathbb{R}^n$? Cette équation est équivalente au système d'équations : $x_1 = x_2 . . . = x_{n-1} = 0$, et par conséquent le sous-ensemble A considéré n'est autre que l'axe des x_n. Alors, pour qu'une fonction f admette en un point a de l'axe des x_n , un maximum ou minimum relatif, il est seulement nécessaire que sa dérivée partielle $\partial f/\partial x_n$ soit nulle au point a , sans qu'aucune condition soit nécessaire relativement à ses autres dérivées partielles. Si, au contraire, on exprimait qu'il existe un multiplicateur de Lagrange λ tel que :

$$\frac{\partial f}{\partial x_i}(a) = \lambda \frac{\partial g}{\partial x_i}(a) \quad i = 1,2,\ldots n,$$

on voit que cette relation exprimerait que toutes les dérivées partielles de f au point a, sont nulles ! Mais nous aurions appliqué le théorème 34 alors qu'il n'est pas applicable, car, en un point a de l'axe des x_n , la différentielle de g est nulle.

Remarquons d'autre part, comme nous l'avons déjà remarqué, pour les maxima et minima non lies, que cette méthode ne nous permet pas d'affirmer que la fonction est maxima ou minima au point a puisqu'après les conditions du premier ordre, il faudrait examiner le développement de Taylor de f au point a ; remarquons aussi que cette méthode laisse complètement échapper les extrema liés d'une fonction définie sur une partie F de E qui ne serait pas ouverte, si ces points appartiennent à la frontière de F .

On pourra utiliser cette méthode pour déterminer les axes d'une quadrique dans un espace affine euclidien à n dimensions, en cherchant les points de cette quadrique dont la distance au centre est extrema. On pourra de même chercher les axes d'une section plane d'une quadrique. Plus généralement cherchons le point x d'une hvpersurface d'équation

(III,10;6) $$g(x_1, x_2, \ldots, x_n) = 0 \quad (g \text{ continue})$$

dans l'espace& , dont la distance euclidienne à un point donné $b = (b_1, b_2, \ldots, b_n)$ est minima. D'après ce qui a été vu page 82 la surface étant un ensemble fermé (comme image réciproque de 0 par l'application continue g de $\mathbb{R}^n$ dans $\mathbb{R}$), il existe nécessairement au moins un point de cette surface dont la distance au point b est minima. Nous devons chercher le point x réalisant le minimum de la quantité

(III,10;7) $$f(x) = \sum_{i=1}^{n} (x_i - b_i)^2 ,$$

parmi tous les points qui vérifient l'équation de la surface. Sir. répond à la question, et si en ce point la différentielle dg n'est pas nulle, c'est-à-dire si les n dérivées partielles $\partial g/\partial x_i$ ne sont pas simultanément nulles, il existe nécessairement un multiplicateur de Lagrange λ, tel que l'on ait les équations

(III,10;8) $$x_i - b_i = \lambda \frac{\partial g}{\partial x_i}(x) , \qquad i = 1, 2, \ldots, n .$$

En éliminant λ on trouvera le point x cherché en écrivant les n équations'& n inconnues :

(III,10;9) $$\frac{\frac{\partial g}{\partial x_1}(x)}{x_1 - b_1} = \frac{\frac{\partial g}{\partial x_2}(x)}{x_2 - b_2} = \cdots = \frac{\frac{\partial g}{\partial x_n}(x)}{x_n - b_n} ; \qquad g(x) = 0 .$$

Ces équations expriment tout simplement que le point cherché, si ce n est pas un point singulier, est le pied d'une normale issue du point b à l'hypersurface C'est conforme à l'interprétation géométrique donnée aprìs l'énoncé du théorème 34 : au point x cherché l'hypersurface est tangente à la sphère de centre b qui passe en ce point. Le problème étant ainsi résolu, on devra cependant remarquer ce qui suit :

1°/ Tout point de la surface qui est le pied d'une normale issue du point b ne donne pas nécessairement naissance à un minimum de la distance, il peut donner lieu à un minimum relatif ou à un maximum relatif, ou à un col.

2°/ Il peut arriver qu'aucun des pieds des normales ne donne effectivement le minimum absolu de la distance, et que celui-ci soit réalisé par un point singulier de la surface, c'est-a-dire par un point où les n dérivées partielles $\frac{\partial g}{\partial x_i}$ sont simultanément nulles, et où par conséquent le procédé du multiplicateur n'est plus valable. Remarquons d'ailleurs que, lorsqu'on a éliminé λ pour trouver (III,10;9), on a introduit ces solutions nouvelles. Si la surface n'a aucun point singulier, comme nous avons vu que le minimum existait nécessairement, on en déduira que, par tout point b , on peut mener à la surface au moins une normale. Si, en outre, la surface est compacte, la distance d'un point de la surface au point b , fonction continue sur un compact, admet un maximum, et dans ce cas, la surface étant toujours supposée sans point singulier, on peut certainement du point b mener au moins une deuxième normale, dont le pied donnera le maximum de la distance d'un point de la surface au point b .

Applications de la théorie des maxima liés; inégalités de Hölder et de Minkowski

Soient $c_1, c_2, \dots, c_n$, n nombres > 0 donnés, fixés une fois pour toutes. Soit d'autre part p un nombre > 0 fini. Pour tout point $\vec{X} = (X_1, X_2, \dots, X_n)$ de $\mathbb{C}^n$, on pose :

$$\|\vec{X}\|_p = (c_1 |X_1|^p + c_2 |X_2|^p + \dots + c_n |X_n|^p)^{\frac{1}{p}} \tag{III,10;10}$$

et on appelle cette quantité la norme d'ordre p de $\vec{X}$ relative aux coefficients $c_i > 0$. C'est manifestement un nombre ≥ 0 . On a d'autre part, trivialement $\|\lambda\vec{X}\| = |\lambda| \, \|\vec{X}\|$. Enfin $\|\vec{X}\| = 0$ est équivalent à $\vec{X} = \vec{0}$ Pour justifier le nom de norme, il reste à démontrer l'inégalité de convexité.

Comme nous le verrons plus loin, cette inégalité est seulement vérifiée pour $p \geq 1$; c'est donc seulement dans ce cas que le nom de norme est justifié. Néanmoins, nous continuerons à l'appeler norme d'ordre p , et à la noter de la même manière, même si p est < 1 !

Bien entendu, l'application $(\vec{X}, p) \Rightarrow \|\vec{X}\|_p$ est continue de $\mathbb{C}^n \times \mathbb{R}_+^*$ dans $\mathbb{R}$, où $\mathbb{R}_+^*$ est l'ensemble des nombres réels > 0 . Cherchons ce que devient cette norme lorsque p tend vers $+\infty$. Posons $M = \operatorname{Sup}(|X_1|, |X_2|, \dots |X_n|)$, supposons pour fixer les idées, que i soit un indice pour lequel $|X_i| = M$. On a alors les inégalités :

$$\|\vec{X}\|_p \leq \left(\sum_{i=1}^{n} c_i\right)^{\frac{1}{p}} M, \text{ et } \|\vec{X}\|_p \geq c_i^{\frac{1}{p}} M . \tag{III,10;11}$$

Lorsque? tend vers $+\infty$, les deuxièmes membres de ces inégalités tendent tous les deux vers M ; on en déduit que, si $\vec{X}$ est fixe, et que p tend vers $+\infty$, $\|\vec{X}\|_p$ tend vers M .

C'est pourquoi on pose habituellement :

(III,10;12) $$\|\vec{X}\|_\infty = \operatorname{Sup}\left(|X_1|, |X_2|, \ldots, |X_n|\right).$$

On remarquera alors que les trois normes, considérées sur $\mathbb{C}^n$ à la page 40 , ne sont autres que les normes $\|\vec{X}\|_\infty$, $\|\vec{X}\|_1$, $\|\vec{X}\|_2$. Voyons maintenant ce que devient $\|\vec{X}\|_p$ lorsque p tend vers 0 . La chose n'a pas d'intérêt dans le cas général, elle n'est interessante que dans le cas particulier où la somme $\sum_{i=1}^{n} c_i$ vaut 1. Dans ce cas la norme s'appelle aussi la moyenne d'ordre p des nombres $|X_1|, |X_2|, \ldots, |X_n|$, par rapport aux coefficients ci.

Il s'agit là d'une moyenne pondérée.

La moyenne arithmétique habituelle correspond aux c_i égaux à $\frac{1}{n}$ et à $p = 1$; pour des c_i égaux à $\frac{1}{n}$, et p quelconque, la moyenne d'ordre p est le nombre dont la puissance p-ième est la moyenne arithmétique des puissances p-ièmes des $|X_i|$.

Posons donc $\sum_{i=1}^{n} c_i = 1$, et cherchons la limite de $\|\vec{X}\|_p$ lorsque p tend vers 0 . Supposons d'abord qu'aucun des nombres X_i ne soient nuls. On a alors le développement limité:

(III,10;13) $$|X_i|^p = e^{p \log|X_i|} \doteq 1 + p \log|X_i| + \cdots$$

on en déduit :

(III,10;14) $$\sum_{i=1}^{n} c_i |X_i|^p = \sum_{i=1}^{n} c_i + p \sum_{i=1}^{n} c_i \log|X_i| + \cdots = 1 + p \sum_{i=1}^{n} c_i \log|X_i| + \cdots$$

et, compte tenu de ce que $\sum_{i=1}^{n} c_i = 1$, on en déduit le développement :

(III,10;15) $$\log\left(\sum_{i=1}^{n} c_i |X_i|^p\right) = p \sum_{i=1}^{n} c_i \log|X_i| + \cdots .$$

On voit par conséquent que, lorsque p tend vers 0, $\log\|\vec{X}\|_p = \frac{1}{p}\log\left(\sum_{i=1}^{n} c_i |X_i|^p\right)$ tend vers $\sum_{i=1}^{n} c_i \log|X_i|$, t par conséquent $\|\vec{X}\|_p$ tend vers :

(III,10;16) $$\prod_{i=1}^{n} |X_i|^{c_i} .$$

Cette quantité n'est autre que la moyenne géométrique pondérée des nombres $|X_i|$ par rapport aux poids c_i .

Si tous les c_i sont égaux à $\frac{1}{n}$, c'est la moyenne géométrique habituelle $\sqrt[n]{|X_1 X_2 \dots X_n|}$. On vérifie aisément que le résultat reste valable si certains des X_i sont nuls, la limite de $\|\vec X\|_p$ étant alors nulle.

On est donc amené à poser, lorsque $\sum_{i=1}^{n} c_i = 1$

(III,10;17) $$\|\vec X\|_0 = \prod_{i=1}^{n} |X_i|^{c_i}$$

Théorème 35 (Inégalité de Minkowski)

Pour $p \geq 1$ * on a l'inégalité de convexité :

(III,10;18) $$\|\vec X + \vec Y\|_p \leqslant \|\vec X\|_p + \|\vec Y\|_p ;$$

en outre, pour $1 < p < \infty$, on a l'inégalité stricte $<$, sauf si les vecteurs $\vec X$ et $\vec Y$ sont proportionnels, avec un rapport de proportionnalité ≥ 0.

Démonstration - L'inégalité est évidente pour $p = 1$ et $p = +\infty$.

Il reste donc à étudier le cas p fini > 1, et il suffit naturellement de démontrer l'inégalité lorsque les vecteurs $\vec X$ et $\vec Y$ sont tous les deux $\neq 0$. Elle sera évidemment montrée, si on la montre pour des X_i, Y_i, réels ≥ 0.

* Pour $p < 1$, le résultat ne subsiste pas. Prenons par exemple $p = \frac{1}{2}$, $n = 2$, $c_1 = c_2 = 1$, $X_1, X_2, Y_1, Y_2 > 0$.

Alors (III,10;8) signifierait :

$$\left(\sqrt{X_1+Y_1} + \sqrt{X_2+Y_2}\right)^2 \leqslant \left(\sqrt{X_1} + \sqrt{X_2}\right)^2 + \left(\sqrt{Y_1} + \sqrt{Y_2}\right)^2 \quad \text{, ou}$$

$$\sqrt{(X_1+Y_1)(X_2+Y_2)} \leqslant \sqrt{X_1 X_2} + \sqrt{Y_1 Y_2} \quad \text{ou}$$

$$(X_1+Y_1)(X_2+Y_2) \leqslant X_1X_2 + Y_1 Y_2 + 2\sqrt{X_1X_2Y_1Y_2} \quad \text{ou}$$

$$X_1 Y_2 + X_2 Y_1 \leqslant 2\sqrt{X_1 X_2 Y_1 Y_2}$$, alors qu'on a juste l'inégalité $\geqslant$, en vertu de $a + b \geq 2\sqrt{ab}$

Mais l'ensemble des points de coordonnées 30 de $\mathbb{R}^n$ n'est pas un ouvert dans un espace vectoriel normé, et par con équent les théorèmes sur les extréma ne seraient pas applicables. C'est pourquoi nnus allons démontrer l'inégalité sur l'espace vectoriel $\mathbb{R}^n$ entier.

Posons $\|\vec{X}\|_p = \alpha$, $\|\vec{Y}\|_p = \beta$, et $\|\vec{X}+\vec{Y}\|_p = \gamma$.

Considérons l'ensemble A des points $(\vec{X},\vec{Y})$ de $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$, pour lesquels a et β prennent des valeurs données α_0, β_0, distinctes de 0 . C'est manifestement une <u>partie compacte</u> de $\mathbb{R}^{2n}$. En effet, elle est fermée, comme image réciproque du point (α_0 , β_0) de $\mathbb{R} \times \mathbb{R}$, par l'application continue $(\vec{X},\vec{Y}) \Rightarrow (\|\vec{X}\|_p , \|\vec{Y}\|_p)$ de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R} \times \mathbb{R}$; et elle

est bornée, car, si α_0 et β_0 sont donnés-, cela-donne une majoration de toutes les coordonnées de Xet de Y ; elle est donc compacte d'après le théorème 23 du chapitre II. Comme alors γ est une fonction continue du point $(\vec{X},\vec{Y})$ de $\mathbb{R}^n \times \mathbb{R}^n$, elle admet nécessairement sur le compact A un maximum, et un minimum. Appelons $(\vec{X},\vec{Y})$ un point où ce maximum ou ce minimum est atteint. C'est un extremum lié. C'est en effet un point où la fonction γ^p , continue sur $\mathbb{R}^n \times \mathbb{R}^n$, atteint un maximum ou un minimum, parmi les points liés par les deux relations $\alpha^p(\vec{X},\vec{Y}) = \alpha^p(\vec{X}) = \alpha_0^p$, $\beta^p(\vec{X},\vec{Y}) = \beta^p(\vec{Y}) = \beta_0^p$.

<u>Or les trois fonctions α^p, β^p, γ^p sont continuement dérivables.</u>

Considérons par exemple la fonction α^p . Elle admet une dérivée partielle par rapport à X_i , dans l'ouvert $X_i \neq 0$, qui est :

$$\frac{\partial}{\partial X_i}(\alpha^p) = p\, c_i\, \varepsilon_i\, |X_i|^{p-1} ,\ \varepsilon_i = \text{signe de } X_i \ . \tag{III,10;19}$$

Mais cette dérivée posséde une limite qui est nulle lorsque X_i tend vers 0 ; donc, d'après le théorème 14 appliqué à la fonction α^p lorsque X_i seule varie, il existe aussi une dérivée partielle sur l'hyperplan $X_i = 0$, et cette dérivée est nulle * . La dérivée de α^p, par

* Un examen direct montre aussi immédiatement qu'en un point où $X_i = 0$ la dérivée $\frac{\partial}{\partial X_i}$ est nulle !

rapport à X_i, existe donc partout; la fonction dérivée est définie par (III,10;19). Cette fonction est continue dans tout $\mathbb{R}^n \times \mathbb{R}^n$, et par conséquent α^p qui admet des dérivées partielles du premier ordre continues, est continuement dérivable d'après le théorème 15. Il en est de même de β^p et γ^p, dont les dérivées sont :

$$\text{(III,10;20)} \quad \frac{\partial}{\partial Y_i}(\beta^p) = p\, c_i\, \eta_i\, |Y_i|^{p-1}, \qquad \eta_i = \textit{signe de } Y_i .$$

$$\text{(III,10;21)} \quad \frac{\partial}{\partial X_i}(\gamma^p) = \frac{\partial}{\partial Y_i}(\gamma^p) = p\, c_i\, \zeta_i\, |X_i + Y_i|^{p-1}, \quad \zeta_i = \text{signe de } X_i + Y_i .$$

(On voit que les résultats précédents n'auraient pas été exacts pour $p \leq 1$). Alors, au point $(\vec{X}, \vec{Y})$ où est réalisé, soit le maximum, soit le minimum, il existe deux multiplicateurs de Lagrange λ et μ, tels que l'on ait * :

$$\text{(III,10;22)} \qquad d(\gamma^p) = \lambda\, d(\alpha^p) + \mu\, d(\beta^p) , \qquad \text{ou}$$

$$\text{(III,10;23)} \quad \begin{cases} p\, c_i\, \zeta_i\, |X_i + Y_i|^{p-1} = \lambda\, p\, c_i\, \varepsilon_i\, |X_i|^{p-1}, & i = 1,2,\ldots,n\ ; \\ p\, c_i\, \zeta_i\, |X_i + Y_i|^{p-1} = \mu\, p\, c_i\, \eta_i\, |Y_i|^{p-1}, & i = 1,2,\ldots,n . \end{cases}$$

Si, pour un indice i X_i est nul, la première équation (III,10;23) pour l'indice i donne $X_i + Y_i = 0$, donc Y_i est nul aussi. De même, si Y_i est nul la seconde équation donne $X_i = 0$. Donc, au point où est réalisé le maximum ou le minimum, l'une des deux coordonnées X_i, Y_i, ne peut être nulle sans que l'autre le soit aussi.

* A condition d'être sûr que $d(\alpha^p)$ et $d(\beta^p)$ sont indépendantes. Mais l'une ne dépend que des dX_i, l'autre des dY_i, elles ne peuvent être dépendantes que si l'une d'elles est nulle c'est-a-dire si $\vec{X} = \vec{0}$ ou $\vec{Y} = \vec{0}$, ce que nous avons exclu en prenant $\alpha_0 \neq 0$, $\beta_0 \neq 0$.

SI aucune n'est nulle, les équations (III,10;23) pour l'indice i sont équivalentes à

$$\zeta_i \left| 1 + \frac{Y_i}{X_i} \right|^{p-1} = \varepsilon_i \lambda \,, \quad \zeta_i \left| 1 + \frac{X_i}{Y_i} \right|^{p-1} = \eta_i \mu \,. \tag{III,10;24}$$

De toute façon l'ensemble des équations, que les coordonnées soient nulles ou non, peut s'écrire sous la forme équivalente, après élimination de λ et μ en prenant les modules des quantités écrites :

$$\frac{Y_1}{X_1} = \frac{Y_2}{X_2} = \cdots = \frac{Y_n}{X_n} \,, \tag{III,10;25}$$

certains des rapports étant éventuellement indéterminés parce qu'égaux à $\frac{0}{0}$ (mais pas tous, puisque $\vec{X}$ et $\vec{Y}$ ne sont pas nuls). Soit s la valeur commune de tous ces rapports. On a alors nécessairement : $\beta_0 = |s|\,\alpha_0$, et alors :

$$\gamma = \left(\sum_{i=1}^{n} c_i \, |1+s|^p \, |X_i|^p \right)^{\frac{1}{p}} = |1+s|\,\alpha_0 \,. \tag{III,10;26}$$

Cela montre nécessairement que le maximum de γ est $\alpha_0+\beta_0$, atteint si $s > 0$ et que son minimum est $|\alpha_0-\beta_0|$, atteint si $s<0$. Il en résulte que l'on a bien toujours l'inégalité $\gamma \leq \alpha + \beta$, ce qui est l'inégalité de convexité (III,10;18). On a toujours l'inégalité stricte $\gamma < \alpha + \beta$, sauf aux points précis où est réalisé le maximum. Or, en un quelconque de ces points, on a (III,10;25), la valeur commune des rapports étant $s > 0$; autrement dit les vecteurs $\vec{X}$ et $\vec{y}$ sont proportionnels avec un coefficient de proportionnalité $\geqslant 0$ * . Ceci, il est vrai, ne démontre la deuxième partie de l'énoncé du théorème que si les vecteurs $\vec{x}$ et $\vec{y}$ ont des coordonnées réelles; mais, même dans le cas général, s'ils ont des coordonndes complexes, on a la suite d'inégalités :

$$\|\vec{X}+\vec{Y}\|_p \leqslant \left(\sum_{i=1}^{n} c_i \left(|X_i| + |Y_i| \right)^p \right)^{\frac{1}{p}} \leqslant \|\vec{X}\|_p + \|\vec{Y}\|_p \,, \tag{III,10;27}$$

la première étant une inégalité stricte, sauf si tous les rapports Y_i / X_i sont réels $\geqslant 0$, et ceci démontre le théorème dans le cas général.

<u>Théorème 36 (Inégalités de Hölder)</u>

<u>Pour p et q quelconques, et r tel que $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, si l'on convient de noter par $\vec{X}\,\vec{Y}$ le point de coordonnées</u> $X_1 Y_1$,

* Nous disons $\geqslant 0$ et non > 0 , pour récupérer le cas évident, laissé de côté au début, où l'un au moins des 2 vecteurs $\vec{X}$, $\vec{Y}$, est nul.

$X_2 Y_2, \ldots, X_n Y_n$ on a

(III,10;28) $$\|\vec{X}\vec{Y}\|_r \leqslant \|\vec{X}\|_p \|\vec{Y}\|_q \quad ;$$

de plus, si $0 < p < +\infty$, $0 < q < +\infty$, on a toujours l'inégalité stricte $<$, sauf si

(III,10;29) $$\frac{|Y_1|^q}{|X_1|^p} = \frac{|Y_2|^q}{|X_2|^p} = \cdots = \frac{|Y_n|^q}{|X_n|^p}$$

Démontrons d'abord le théorème, lorsque les 3 nombres p, q, r sont > 1 et finis. Nous emploierons, comme pour le théorème 35, la méthode des multiplicateurs de Lagrange. Nous poserons

$$\alpha = \|\vec{X}\|_p, \quad \beta = \|\vec{Y}\|_q, \quad \gamma = \|\vec{X}\vec{Y}\|_r$$

Si nous nous plaçons au point où est réalisé soit un maximum, soit un minimum de γ, lorsque a et β prennent des valeurs données α_0 et $\beta_0 \neq 0$, il existe deux multiplicateurs λ et μ tels que $d(\gamma^r) = \lambda\, d(\alpha^p) + \mu\, d(\beta^q)$, ou encore

(III,10;30) $$\begin{cases} r\, c_i\, \varepsilon_i\, |X_i|^{r-1} |Y_i|^r = \lambda\, p\, c_i \varepsilon_i\, |X_i|^{p-1}, & i = 1, 2, \ldots n\,; \\ r\, c_i\, \eta_i\, |X_i|^r\, |Y_i|^{r-1} = \mu\, q c_i \eta_i |Y_i|^{q-1}, & i = 1, 2, \ldots n\,. \end{cases}$$

En fait la recherche du minimum est immédiate. Si en effet, pour tout i, l'une au moins des deux coordonnées X_i, Y_i est nulle, mais de manière que l'on ait à la fois $d = \alpha_0$ et $\beta = \beta_0$ (par exemple si $X_1 = \frac{\alpha_0}{c_1^{\frac{1}{p}}}$ et $X_2 = X_3 = \cdots X_n = 0$, puis $Y_1 = 0$, $Y_2 = \frac{\beta_0}{c_2^{\frac{1}{q}}}$, $Y_3 = \cdots Y_n = 0$) alors $\vec{X}\,Y$ est nul, et le minimum de γ^r est nul. Nous écarterons donc ce cas, et 'chercherons le maximum seulement.

Si, pour un indice i, X_i est nulle, la deuxième équation pour l'indice i donne $Y_i = 0$ ou $\mu = 0$ • mais, si $\mu = 0$, alors, quel que soit i, l'une au moins des deux coordonnées X_i Y_i est nulle, et nous nous trouvons dans le cas du minimum, écarté par hypothèse. On a donc $\mu \neq 0$, et par suite $X_i = 0$ entraîne $Y_i = 0$. De même $Y_i = 0$ entraîne $X_i = 0$.

Si par ailleurs X_i et Y_i sont toutes les deux $\neq 0$, alors les deux équations (III,10;30) donnent la relation :

(III,10;31) $$p\lambda |X_i|^p = q\mu |Y_i|^q, \quad i = 1,2,\ldots n, \quad \text{donc (III,10;29),}$$

certains de ces rapports, mais pas tous, pouvant se présenter sous la forme indéterminée $\frac{0}{0}$

Posons : $|Y_i|^q = s\ |X_i|^p$; alors $\beta_0^q = s\alpha_0^p$.

On en déduit

(III,10;32) $$\gamma^r = s^{\frac{r}{q}} \sum_{i=1}^{n} c_i |X_i|^{r+\frac{pr}{q}} = s^{\frac{r}{q}} \sum_{i=1}^{n} c_i |X_i|^p = s^{\frac{r}{q}} \alpha_0^p$$

ce qui donne $\gamma^r = \alpha_0^r \beta_0^r$ ou $\gamma = \alpha_0 \beta_0$. On en déduit que, pour tout point $(\vec{X}, \vec{Y})$ de $\mathbb{R}^n \times \mathbb{R}^n$, on a l'inégalité $\gamma \leqslant \alpha_0 \beta_0$, ce qui est l'inégalité (III,10;28).

De plus on a toujours l'inégalité stricte $\gamma < \alpha_0 \beta_0$, sauf aux points précis où est réalisé le maximum; et en ces points on a nécessairement les égalités de proportionnalité (III, 10;29).

Le théorème n'a été démontré que si p, q, r sont > 1 et finis. Mais, si nous supposons le théorème démontré pour des indices p, q, et r déterminé par $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, alors, si nous posons $|X_i| = |\xi_i|^k$, $|Y_i| = |\eta_i|^k$, $0 < k < +\infty$, on a l'inégalité

(III,10;33) $$\|\vec{\xi}\,\vec{\eta}\|_{kr}^k \leqslant \|\vec{\xi}\|_{kp}^k \ \|\vec{\eta}\|_{kq}^k ,$$

ce qui montre, en élevant les 2 membres à la puissance $\frac{1}{k}$, que le théorème est encore démontré pour les exposants kp , kq, kr avec $\frac{1}{kp} + \frac{1}{kq} = \frac{1}{kr}$.

Il en résulte que, pour démontrer le théorème pour des exposants p,q quelconques, >0 finis, on pourra se ramener au cas déjà démontré des exposants $\frac{p}{k}$, $\frac{q}{k}$, en choisissant k de manière que les trois nombres $\frac{p}{k}$, $\frac{q}{k}$, $\frac{r}{k}$, soient > 1 .

Ceci toutefois ne démontre pas encore le théorème si certains des exposants p, q, r sont égaux à $+\infty$, ou à 0 si $\sum_{i=1}^{n} c_i = 1$.

Si par exemple $p = +\infty$, on a nécessairement $q = r$, et l'inégalité (III,10;28) s'écrit sous la forme :

(III,10;34) $$\|\vec{X}\,\vec{Y}\|_q \leqslant \|\vec{X}\|_\infty \ \|\vec{Y}\|_q ,$$

qui est triviale, même si q est aussi égal à $+\infty$.

De toute façon les inégalités dans ces différents cas peuvent toujours se démontrer par passage à la limite.

Corollaire 1 - On dit que deux exposants $p, p' \geq 1$, sont conjugués, si l'on a la relation :

(III,10;35) $$\frac{1}{p} + \frac{1}{p'} = 1 \text{ , ou } \quad p' = \frac{p}{p-1}$$

alors on a le cas particulier de l'inégalité de Hölder :

(III,10;36) $$\left|\sum_{i=1}^{n} c_i X_i Y_i\right| \leq \|\vec{X}\|_p \; \|\vec{Y}\|_{p'} \; ;$$

en outre, pour $p \neq 1$ et $\neq \infty$, on a l'inégalité stricte <, sauf si l'on a la relation de proportionnalité :

(III,10;37) $$\frac{|Y_1|^{p'}}{|X_1|^p} = \frac{|Y_2|^{p'}}{|X_2|^p} = \cdots = \frac{|Y_n|^{p'}}{|X_n|^p} \; ,$$

tous les produits $X_i Y_i$ ayant le même argument. *

Il suffit d'appliquer l'inégalité générale dans laquelle $r = 1$, elle donne alors la relation : $\|\vec{X}\vec{Y}\|_1 \leq \|\vec{X}\|_p \|\vec{Y}\|_{p'}$, et par conséquent a fortiori la relation (III,10;36). Mais

(III,10;38) $$\left|\sum_{i=1}^{n} c_i X_i Y_i\right| < \|\vec{X}\vec{Y}\|_1$$

sauf si tous les nombres complexes $X_i Y_i$ ont le même argument, on voit bien que, sauf dans le cas indiqué dans l'énoncé, (III,10;36) est bien une inégalité stricte.

Corollaire 1 bis - (Inégalité de Cauchy-Schwarz)

On a l'inégalité :

(III,10;39) $$\left|\sum_{i=1}^{n} c_i X_i \bar{Y}_i\right| \leq \|\vec{X}\|_2 \; \|\vec{Y}\|_2 \; ,$$

et l'inégalité stricte <, sauf si les vecteurs $\vec{X}$ et $\vec{Y}$ sont proportionnels.

Il suffit d'appliquer le corollaire 1 aux exposants conjugués $p = p' = 2$; l'argument de $X_i \bar{Y}_i$ est celui de $\frac{X_i}{Y_i}$, et alors tous les $\frac{Y_i}{X_i}$ ont même module et même argument, si et seulement si $\vec{X}$ et $\vec{Y}$ sont proportionnels. Nous avons donné une autre démonstration (formule (III,1;11)) ; il suffit, pour retrouver (III,10;39), d'appliquer (III,1;11) à la forme sesquilinéaire hermitienne définie positive :

(III,10;40) $$B(\vec{X},\vec{Y}) = \sum_{i=1}^{n} c_i X_i \bar{Y}_i \; .$$

* Si tous les X_i ne sont pas nuls, (III,10;37) exprime qu'il existe une constante k telle que $|Y_i| = k|X_i|^{p-1}$

Corollaire 2 - SI la somme des poids $\sum_{i=1}^{n} c_i$ est 1, alors, pour $\vec{x}$ fixé, la moyenne d'ordre $\not p$: $\|\vec{x}\|_{\not p}$ est une fonction continue croissante de $\not p$; elle est strictement croissante, sauf si $|X_1| = |X_2| = \cdots = |X_n|$, auquel cas c'est une constante.

En efr'et, il suffit d'appliquer l'inégalité de Hölder au cas où $Y_1 = Y_2 = \ldots = Y_n = 1$; compte tenu de ce que $\sum_{i=1} c_i = 1$, on a alors $\|\vec{Y}\|_q = 1$; cela donne, quel que soit $q \geq 0$ et par conséquent quel que soit $r \leq \not p$, l'inégalité :

$$\|\vec{X}\|_r \leq \|\vec{X}\|_{\not p} \quad ; \tag{III,10;41}$$

en outre il s'agit d'une inégalité stricte, sauf si l'on a :

$$\frac{1}{|X_1|^{\not p}} = \frac{1}{|X_2|^{\not p}} = \cdots = \frac{1}{|X_n|^{\not p}} , \tag{III,10;42}$$

c'est-à-dire si tous les $|X_i|$ sont égaux, auquel cas manifestement la moyenne est une constante égale à la valeur commune de ces nombres.

Corollaire 2 bis - La moyenne géométrique d'un nombre fini de nombres $\geqslant 0$ est $\leqslant$ a leur moyenne arithmétique, et $<$, sauf si tous ces nombres sont egaux.

Il suffit en effet d'appliquer (III,10;41) à $r=0$, $\not p = 1$.

Démonstration des Inégalités de Hölder et de Minkowski par la théorie des fonctions convexes.

La fonction $x \to x^s$, $s \geq 1$, est convexe pour $x \geq 0$; elle est en effet continue, et sa dérivée seconde pour $x > 0$ existe, et vaut $s(s-1)\, x^{s-2} > 0$ (théorème 7 bis).

On a donc, pour $z_1, z_2, \ldots, z_n \geq 0$, et $\alpha_1, \alpha_2, \ldots, \alpha_n \geq 0$:

$$\left(\frac{\alpha_1 z_1 + \cdots + \alpha_n z_n}{\alpha_1 + \cdots + \alpha_n}\right)^s \leqslant \frac{\alpha_1 z_1^s + \cdots + \alpha_n z_n^s}{\alpha_1 + \cdots + \alpha_n} \tag{III,10;43}$$

ou, en multipliant par $(\alpha_1 + \ldots + \alpha_n)^s$ et en élevant à la puissance $\frac{1}{s}$:

$$(\alpha_1 z_1 + \cdots + \alpha_n z_n) \leqslant \left(\alpha_1 z_1^s + \cdots + \alpha_n z_n^s\right)^{\frac{1}{s}} \times (\alpha_1 + \cdots + \alpha_n)^{1-\frac{1}{s}}. \tag{III,10;44}$$

Posons alors, pour $X_i, Y_i \geq 0$, $c_i \geq 0$, $\not p, q > 0$, $\frac{1}{\not p} + \frac{1}{q} = \frac{1}{r}$:

$$\alpha_i = c_i Y_i^q , \quad z_i = X_i^r Y_i^{r-q} , \quad s = \frac{\not p}{r}. \tag{III,10;45}$$

On en déduit

(III,10;46) $\frac{1}{s} = \frac{r}{p}$, $\quad 1 - \frac{1}{s} = \frac{r}{q}$ $\left(\text{car } \frac{r}{p} + \frac{r}{q} = 1\right)$

$$\alpha_i z_i = c_i (X_i Y_i)^r$$

$$\alpha_i z_i^s = c_i X_i^{rs} Y_i^{q+s(r-q)} = c_i X_i^p Y_i^{p+q-\frac{pq}{r}} = c_i X_i^p$$

En élevant alors (III,10;44) à la puissance $\frac{1}{r}$, on obtient

(III,10;47) $$\left(c_1 (X_1 Y_1)^r + \cdots + c_n (X_n Y_n)^r\right)^{\frac{1}{r}}$$

$$\leqslant \left(c_1 X_1^p + \cdots + c_n X_n^p\right)^{\frac{1}{p}} \left(c_1 Y_1^q + \cdots + c_n Y_n^q\right)^{\frac{1}{q}} ,$$

ce qui est l'inégalité de Hölder.

La fonction $x \longrightarrow \left(1 - x^{\frac{1}{p}}\right)^p$, $p \geqslant 1$, est convexe dans $[0, 1]$.

En effet, sa dérivée seconde dans $]0, 1[$

$$\left(1 - x^{\frac{1}{p}}\right)^{p-2} \left(1 - \frac{1}{p}\right) x^{\frac{1}{p}-2} \geqslant 0 .$$

On a donc, pour $0 \leqslant z_1, z_2, \ldots, z_n \leqslant 1$, et $\alpha, \ldots, \alpha_n \geqslant 0$:

(III,10;48) $$\left(1 - \left(\frac{\alpha_1 z_1 + \cdots + \alpha_n z_n}{\alpha_1 + \cdots + \alpha_n}\right)^{\frac{1}{p}}\right)^p \leqslant \frac{\alpha_1\left(1 - z_1^{\frac{1}{p}}\right)^p + \cdots + \alpha_n\left(1 - z_n^{\frac{1}{p}}\right)^p}{\alpha_1 + \cdots + \alpha_n}$$

En élevant a la puissance $\frac{1}{p}$, puis en multipliant par $(\alpha_1 + \cdots + \alpha_n)^{\frac{1}{p}}$, on obtient

(III,10;49) $$(\alpha_1 + \cdots + a_n)^{\frac{1}{p}} \leqslant \left(\alpha_1 z_1 + \cdots + \alpha_n z_n\right)^{\frac{1}{p}}$$

$$t \left(\alpha_1 \left(1 - z_1^{\frac{1}{p}}\right)^p + \cdots + \alpha_n \left(1 - z_n^{\frac{1}{p}}\right)^p\right)^{\frac{1}{p}}$$

Pour c_i , X_i , Y_i , 0 , posons

(III,10;50) $\alpha_i = c_i (X_i + Y_i)^p$, $\qquad \alpha_i \geqslant 0$.

$z_i = \left(\frac{X_i}{X_i + Y_i}\right)^p$, $\qquad 0 \leqslant z_i \leqslant 1$.

On en déduit

(III,10;51) $$\alpha_i z_i = c_i X_i^p$$

$$\alpha_i \left(1 - z_i^{\frac{1}{p}}\right)^p = c_i Y_i^p$$

Alors (III,10;49) devient

$$(\text{III},10;52)\qquad \left(c_1(X_1+Y_1)^p+\cdots+c_n(X_n+Y_n)^p\right)^{\frac{1}{p}}$$
$$\leqslant \left(c_1X_1^p+\cdots+c_nX_n^p\right)^{\frac{1}{p}} \left(c_1\ Y_1^p+\cdots+c_nY_n^p\right)^{\frac{1}{p}},$$

ce qui est l'inégalité de Minkowski.

On peut aussi déduire Minkowski de Hölder.

On a en effet :

$$(\text{III},10;53)\qquad \sum_{i=1}^{n}(X_i+Y_i)^p=\sum_{i=1}^{n}(X_i+Y_i)^{p-1}X_i+\sum_{i=1}^{n}(X_i+Y_i)^{p-1}Y_i\ .$$

D'après Hölder, relatif aux exposants p', p, $\frac{1}{p}+\frac{1}{p'}=1$ $\left(p'=\frac{p}{p-1}\right)$, on aura

$$(\text{III},10;54)\qquad \sum_{i=1}^{n}(X_i+Y_i)^p\leqslant\left(\sum_{i=1}^{n}(X_i+Y_i)^{(p-1)p'}\right)^{\frac{1}{p'}}\left(\sum_{i=1}^{n}X_i^p\right)^{\frac{1}{p}}$$
$$+\left(\sum_{i=1}^{n}(X_i+Y_i)^{(p-1)p'}\right)^{\frac{1}{p'}}\left(\sum_{i=1}^{n}Y_i^p\right)^{\frac{1}{p}}\ .$$

Comme $(p-1)p'=p$, il suffit de multiplier les 2 membres par $\left(\sum_{i=1}^{n}(X_i+Y_i)^p\right)^{-\frac{1}{p'}}$ pour ootenir

$$(\text{III},10;55)\qquad \left(\sum_{i=1}^{n}(X_i+Y_i)^p\right)^{1-\frac{1}{p'}}\leqslant\left(\sum_{i=1}^{n}X_i^p\right)^{\frac{1}{p}}+\left(\sum_{i=1}^{n}Y_i^p\right)^{\frac{1}{p}},$$

ce qui est Minkowski puisque $1-\frac{1}{p'}=\frac{1}{p}$.

§ 11 CALCUL DES VARIATIONS

Position du problème

Considérons le problème de mécanique suivant :

Considérons, dans l'espace à trois dimensions, deux points A et B et une courbe $\mathscr{C}$, de classe C^1 , joignant ces deux points. Un point matériel est abandonné sur cette courbe sans frottement, sous l'influence de la pesanteur.

Partant du point A avec une vitesse initiale nulle, il arrive donc au point B au bout d'un certain temps t . A et B étant donnés, ce temps est une fonction de la courbe considérée; comment choisir cette courbe pour que ce temps t soit minimum ? Une telle courbe s'appelle brachistochrone ("temps le plus court'). Considérons pour l'instant comme évident que la courbe qui réalise le minimum doit être située dans le plan vertical du segment AB . Prenons alors, dans ce plan l'axe des x horizontal, l'axe des z vertical dirigé vers le bas; soient $z=a$, $x=\alpha$, et $z=b$, $x=\beta$, les coordonnées de A et B respectivement. Les théorèmes classiques de la chute des corps expriment qu'en un point donné de la courbe, compte tenu de l'absence de frottement, la vitesse v du point matériel est donnéepar la formule :

(III,11;1) $$v^2 = 2g(z-a) .$$

On en déduit la différentielle du temps en fonction de la différentielle de l'abscisse curviligne : $dt = \frac{ds}{\sqrt{2g(z-a)}}$, de sorte que finalement le temps mis par le mobile pour aller de A à B est donné par l'intégrale :

(III,11;2) $$\int_a^b \frac{\sqrt{1+x'^2}}{\sqrt{2g(z-a)}}\,dz = \int_a^b \frac{\sqrt{1+f'^2(z)}}{\sqrt{2g(z-\alpha)}}\,dz ,$$

dans laquelle on suppose qu'on s'est donné la courbe en se donnant x en fonction de z, $x = f(z)$ * . On est donc ramené au problème mathématique suivant : Trouver une fonction f réelle de la variable réelle z, telle que l'intégrale (III,11;2), où a, b et g sont donnés, ainsi que $f(a) = \alpha$, $f(b)=\beta$, soit aussi petite que possible.

Nous pouvons alors poser le problème général de la façon suivante : Soient $[a,b]$ un intervalle réel, F un espace affine normé sur le corps des réels, $\mathcal{U}$ un ouvert de $F\times\vec{F}$; soit, d'autre part, $L : (x,(y,\vec{Y})) \rightarrow L(x,y,\vec{Y})$, une fonction <u>réelle</u> continue donnée, définie sur $[a,b]\times\mathcal{U}$. Soit alors $f: x \rightarrow y = f(x)$, une fonction continuement dérivable définie sur $[a,b]$, à valeurs dans F , telle que tous les $(f(x), \vec{f}'(x))$, pour $x \in [a,b]$, soient dans $\mathcal{U}$ **

Dans ces conditions, la fonction composée $x \rightarrow L(x,f(x),\vec{f}'(x))$ est une fonction continue sur $[a,b]$ à valeurs dans $\mathbb{R}$. On peut alors calculer l'intégrale :

* La possibilité d'exprimer x en fonction de z est contestable : voir discussion page 367 .

** Comme il s'agit d'une fonction sur $\mathbb{R}$, $\vec{f}'(x)$ est le <u>vecteur dérivé</u> au point x , élément de $\vec{F}$, défini par (III,3;1); il vaut $f'(x)\cdot 1$, si $f'(x) \in \mathcal{L}(\mathbb{R};\vec{F})$ est l'application dérivée. D'autre part $[a,b]$ n'est pas un ouvert de $\mathbb{R}$; voir à ce sujet le début du § 2 page 184.

$$J(f) = \int_a^b L(x, f(x), \vec{f}'(x))\, dx \,, \tag{III,11;3}$$

qui est un nombre réel.

Ce nombre réel dépend du choix de la fonction f ; il définit donc une fonction $J : f \longrightarrow J(f)$; on se propose de chercher, a et b étant fixés ainsi que L, parmi toutes les fonctions f prenant des valeurs fixées α, β aux deux extrémités a, b celle qui rend l'intégrale $J(\;)$ maxima ou minima. Si F est un espace de dimension finie m, on peut, par le choix d'un référentiel, l'identifier à $\mathbb{R}^m$; alors une fonction définie sur $[a,b]$, à valeurs dans F, n'est pas autre chose qu'un système de m fonctions réelles f_i, $i=1,2,\ldots m$, définies sur $[a,b]$, et la fonction dérivée est le système des m fonctions dérivées f'_i ; L est alors une fonction réelle de $2m+1$ variables indépendantes x, y_i, Y_i, $i = 1,2,\ldots m$; et l'on cherche, parmi tous les systèmes de fonctions f_i, $i=1,2,\ldots m$, continuement dérivables sur $[a,b]$ et prenant des valeurs données en a et b, celui qui rend extrema l'intégrale

2

$$J(f_1, f_2, \ldots f_m) = \int_a^b L(x, f_1(x), f_2(x), \ldots f_m(x), f'_1(x), f'_2(x), \ldots f'_m(x))\, dx . \tag{III,11;4}$$

Rappelons que l'espace $E = \left(F^{[a,b]}\right)_{cb;1}$, des fonctions continuement dérivables définies sur $[a,b]$ à valeurs dans F, est un espace affine normé, d'espace vectoriel normé associé $\left(\vec{F}^{[a,b]}\right)_{cb;1}$ * ; la distance de deux éléments f, g, de l'espace affine E, est donnée par la formule (III,3;41) :

$$|||\overrightarrow{f-g}|||_1 = \sup_{a \leqslant x \leqslant b} \left(\|\overrightarrow{f(x)-g(x)}\|, \|\vec{f}'(x) - \vec{g}'(x)\|\right). \tag{III,11;5}$$

Nous avons ici à considérer d'abord un sous-ensemble de l'espace affine E, l'ensemble Ω formé des f telles que l'image par $(f, \vec{f}')$ de l'intervalle $[a,b]$ soit contenue dans $\mathcal{U}$, ouvert de $F \times \vec{F}$.

* Comme $[a,b]$ est compact, une fonction continue sur $[a,b]$ est bornée, le symbole cb, signifiant continue bornée, pourrait être remplacé par c, signifiant continue.

Théorème 37 - soit $[a,b]$ un segment de $\mathbb{R}$, F un espace affine normé sur le corps des réels, $\mathcal{U}$ un ouvert de $F \times \vec{F}$. L'ensemble Ω des fonctions f définies sur $[a,b]$ à valeurs dans F, telles que $(f(x), \vec{f}'(x))$ soit dans $\mathcal{U}$ pour tout x de $[a,b]$, est un ensemble ouvert dans l'espace $E = (F^{[a,b]})_{cb;1}$ des fonctions continuement dérivables définies sur $[a,b]$ à valeurs dans F.

Soit en effet f_0 un point de cet ensemble. L'image par $(f_0, \vec{f}_0')$ de l'intervalle compact $[a,b]$ est un compact K de $\mathcal{U}$ (théorème 28 du chapitre II). D'après ce que nous avons vu page 83, la distance de K à $\complement\mathcal{U}$ est un nombre $\delta > 0$. Considérons alors, dans E, la boule ouverte de centre f_0 et de rayon δ, formée de toutes les fonctions vérifiant $|||f - f_0||| < \delta$; pour une telle f, nous voyons que $(f, \vec{f}')$ prend toujours ses valeurs dans $\mathcal{U}$; autrement dit toute cette boule appartient à Ω, et Ω est bien ouvert dans E. L'ensemble des fonctions $f \in E$ qui vérifient $f(a)=\alpha$, $f(b)=\beta$ donnés, est un sous-espace affine de E_0 de; son sous-espace vectoriel associé est l'ensemble $\vec{E}_0$ des fonctions de $\vec{E}$ qui sont nulles en a et b. Alors Ω découpe sur E_0 un ouvert Ω_0 de E_0. Alors $J : f \longrightarrow J(f)$, est une fonction réelle définie ~~sur~~ l'ouvert Ω_0 de l'espace affine normé E_0, et nous cherchons un maximum ou un minimum de cette fonction. Le théoréme 23 doit nous donner une condition nécessaire. Nous voyons ici l'intérêt qu'il y a eu à démontrer le théorème 23 et tous les théorèmes du calcul différentiel dans le cas d'espaces normés de dimension infinie: en effet, même si F est de dimension finie, même si c'est, par exemple, la droite réelle $\mathbb{R}$, l'espace E, qui Intervient ici, est, lui, de dimension infinie.

Pour alléger les énoncés, nous ne répéterons pas les données $[a,b] \subset \mathbb{R}, F, \mathcal{U}, \Omega, L, J$, quand ce ne sera pas indispensable.

Dérivabilité de J

Nous allons donc étudier la dérivabilité de la fonction J; nous nous placerons plus généralement sur l'espace affine E, et nous donnerons la dérivée $J'(f_0)$ en f_0; la restriction de J à E_0 aura alors pour dérivée en $f_0 \in E_0$ la restriction de $J'(f_0)$ à $\vec{E}_0$. Nous appellerons $\vec{\delta f}$ un accroissement donné à f; c'est un élément de $\vec{E}$, donc une fonction $x \longrightarrow \vec{\delta f}(x)$ sur $[a,b]$ à valeurs dans $\vec{F}$. Nous chercherons la "partie principale" δJ de l'accroissement ΔJ de J, quand nous remplaçons f_0 par $f_0 + \vec{\delta f}$. Il y aurait danger à utiliser la notation $\vec{df}$, qui, f étant une application de $[a,b]$ dans F, signifie habituellement la différentielle de cette application; $\vec{df} = \vec{f}'(x)\,dx \in \vec{F}$. Ayant mis $\vec{\delta f}$, nous mettons δJ par analogie.

Théorème 38 - La fonction réelle J , définie sur Ω par (III,11;3), est continuement dérivable, et sa dérivée $\overleftarrow{J'(f_0)} \in \mathcal{L}(\vec{E};\mathbb{R}) = \overleftarrow{E'}$, au point f_0 , est donnée, par :

$$(\text{III},11;6) \qquad \overrightarrow{\delta f} \longrightarrow \delta J = \overleftarrow{J'(f_0)}\cdot\overrightarrow{\delta f}$$

$$= \int_a^b \left(L'(x,f_0(x),\overrightarrow{f_0'(x)})\cdot(0,\overrightarrow{\delta f(x)},\overrightarrow{\delta f'(x)})\right)dx = \int_a^b \left(\overleftarrow{\partial_2 L}(x,f_0(x),\overrightarrow{f_0'(x)})\cdot\overrightarrow{\delta f}(x) + \overleftarrow{\partial_3 L}(x,f_0(x),\overrightarrow{f_0'(x)})\cdot\overrightarrow{\delta f'}(x)\right)dx$$

où L' est la dérivée totale, et où $\overleftarrow{\partial_2 L}$ et $\overleftarrow{\partial_3 L}$ sont les dérivées partielles de la fonction L (définie sur un ouvert de $[a,b]\times F\times\vec{F}$) par rapport à la 2ème et à la 3ème variable.

Notons qu'au lieu de $\overleftarrow{\partial_2 L}$ et $\overleftarrow{\partial_3 L}$, on peut écrire $\overleftarrow{\frac{\partial L}{\partial y}}$ et $\frac{\partial L}{\partial \vec{Y}}$, puisque nous avons appelé y et $\vec{Y}$ les variables de F et $\vec{F}$.

Comme, au lieu de L $(x,f(x),\vec{f'}(x))$, on écrit aussi en abrégé, L $(x, y, \vec{y}')$, on pourra aussi écrire $\frac{\partial L}{\partial y}$ et $\overleftarrow{\frac{\partial L}{\partial \vec{y}'}}$ au lieu de $\overleftarrow{\partial_2 L}$ et $\overleftarrow{\partial_3 L}$; ce sont les dérivées de la fonction $(x,y,\vec{y}') \longrightarrow L(x,y,\vec{y}')$ quand on considère $x\in[a,b]$, $y\in F$, $\vec{y}'\in\vec{F}$, comme des variables indépendantes. Une fois qu'elles sont calculées, on remplace y par $f(x)$, $\vec{y}'$, par $f_0'(x)$.

2

Comme $y \longrightarrow L(x,y,\vec{Y})$ est une application de F dans $\mathbb{R}$, $\overleftarrow{\partial_2 L}(x,y,\vec{Y})$ est un élément de $\mathcal{L}(\vec{F};\mathbb{R}) = \overleftarrow{F'}$, dual de?; d'où la flèche $\overleftarrow{\partial_2 L}$. De même pour $\overleftarrow{\partial_3 L}$. Alors, pour tout $x\in[a,b]$, $\overleftarrow{\partial_2 L}(x,f_0(x),\overrightarrow{f_0'}(x))\in\overleftarrow{F'}$, et $\overrightarrow{\delta f}(x)\in\vec{F}$, ce qui fait que $\overleftarrow{\partial_2 L}(x,f_0(x),\overrightarrow{f_0'}(x))\cdot\overrightarrow{\delta f}(x)$, dans (III,11;6) est un nombre réel; (III,11;6) sera l'intégrale d'une fonction réelle sur $[a,b]$.

Démonstration. La variation ΔJ de l'intégrale correspondant à la variation $\overrightarrow{\delta f}$ de f est donnée par la formule

(III,11;7) $$\Delta J = \int_a^b \left[L\left(x, f_0(x) + \overrightarrow{\delta f}(x), \overrightarrow{f_0'}(x) + \overrightarrow{\delta f'}(x)\right) - L\left(x, f_0(x), \overrightarrow{f_0'}(x)\right)\right] dx \, .$$

La formule des accroissements finis (corollaire 1 du théorème 13), pour x fixé, donne immédiatement :

(III,11;8) $$\begin{cases} L\left(x, f_0(x) + \overrightarrow{\delta f}(x), \overrightarrow{f_0'}(x) + \overrightarrow{\delta f}'(x)\right) - L\left(x, f_0(x), \overrightarrow{f_0'}(x)\right) \\ = L'\left(x, f_0(x), \overrightarrow{f_0'}(x)\right) . \left(0, \overrightarrow{\delta f}(x), \overrightarrow{\delta f}'(x)\right) + R(x) \\ = \overleftarrow{\partial_2 L}\left(x, f_0(x), \overrightarrow{f_0'}(x)\right) . \overrightarrow{\delta f}(x) \\ + \overleftarrow{\partial_3 L}\left(x, f_0(x), \overrightarrow{f_0'}(x)\right) . \overrightarrow{\delta f}'(x) + R(x) \quad , \end{cases}$$

où $R(x)$ admet la majoration

(III,11;9) $$|R(x)| \leq \operatorname*{Sup}_{\substack{\|\vec{\xi}\| \leq \|\overrightarrow{\delta f}(x)\| \\ \|\vec{\xi'}\| \leq \|\overrightarrow{\delta f}'(x)\|}} \left\| L'\left(x, f_0(x) + \vec{\xi}, \overrightarrow{f_0'}(x) + \vec{\xi'}\right) - L\left(x, f_0(x), \overrightarrow{f_0'}(x)\right)\right\| \times \left\|\left(0, \overrightarrow{\delta f}(x); \overrightarrow{\delta f}'(x)\right)\right\|^{*}$$

Utilisons alors le théorème d'uniforme continuité, théorème 31 du chapitre II, sous sa forme améliorée (**). La fonction L' est supposée continue sur $[a,b] \times \mathcal{U}$; l'ensemble des points $\left(x, f_0(x), \overrightarrow{f_0'}(x)\right)$, $x \in [a,b]$, est contenu dans le compact $[a,b] \times K$ de $[a,b] \times \mathcal{U}$ (notation de la démonstration du théorème 37).

* Nous prenons, sur $\mathbb{R} \times \vec{F} \times \vec{F}$, la norme $\|(X, \vec{Y}, \vec{Z})\| = \operatorname{Sup}.(|X|, \|\vec{Y}\|, \|\vec{Z}\|)$

(**) Ce théorème est le suivant : Si f est une application continue d'un espace métrique E dans un espace métrique F, et si K est un compact de E, alors, quel que soit $\varepsilon > 0$, il existe $\eta > 0$ tel que $d(x', x'') \leq \eta$, $x' \in K$, entraine $d(f(x'), f(x'')) \leq \varepsilon$; le théorème 31 lui-même est relatif à $E = K$. La démonstration est la même. Si c'était inexact, on pourrait trouver deux suites $x_n' \in K$, $x_n'' \in E$, $d(x_n', x_n'') \leq \frac{1}{n}$, $d(f(x_n'), f(x_n'')) > \varepsilon$. Comme K est compact, on pourrait extraire des x_n' une suite partielle, convergeant vers un élément c de K. Alors x_n'' tendrait aussi vers c dans E. Comme f est continue en c, $f(x_n')$ et $f(x_n'')$ tendraient vers $f(c)$, donc $d(f(x_n'), f(x_n''))$ tendrait vers 0, d'où une contradiction.

Donc, étant donné $\varepsilon > 0$, il existe η, $0 < \eta < \delta$, tel que $x \in [a,b]$, $\|\vec{\xi}\| \leqslant \eta$, $\|\vec{\xi}'\| \leqslant \eta$, entraîne :

(III, 11; 10) $$\left\| L'(x, f_0(x) + \vec{\xi}, \vec{f}'_0(x) + \vec{\xi}') - L'(x, f_0(x), \vec{f}'_0(x)) \right\| \leqslant \frac{\varepsilon}{b-a} .$$

Alors $\left\|\left| \overrightarrow{\delta f} \right\|\right|_1 \leqslant \eta$ entraîne

(III, 11; 11) $$|R(x)| \leqslant \frac{\varepsilon}{b-a} \left\|\left| \overrightarrow{\delta f} \right\|\right|_1$$

On a donc

(III, 11; 11bis) $$\Delta J = \int_a^b \left(\overleftarrow{\partial_2 L} . \overrightarrow{\delta f} + \overleftarrow{\partial_3 L} . \overrightarrow{\delta f}' \right) dx + \int_a^b R(x)\, dx$$

chacune des fonctions $x \longrightarrow \overleftarrow{\partial_2 L}(x, f_0(x), \vec{f}'_0(x)) . \overrightarrow{\delta f}(x)$, $x \longrightarrow \overleftarrow{\partial_3 L}(x, f_0(x), \vec{f}'_0(x)) . \overrightarrow{\delta f}'(x)$, et par conséquent aussi $x \longrightarrow R(x)$, étant continue (en effet, $\partial_2 L$ est continue de $[a,b] \times \mathcal{U}$ dans $\overleftarrow{F}'$ puisque L est continuement dérivable ; f_0, $\vec{f}'_0$, $\overrightarrow{\delta f}$, sont continues de $[a,b]$ dans $F, \vec{F}, \vec{F}$; donc $x \longrightarrow \overleftarrow{\partial_2 L}(x, f_0(x), \vec{f}'_0(x))$ est continue de $[a,b]$ dans $\overleftarrow{F}'$, et comme $(\overleftarrow{u}, \vec{X}) \longrightarrow \overleftarrow{u} . \vec{X} = \langle \overleftarrow{u}, \vec{X} \rangle$ est continue de $(\overleftarrow{F}' = \mathcal{L}(\vec{F}; \mathbb{R})) \times \vec{F}$ dans $\mathbb{R}$, $x \longrightarrow \overleftarrow{\partial_2 L}(x, f_0(x), \vec{f}'_0(x)) . \overrightarrow{\delta f}(x)$ est bien continue de $[a,b]$ dans $\mathbb{R}$).

La 2ème intégrale est majorée par $\varepsilon \left\|\left| \overrightarrow{\delta f} \right\|\right|_1$, pour $\left\|\left| \overrightarrow{\delta f} \right\|\right|_1 \leqslant \eta$; donc elle est infiniment petite devant $\left\|\left| \overrightarrow{\delta f} \right\|\right|_1$, quand $\left\|\left| \overrightarrow{\delta f} \right\|\right|_1$ tend vers 0 .

Par ailleurs la 1ère intégrale définit une forme linéaire continue sur $\vec{E}$; en effet, on a

(III,11;11ter) $$\left| \int_a^b \left(\overleftarrow{\partial_2 L} \cdot \overrightarrow{\delta f} + \overleftarrow{\partial_3 L} \cdot \overrightarrow{\delta f'} \right) dx \right|$$

$$= \left| \int_a^b \left(L'(x, f_0(x), \overrightarrow{f_0'}(x)) \right) \cdot \left(0, \overrightarrow{\delta f}(x), \overrightarrow{\delta f}'(x) \right) dx \right|$$

$$\leqslant (b-a) M \, |||\overrightarrow{\delta f}|||_1 \quad , \qquad M = \sup_{a \leqslant x \leqslant b} \| L'(x, f_0(x), \overrightarrow{f_0'}(x)) \|$$

Nous avons donc bien montré que J est dérivable en $f_0 \in \Omega$, et que sa dérivée est donnée par (III,11;6).

Montrons maintenant que J est continuement dérivable sur Ω Considérons deux éléments f_0 , f de Ω..

(III,11;11quarto) $$\overleftarrow{J'(f)} \cdot \overrightarrow{\delta f} - \overleftarrow{J'(f_0)} \cdot \overrightarrow{\delta f}$$

$$= \int_a^b \left(L'(x, f(x), \overrightarrow{f}'(x)) \quad L'(x, f_0(x), \overrightarrow{f_0'}(x)) \right) \cdot \left(0, \overrightarrow{\delta f}(x), \overrightarrow{\delta f}'(x) \right) dx .$$

Alors, d'après (III,11;10), $|||\overrightarrow{f - f_0}|||_1 \leqslant \eta$ entraine

(III,11;12) $$\left| \left(\overleftarrow{J'(f)} - \overleftarrow{J'(f_0)} \right) \cdot \overrightarrow{\delta f} \right| \leqslant \varepsilon \, |||\overrightarrow{\delta f}|||_1 \text{ et}$$

(III,11;12bis) $$\| \overleftarrow{J'(f)} - \overleftarrow{J'(f_0)} \| = \sup_{|||\overrightarrow{\delta f}|||_1 \leqslant 1} \left| \left(\overleftarrow{J'(f)} - \overleftarrow{J'(f_0)} \right) \cdot \overrightarrow{\delta f} \right| \leqslant \varepsilon \; ;$$

Ainsi $\overleftarrow{J'(f)} - \overleftarrow{J'(f_0)}$ converge bien vers $\overleftarrow{0}$ dans $\overleftarrow{E'}$ quand f tend vers f_0 dans **E** , et **J** est bien de classe C^1.

<u>Exemple</u> Il n'est pas inutile de revoir rapidement ce que nous venons de faire dans le cas simple où $F = \mathbb{R}$, $\mathcal{U} = [a,b] \times \mathbb{R} \times \mathbb{R}$ tout entier. On suppose donc que L est une fonction de classe C^1 de **3** variables réelles, que nous écrivons x, y, y', $x \in [a,b]$, $y \in \mathbb{R}$, $y' \in \mathbb{R}$. On considère la fonction J définie sur $(\mathbb{R}^{[a,b]})_{cb;1}$:

(III,11,12ter) $$J(f) = \int_a^b L(x, f(x), f'(x)) \, dx .$$

Un donne a f un accroissement δf , qui est aussi une fonction C^1 sur $[a,b]$. L'accroissement ΔJ de J est donné par

(III,11;12quarto) $$\Delta J = \int_a^b \left(L(x, f_0 + \delta f, f_0' + \delta f') - L(x, f_0, f_0') \right) dx .$$

Sa "partie principale" est la différentielle

(III,11;12quinto) $$\left\{ \begin{aligned} \delta J = \int_a^b \Big(& \frac{\partial L}{\partial y}(x, f_0(x), f_0'(x))\, \delta f(x) \\ & + \frac{\partial L}{\partial y'}(x, f_0(x), f_0'(x))\, \delta f'(x) \Big) dx . \end{aligned} \right.$$

La justification de cette formule est une majoration du type

(III,11,12sexto) $$|\delta J - \Delta J| \leq \varepsilon \|\delta f\|_1 \quad \text{pour} \quad \|\delta f\|_1 \leq \eta \qquad \text{(page 356)} ;$$

elle s'obtient par l'application a L de la formule des accroissements finis.

Nous allons maintenant transformer (III,11;6), en faisant sur L , et sur l'élément f_0 de Ω où nous calculons la dérivée de J , une hypothèse restrictive : L et f_0 sont de classe C^2 .

Théorème 39 - Si, dans les conditions de l'énoncé du théorème 38, L est en outre de classe C^2 sur $[a,b] \times \mathcal{U}$, et si $f_0 \in \Omega$ est de classe C^2 sur $[a,b]$, la dérivée $J'(f_0)$ peut s'écrire :

(III,11;13)
$$\delta J = \overleftarrow{J'(f_0)} \cdot \overrightarrow{\delta f}$$
$$= \left[\overleftarrow{\partial_3 L}\left(x, f_0(x), \vec{f}'_0(x)\right) \cdot \overrightarrow{\delta f}(x) \right]_{x=a}^{x=b}$$
$$+ \int_a^b \left[\overleftarrow{\partial_2 L}\left(x, f_0(x), \vec{f}'_0(x)\right) - \frac{d}{dx}\left(\overleftarrow{\partial_3 L}\left(x, f_0(x), \vec{f}'_0(x)\right)\right) \right] \cdot \overrightarrow{\delta f}(x)\, dx$$

Le crochet $[\ \ ]_{x=a}^{x=b}$ veut dire la différence des valeurs pour $x = b$ et pour $x = a$.

Démonstration - Il suffit de faire une intégration par parties dans (III,11;6). Dans la formule , $\frac{d}{dx}$ est la dérivée de l'application $x \to \overleftarrow{\partial_3 L}(x, f(x), \vec{f}'_0(x))$ de $[a,b]$ dans $\overleftarrow{F'}$; ce n'est donc pas une dérivée partielle $\partial_1 L = \frac{\partial L}{\partial x}$, mais une dérivée totale, qui tient compte de ce que f_0 et $\vec{f}'_0$ sont des fonctions de x. L'intégration par parties que nous venons de faire est un peu plus générale que celle qui a été vue en Mathématiques Spéciales. C'est la formule

(III,11;14)
$$\int_a^b B(\vec{u}(x), \vec{v}'(x))\, dx = \left[B(\vec{u}(x), \vec{v}(x)) \right]_{x=a}^{x=b} - \int_a^b B(\vec{u}'(x), \vec{v}(x))\, dx,$$

où B est une forme bilinéaire continue; elle résulte immédiatement de ce que la dérivée de $x \longrightarrow B(\vec{u}(x), \vec{v}(x))$ est $x \to B(\vec{u}'(x), \vec{v}(x)) + B(\vec{u}(x), \vec{v}'(x))$ (théorème 12). Ici, il s'agit, pour B, de l'application bilinéaire continue $(\overleftarrow{u}, \vec{v}) \longrightarrow \overleftarrow{u} \cdot \vec{v}$ ou $\langle \overleftarrow{u}, \vec{v} \rangle$ de $\overleftarrow{F'} \times \vec{F}$ dans $\mathbb{R}$. Naturellement, on ne peut effectuer cette intégration par parties que si les fonctions u et v considérées sont continuement dérivables. C'est pourquoi nous avons supposé f_0 de classe C^2 : comme L est de classe C^2, $\partial_3 L$ est de classe C^1, et f_0, $\vec{f}'_0$ sont de classe C^1, alors $u : x \longrightarrow \partial_3 L(x, f_0(x), \vec{f}'_0(x))$ est bien de classe C^1, d'après le corollaire 5 du théorème 11.

<u>Exemple</u> - si $F = \mathbb{R}$, $\mathcal{U} = \mathbb{R} \times \mathbb{R}$, la formule (III,11;12 quinto), traitée suivant le théorème 39, donnera le cas particulier suivant de (III,11;13) :

$$\delta J = \left[\frac{\partial L}{\partial y'}(x, f_0(x), f_0'(x))\,\delta f(x)\right]_{x=a}^{x=b}$$

$$+ \int_a^b \left[\frac{\partial L}{\partial y}\left(x, f_0(x), f_0'(x)\right) - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}(x, f_0(x), f_0'(x))\right)\right]\delta f(x)\,dx .$$

Condition nécessaire d'extremum

Revenons au **problème** d'extremum **initialement** posé. Une condition nécessaire pour que $f_0 \in \Omega_0$ (c'est-à-dire $f_0 \in \Omega$. , avec $f_0(a) = \alpha$ et $f_0(b) = \beta$ donnés), ouvert de $\vec{E}_0$, rende J extrema sur Ω_0, est que $J'(f_0)$ soit nulle sur $\vec{E}_0$ (théorème 23).

Théorème 40 - Soient $[a, b]$ un segment de $\mathbb{R}$, F un espace affine normé, $\mathcal{U}$ un ouvert de $F \times \vec{F}$, Ω l'ensemble des applications f continuement dérivables de $[a,b]$ dans F telles que $(f(x), f'(x)) \in \mathcal{U}$ pour tout x de $[a,b]$, Ω_0 le sous-ensemble de Ω formé des fonctions f vérifiant $f(a) = \alpha$, $f(b) = \beta$ donnés, L une fonction de classe C^2 sur $[a,b] \times \mathcal{U}$. Pour qu'une fonction $f_0 \in \Omega_0$ de classe C^2, rende maxima ou minima, ou plus généralement stationnaire sur Ω_0 l'intégrale J de (III,11;3), il est nécessaire que f_0 soit solution de l'équation différentielle du 2ème ordre:

(III,11;15)
$$\overleftarrow{\partial_2 L}\left(x, f_0(x), \vec{f}_0'(x)\right) - \frac{d}{dx}\left(\overleftarrow{\partial_3 L}\left(x, f_0(x), \vec{f}_0'(x)\right)\right) = \overleftarrow{0}$$

ou

(III,11;16)
$$\frac{\overleftarrow{\partial L}}{\partial y} - \frac{d}{dx}\left(\frac{\overleftarrow{\partial L}}{\partial \vec{y}'}\right) = \overleftarrow{0}$$

Cette équation s'appelle **équation d'Euler** et une solution de cette **équation** s'appelle une extremale de **l'intégrale** J ou de la fonction L ,

Démonstration - Si $\overrightarrow{\delta f} \in \vec{E}_0$, on a $\overrightarrow{\delta f}(a) = \overrightarrow{\delta f}(b) = \vec{0}$

Alors, dans (III,11;13), $\Big[\quad\Big]_{x=a}^{x=b}$ est nul.

Alors la condition **nécessaire** d'extremum $\delta J = 0$ **exprime** que **l'intégrale qui** figure dans le 2ème membre de (III,11;13) est nulle, quelle que soit la fonction $\overrightarrow{\delta f} \in \vec{E}_0$. Pour en déduire que f_0 satisfait **à l'équation différentielle** (III,11;15), nous n'aurons qu'à appliquer **à** la **fonction** Φ:

(III,11;17)
$$x \longrightarrow \overleftarrow{\Phi}(x) = \overleftarrow{\partial_2 L}\left(x, f_0(x), \vec{f}_0'(x)\right) - \frac{d}{dx}\left(\overleftarrow{\partial_3 L}\left(x, f_0(x), \vec{f}_0'(x)\right)\right)$$

le **lemme** suivant :

Lemme de Haar

Si $\overleftarrow{\Phi}$ est une fonction continue sur un intervalle $[a,b]$, à valeurs dans le dual $\overleftarrow{F}'$ d'un espace vectoriel normé $\vec{F}$, et si, pour toute fonction $\vec{\eta}$, définie sur $[a,b]$, à valeurs dans $\vec{F}$, continuement dérivable, s'annulant aux deux extrémités de l'intervalle, on a :

(III,11;18) $$\int_a^b <\overleftarrow{\Phi(x)},\ \overrightarrow{\eta(x)}>\, dx = 0\ ,$$

alors la fonction $\overleftarrow{\Phi}$ est identiquement nulle sur $[a,b]$.

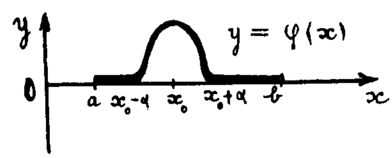
Supposons que $\overleftarrow{\Phi}$ ne soit pas identiquement nulle, et montrons que nous aboutissons à une contradiction.
Soit donc x_0 un point tel que $\overleftarrow{\Phi}(x_0) \neq 0$.

On peut évidemment toujours supposer x_0 différent des extrémités a et b, car si $\overleftarrow{\Phi}$, supposée continue, était identiquement nulle dans l'intervalle ouvert $]a,b[$, elle serait aussi identiquement nulle dans l'intervalle fermé.
Dire que $\overleftarrow{\Phi}(x_0)$, forme linéaire continue sur $\vec{F}$, est distincte de 0 c'est dire qu'il existe un point $\vec{e}$ de $\vec{E}$ tel que la valeur de cette forme linéaire sur $\vec{e}$ soit $\neq 0$.
On peut d'ailleurs choisir $\vec{e}$ tel que la valeur de la forme linéaire en $\vec{e}$ soit > 0 (sinon, on remplace $\vec{e}$ par $-\vec{e}$).
Il existe alors, d'après la continuité de $\overleftarrow{\Phi}$, un nombre α assez faible pour que l'intervalle $[x_0-\alpha, x_0+\alpha]$ soit dans $[a,b]$, et que la fonction continue scalaire $x \longrightarrow <\overleftarrow{\Phi}(x), \vec{e}>$ y soit > 0.

Soit alors Q une fonction scalaire continuement dérivable, >0 dans $]x_0-\alpha, x_0+\alpha[$, et nulle en dehors.

On pourra prendre, par exemple, la fonction

(III,11;19) $$\varphi(x) = \begin{cases} 0 \text{ si } x \notin\ ]x_0-\alpha\ ,\ x_0+\alpha[\\ \cos^2\left(\dfrac{\pi}{2\alpha}(x-x_0)\right) \text{ ou } \left(\alpha^2-(x-x_0)^2\right)^2 \text{ si } x \in\ ]x_0-\alpha, x_0+\alpha[\end{cases}$$



Prenons alors, pour fonction $\vec{\eta}$, la fonction :
$\vec{\eta}(x) = \vec{e}\,\varphi(x)$. Elle est évidemment continuement dérivable dans $[a,b]$, et s'annule aux deux extrémités. La fonction $x \longrightarrow <\overleftarrow{\Phi}(x), \overrightarrow{\eta(x)}>$ est continue, >0 dans $]x_0-\alpha, x_0+\alpha[$, nulle ailleurs : son intégrale est alors > 0.

Nous sommes donc arrivés à une contradiction; il est donc exact que $\overleftarrow{\Phi}$ est identiquement nulle, le lemme est démontré.

Exemple - Pour que f_0 rende extrema l'intégrale J parmi toutes les $f \in \mathbb{R}^{[a,b]}$ de classe C^1 (voir exémple après le théorème 38) verifiant les conditions aux limites $f(a) = \alpha$, $f(b) = \beta$ donnés, il faut et il suffit, si f_0 est de classe C^2, qu'elle vérifie

(III,11;19. a)
$$\int_a^b \left[\frac{\partial L}{\partial y}(x, f_0(x), f_0'(x)) - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}(x, f_0(x), f_0'(x))\right)\right] \delta f(x)\, dx = 0,$$

pour toute $\delta f \in \mathbb{R}^{[a,b]}$ de classe C^1 vérifiant $\delta f(a) = \delta f(b) = 0$; le lemme de Haar exprime que, pour cela, il est nécessaire et suffisant que f_0 vérifie $f_0(a) = \alpha$, $f_0(b) = \beta$, et soit solution de l'équation différentielle d'Euler

(III,11;19 .b)
$$\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) = 0.$$

Corollaire - Soit f_0 une fonction sur $[a,b]$ qui rende J stationnaire, dans les conditions du théorème 40. Alors, si $a_1 \geqslant a$, $b_1 \leqslant b$, f_0 rend aussi stationnaire l'intégrale

(III,11;19 bis)
$$J_1(f) = \int_{a_1}^{b_1} L(x, f(x), \vec{f}'(x))\, dx,$$

sur l'espace de toutes les fonctions f de classe C^1 définies sur $[a_1, b_1]$, à valeurs dans F, telles que (f, f') applique $[a_1, b_1]$ dans $\mathcal{U}$, et vérifiant $f(a_1) = \alpha_1 = f_0(a_1)$, $f(b_1) = \beta_1 = f_0(b_1)$.

En effet l'équation d'Euler ne fait pas intervenir les extrémités de l'intervalle.

Ce résultat n'était pas évident à priori, même dans le cas d'un maximum ou d'un minimum absolu. Supposons par exemple que f_0 soit un minimum pour J.

Alors, pour f donnée comme ci-dessus sur $[a_1, b_1]$, appelons $\tilde{f}$ la fonction égale à f dans $[a_1, b_1]$, et à f_0 dans $[a, a_1[$ et $]b_1, b]$. C'est une fonction continue, puisque $f(a_1) = f_0(a_1)$, $f(b_1) = f_0(b_1)$. Si elle est de classe C^1, nous aurons $J(\tilde{f}) \geqslant J(f_0)$ en vertu de la propriété de minimum, donc, en retranchant aux 2 intégrales la partie commune $\int_a^{a_1} + \int_{b_1}^{b}$, $J_1(f)$ 3 $J_1(f_0)$, et f_0 est encore un minimum pour J_1. Mais $\tilde{f}$ a bien, en a_1 et en b_1, une dérivée à droite et une dérivée à gauche, mais pas nécessairement égales, donc $\tilde{f}$ n'est pas de classe C^1, et ce raisonnement n'est pas applicable. Pour le rendre valable, on doit utiliser une approximation de f par des fonctions ae classe C^1 et faire un passage à la limite que nous nedétaillerons pas, et qui **montre, indépendamment** des équations a **Euler,** que f_0 reste un minimum pour J_1 .

Remarques 1°/ Il n'est nullement certain a priori qu'il existe effectivement un f_{10} où l'intégrale atteigne son maximum ou son minimum. En effet, Ω_0 est ouvert, et non compact. Voir remarque 1°/ après le théorème 23.

2°/ SI même on a pu prouver qu'il existe un élément f_0 de l'ouvert Ω_0 **considéré, où** l'intégrale atteigne son maximum ou son minimum, rien ne prouve que cet élément f_0 soit deux fols continuement dérivable, la dérivée seconde **de** f n'**ayant** en effet rien à faire avec le problème. S'il est légitime de faire des restrictions sur L , qui est une **donnée** du problème, il ne l'est pas du tout d'en faire sur f_0 , qui est une Inconnue du problème. Une démonstration un peu plus compliquéeque celle que nous avons **donnée** permet de lever **cette** objection, et de montrer, dans la plupart des cas, ... fonction de classe C' rendant J stationnaire est n3cessairement de classe C^2 .

3°/ En pratique, on écrira l'équation différentielle **(III,11;16),** et on cherchera la solution de cette équation, vérifiant les conditions aux **limites** $f(a) = \alpha$, $f(b) = \beta$. Nous avons, par rapport **à** f**, une équation** différentielle du second ordre; au lieu d'écrire, comme on en avait l'habitude , que la fonction f et **sa dérivée** première f' prennent **en** un point **donné** des valeurs Initiales **données,** on a à écrire deux condition8 d'une autre nature, **à** savoir que la fonction f prend en deux points donnés des valeurs données. Rien ne prouve, d'une fagon générale , l'existence d'une solution possédant ces **propriétés,** ni son **unicité.**

Z

-Revenons au cas **étudié à** la formule **(III,11;4). Ici** δf est un système **de** m **accroissements** $(\delta f_1, \delta f_2, \ldots, \delta f_m)$ La **dérivée** partielle $\frac{\partial f}{\partial y}(x, y, \vec{Y})$**, vecteur de** $F' = \mathbb{R}^m$, est ie système des m **dérivées** $\frac{\partial f}{\partial y_i}(x, y_1, y_2, \ldots y_m, Y_1, Y_2, \ldots Y_m)$, $i = 1, 2, \ldots m$; **tandis** que $\frac{\partial f}{\partial \vec{y}'}(x, y, \vec{Y})$ **est le système des** m **dérivées** $\frac{\partial f}{\partial y'_i}$ ou $\frac{\partial f}{\partial Y_i}(x, y_1, y_2, \ldots y_m, Y_1, Y_2, \ldots Y_m)$, i $1, 2, \ldots m$**. Enfin le produit scalaire** $\langle \overleftarrow{\alpha}, \vec{X} \rangle$ **ou** $\overleftarrow{\alpha} \cdot \vec{X}$ **, pour** $\overleftarrow{\alpha} = (\alpha_1, \alpha_2, \ldots \alpha_m)$, $\vec{X} = (X_1, X_2, \ldots X_m)$ est $\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_m X_m$ **. L'équation d'Euler** relative **à** la fonction $f = (f_1, f_2, \ldots f_m)$ est donc identique au **système** de ses m équations composantes :

(III,11;20) $$\frac{\partial L}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'_i}\right) = 0 \; ; \; i = 1,2,\dots m \; ; \text{ ou}$$

(III,11;21) $$\frac{\partial L}{\partial y_i} - \frac{\partial^2 L}{\partial y'_i \partial x} - \sum_{j=1}^{m} \frac{\partial^2 L}{\partial y'_i \partial y_j} y'_j - \sum_{j=1}^{m} \frac{\partial^2 L}{\partial y'_i \partial y'_j} y''_j = 0 \; ;$$

La solution de ce **système** dépend **de** $2m$ **constantes arbi - traires;** celles-ci doivent satisfaire aux $2m$ **équations**

$$f_i(a) = \alpha_i \; , \; f_i(b) = \beta_i \quad , \; i = 1,2,\dots m .$$

Il **n'est pas** inutile de récrire l'expression de δJ , **pour** $\delta f \in \vec{E}_0$ **:**

(III,11;21bis) $$\delta J = \int_a^b \left(\sum_{i=1}^{m} \left(\frac{\partial L}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'_i}\right)\right) \delta f_i \right) dx .$$

4°/ Nous avons écrit des conditions de variation du premier ordre, et en l'absence de conditions du second ordre (formule de Taylor), nous ne savons pas s'il **s'agit** d'un maximum, d'un minimum ou d'un col. Voir remarque **2°/** après le théorème 23.

Cas simples d'intégrabilité élémentaire des équations d'Euler

1°/ Supposons que L ne **dépende** pas de y **;** alors $\frac{\overleftarrow{\partial L}}{\partial y} \equiv 0$.
L'équation se ramène immédiatement **à** une **équation différentielle** du 1er ordre.

(III,11;21 ter) $$\frac{\overleftarrow{\partial L}}{\partial y'}(x , y') = \text{constante } \overleftarrow{c} \in \overleftarrow{F}$$

(si $F = \mathbb{R}$, $\frac{\partial L}{\partial y'}(x,y') = c$, constante **réelle).**

2°/ Supposons que L ne **dépende** pas de x . Les équations **d'Euler s'écrivent,en** prenant le cas particulier **de (III,11;21) :**

(III,11;21quarto) $$\frac{\partial L}{\partial y_i} - \sum_{j=1}^{m} \frac{\partial^2 L}{\partial y_j \partial y'_i} y'_j - \sum_{j=1}^{m} \frac{\partial^2 L}{\partial y'_i \partial y'_j} y''_j = 0 \; ; \qquad i = 1,2,\dots m$$

Par multiplication de la i **ème** par y'_i **et** addition, on obtient :

(III,11;21 quinto)
$$\sum_{i=1}^{m} \frac{\partial L}{\partial y_i} y'_i - \sum_{i,j=1}^{m} \frac{\partial^2 L}{\partial y_j \, \partial y'_i} y'_i y'_j$$
$$- \sum_{i,j=1}^{m} \frac{\partial^2 L}{\partial y'_i \, \partial y'_j} y'_i y''_j = 0 \quad , \quad \text{ou}$$

(III,11;21 sexto)
$$\frac{d}{dx}\left(L - \sum_{j=1}^{m} \frac{\partial L}{\partial y'_j} y'_j \right) = 0 \quad .$$

On a donc l'intégrale première

(III,11;21 septimo)
$$L - \sum_{j=1} \frac{\partial L}{\partial y'_j} y'_j = c \text{ , ou } L - \frac{\overleftarrow{\partial L}}{\partial y'} \cdot \vec{y}' = c \quad .$$

Nous en verrons une interprétation remarquable dans les problèmes de mécanique (formule (III,11:96)).

En particulier, si $m = 1$, on est entièrement ramené à une équation du premier ordre

(III,11;22)
$$L - \frac{\partial L}{\partial y'} y' = c \quad ,$$

après avoir toutefois introduit (par multiplication par y') la solution étrangère $y =$ constante.

<u>Exemple</u> - <u>Courbe brachistochrone.</u>

Reprenons l'exemple qui nous a servi d'introduction. On doit alors chercher le minimum de l'intégrale (III,11;2). Ici c'est z la variable indépendante,== $f(z)$ la fonction réelle de z à valeur réelle, etx'sa dérivée. L'équation d'Euler s'écrit alors (formule (III,11;21 ter) avec le changement de notation signalé) :

(III,11;23)
$$\frac{\partial L}{\partial x'} = \frac{x'}{\sqrt{1 + x'^2}} \frac{1}{\sqrt{2g(z-a)}} = \text{constante} = \frac{\pm 1}{\sqrt{4gc}} \quad ,$$

où c est une constante arbitraire > 0 . Nous laissons ainsi échapper le cas de la constante 0 , qui ne peut pas s'écrire $\pm \frac{1}{\sqrt{4gc}}$. Elle donne $x =$ constante= α ; elle correspond au cas où B est sur la verticale de A, et où bien entendu la solution du problème est donnée par le segment vertical AB .

Cette équation s'écrit immédiatement sous la nouvelle forme :

(III,11;24) $$x'^2 = \frac{z-a}{2c}(1+x'^2), \text{ ou } x'^2\left(1-\frac{z-a}{2c}\right) = \frac{z-a}{2c}$$

Si l'on fait le changement de variable (justifié par ce qu'on a nécessairement $0 \leqslant \frac{z-a}{c} \leqslant 2$)

(III,11;25) $$z = a + c(1-\cos u) \; ,$$

elle devient la nouvelle dquation

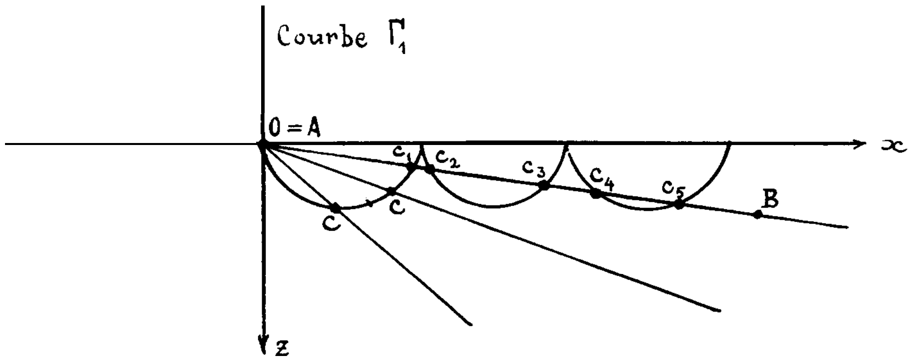
(III,11;26) $$\frac{dx}{dz} = \pm\sqrt{\frac{z-a}{2c-(z-a)}} = \pm\sqrt{\frac{1-\cos u}{1+\cos u}}\,, \quad \frac{dx}{du} = \pm c\sin u\sqrt{\frac{1-\cos u}{1+\cos u}} = \pm 2c\sin^2\frac{u}{2} = \pm c(1-\cos u)$$

qui se résout immédiatement par :

(III,11,27) $$z = a + c(1-\cos u) \quad x = \pm c(u-\sin u) + \text{constante}$$

La constante c et la constante supplémentaire de (III,11;27) doivent être déterminées pour que la courbe passe par **A** et B . Prenons, pour simplifier, A à l'origine des coordonnées. On a $z = a = 0$ pour $u = 2k\pi$; à cause de la périodicité (changement de u en $u-2k\pi$) on peut exiger que ce soit pour $u = 0$, alors on aura $x = 0$ si la constante de (III,11;27) est nulle. On peut prendre le signe+ car le signe $-$ est compensé par un changement de u en $-u$.

Alors on voit que <u>la courbe brachistochrone est une cycloïde Γ , ayant au point de départ A un point de rebroussement à tangente verticale.</u> La constante c doit être déterminée de manière qu'elle passe par B . Lorsque c varie, on a une famille de cycloïdes Γ_c homothétiques de celle Γ_1 qui correspond à $c = 1$ On déterminera alors c géométriquement comme suit : la droite AB coupe Γ en des points C_1, C_2, ... en nombre fini (sauf si B est sur l'horizontale de **A** alors il y a une infinité de solutions). Si C_i est l'un d'eux, l'homothétique de Γ_1 , de centre A , de rapport $\frac{AB}{AC_i}$, passe par B et répond à la question.



A peu près toutes les difficultés du calcul des variations sont condensées dans cet exemple.

1°/ Nous avons admis d'emblée que la courbe était dans le plan vertical de AB, quand nous avons posé le problème au début du paragraphe. Cette objection peut maintenant être. levée aisément. Prenons à priori 3 coordonnées dans $\mathbb{R}^3$, l'axe des z vertical vers le bas. Cherchons la courbe brachistochrone comme arc de courbe paramétrique de classe C^2, en la représentant paramétriquement par 3 fonctions $x(W)$, $y(W)$, $z(W)$. Nous supposerons que A et B correspondent à $W = W_1$ et $W = W_2$ fixés.

Alors le temps mis pour aller de A à B est

(III,11;28)
$$t = \int_{W_1}^{W_2} \frac{ds}{\sqrt{2g(z-a)}} = \int_{W_1}^{W_2} \frac{\sqrt{x'^2+y'^2+z'^2}}{\sqrt{2g(z-a)}}\, dw \quad ;$$

les fonctions x, y, z prennent des valeurs données (coordonnées de A et B) pour $W = W_1$, et $W = W_2$.

Comme il y a 3 fonctions Inconnues, nous avons 3 équations d'Euler. Les 2 premières sont, en vertu de (III,11; 21 ter) :

(III,11; 29)
$$\begin{cases} \dfrac{x'}{\sqrt{x'^2+y'^2+z'^2}} \quad \dfrac{1}{\sqrt{2g(z-a)}} = \text{constante;} \\[2ex] \dfrac{y'}{\sqrt{x'^2+y'^2+z'^2}} \quad \dfrac{1}{\sqrt{2g(z-a)}} = \text{constante;} \end{cases}$$

d'où l'on déduit une relation à coefficients constants $\lambda x' + \mu y' \doteq 0$, d'où $\lambda x + \mu y + \nu = 0$. Ainsi la courbe est sûrement dans un plan vertical, qui est donc celui de AB , et le choix de 2 nouvelles coordonnées dans ce plan était bien légitime.

Sans d'ailleurs rien connaître du calcul des variations on pouvait le voir immédiatement. Reprenons une courbe Γ quelconque, mais prenant un système de coordonnées où le plan des xz soit le plan vertical de AB. La projection orthogonale Γ_0 de la courbe dans ce plan est alors définie par les fonctions $x(W), 0, z(W)$; elle passe encore par A et B ; sur cette projection, le temps de chute t_0 est donné par la même intégrale (III,11;28) que t, en remplaçant y' par 0 : il est strictement plus petit que t si $y'(W)$ n'est pas $\equiv 0$, c'est-à-dire si la courbe Γ n'est pas déjà dans le plan $y = 0$. Ainsi le minimum de t ne peut être réalisé que par une courbe dans le plan vertical de AB.

2°/ Nous avons aussi admis d'emblée que la brachistochrone pouvait se représenter en exprimant x en fonction de z.

Cette hypothèse était entièrement injustifiée, comme le montre le résultat final. Ayant en effet obtenu l'équation d'Euler, nous l'avons transforméeen faisant le changement de variable (III,11;25); les solutions de l'équation transformée sont des cycloïdes, oùx ne s'exprime pas nécessairement en fonction de z. Si, en reprenant la construction de la page AB coupe Γ_1 en un seul point C situé sur la première branche montante de la cycloide, alors il y a une solution unique correspondant aux points A et B , et pour cette solution x ne s'exprime pas en fonction de z . Comment remédier à cette situation ? Nous pouvons procéder comme dans 1°), et chercher Γ, dans le plan vertical de AB , comme courbe paramétrique de classe C^2 ,

On obtient les 2 équations d'Euler en x, z :

(III,11;30)
$$\begin{cases} X = -\dfrac{d}{dW}\left(\dfrac{x'}{\sqrt{x'^2+z'^2}}\;\dfrac{1}{\sqrt{2g(z-a)}}\right) = 0 \quad \text{ou} \quad \dfrac{x'}{\sqrt{x'^2+z'^2}}\;\dfrac{1}{\sqrt{2g(z-a)}} = \text{constante}; \\[2ex] Z = -\dfrac{1}{2}\,\dfrac{\sqrt{x'^2+z'^2}}{\sqrt{2g}\,(z-a)^{3/2}} - \dfrac{d}{dW}\left(\dfrac{z'}{\sqrt{x'^2+z'^2}}\;\dfrac{1}{\sqrt{2g(z-a)}}\right) = 0 \end{cases}$$

Mais il existe alors une infinité de solutions, pour A et B donnés, car, les 2 équations (III,11;30) ne sont pas indépendantes. On se trouve en effet dans les conditions de 2°/ page et il existe une combinaison $x'X + z'Z$ des premiers membres des 2 équations (III,11;30) qui est donnée par (III,11;21 sexto) : $\frac{d}{dw}\left(L - x'\frac{\partial L}{\partial x'} - z'\frac{\partial L}{\partial z'}\right)$; comme

L est homogène et de degré 1 en x', z', l'identité d'Euler, $x'\frac{\partial L}{\partial x'} + z'\frac{\partial L}{\partial z'} \equiv L$, montre que cette combinaison

est identiquement nulle. Ceci n'est d'ailleurs pas étonnant : ce qui importe, c'est Γ et non sa paramétrisation; toutes les paramétrisations équivalentes d'une même courbe Γ, donnent le même temps t de chute, et si l'une est extrémale, toutes le sont. On peut obvier à cet inconvénient comme suit. S'il n'est pas possible de supposer que W représente l'abscisse curviligne s pour toutes les courbes, parce que les extrémités A et B doivent correspondre, pour toutes ces courbes, aux mêmes valeurs W_1 et W_2 de W, et que toutes les courbes n'ont pas même longueur, on peut supposer que le paramètre W, sans signification précise pour les courbes étudiées, est l'abscisse curviligne s pour la courber extrémale que l'on cherche. Alors, sans pouvoir remplacer $\sqrt{x'^2 + z'^2}$ par 1 dans

(III,11;28), on peut le faire dans (III,11;30). Alors on obtient 2 équations simplifiées; toute solution de la première annule aussi la combinaison $x'X + z'Z$ donc

$z'Z$, et par suite vérifie aussi la seconde, sauf si c'est une solution Z = constante (et effectivement z = constante, $x' = 1$, est solution de la première, mais non de la deuxième); ainsi (III,11;30) est équivalent à

$$\frac{\frac{dx}{ds}}{\sqrt{2g(z-a)}} = \text{constante,} \tag{III,11;31}$$

les solutions étrangères Z = constante devant être écartées; c'est exactement (III,11;23), et cela redonne les cycloides par le changement de variable approprié. Ainsi, malgré une fausse hypothèse au début, le résultat final est correct, le changement de variable (III,11;25) ayant, en cours de route rétabli la situation.

En supposant x fonction de z, on ne laisse échapper que des courbes z = constante, pour lesquelles il faut revenir à (III,11;30); ici il n'y en a pas.

3°/ Nous avons dit, après le théorème 40 remarque 1°, que l'existence d'un maximum ou d'un minimum n'était pas assurée. Prenons en effet, pour problème, la minimisation de l'intégrale (III,11;2) par une fonction $x = f(z)$ de classe ... Les seules solutions possibles sont les cycloïdes trouvées. Si alors par A et B ne passe, dans les conditions indiquées à 2° page 367, aucune cycloide permettant d'exprimer x en fonction de z, la minimum cherché n'existe pas; l'intégrale (III,11;2) aura, dans l'ensemble Ω, C E étudié, une borne inférieure > 0, mais pas de minimum.

4°/ A la remarque 3°/ page 369, nous avons dit que les conditions aux limites $y(a) = \alpha$, $y(b) = \beta$, ne donnaient pas nécessairement une solution unique de l'équation différentielle d'Euler. Il peut y avoir 0 solution (nous venons de le voir dans la remarque 4°/), ou un nombre fini ou une infinité (page 372).

5°/ On ne se trouve pas, pour la brachistochrone, dans les conditions d'application du théorème 40, puisque la fonction considérée $L = \frac{\sqrt{1+x'^2}}{\sqrt{2g(z-a)}}$ est ici singulière pour $z = a$, qui est justement ~~une~~ des bornes d'intégration ! C'est la malheureusement un fait assez général en mathématiques; on démontre de très beaux théorèmes généraux, qui ne conviennent guère aux cas particuliers qu'on rencontre.

6°/ Mous avons dit plus haut, que, quand une courbe vérifie l'équation différentielle d'Euler, elle est extremale pour l'intégrale, entre des limites absolument quelconques. Ceci semble contradictoire avec le résultat trouvé, à savoir que la courbe possède nécessairement une tangente verticale au point initial, ce qui d'ailleurs se voit immédiatement sur l'équation différentielle : pour $z = a$, on a $x' = 0$.
L'explication est simple. La courbe est bien extrémale pour l'intégrale

(III,11;33) $$\int_{a_1}^{b_1} \frac{\sqrt{1+x'^2}}{\sqrt{2g(z-a)}}\, dz$$

quelles que soient l'origine a_1 et l'extrémité b_1, mais toujours avec la même fonction à intégrer, c'est-à-dire la même constante a dans $\sqrt{2g(z-a)}$.

Alors on ne résout plus le problème de mécanique considéré, qui supposait qu'au point initial, la vitesse initiale était nulle : la même cycloide, limitée aux extrémités A, et B_1, A_1 n'étant pas le point de rebroussement à tangente verticale, minimise l'intégrale (III, 11: 28), qui est le temps de chute sur la courbe, quand la vitesse initiale en A, est celle $\sqrt{2g(a_1-a)}$, $a_1 > a$, qui correspondrait à une vitesse nulle en A. Ces remarques ne doivent pas avoir un effet démoralisant. Elles montrent seulement qu'en écrivant les equations d'Euler, on a résolu seulement une toute petite partie du problème de maximum ou de minimum pose. Pour résoudre toutes les difficultés, il faut des methodes très puissantes utilisant la théorie des espaces de Hilbert, le théorème 30 bis du chapitre II, la topologie algébrique, la théorie des équations différentielles, etc....

Equation des géodésiques sur une surface

Soit V_n une variété de dimension n sur le corps des réels dans un espace affine euclidien E_N de dimension finie N. Soit $\mathcal{C}$ une courbe paramétrique de classe C' tracée sur V On dit qu'elle est une géodésique entre deux de ses points A et B, si la longueur de l'arc AB de $\mathcal{C}$ est minima * , ou tout au moins stationnaire, parmi tous les arcs de courbe de classe C' joignant A à B sur la variété V. D'après ce que nous avons vu au corollaire du théorème 40, elle est alors géodésique entre deux points quelconques A_1, B_1 de l'arc AB, et finalement on peut appeler géodésique une oourbe qui est géodésique entre 2 quelconques de ses points; elle est une extrémale pour la longueur, solution d'une équation d'Euler.

Soit $\Phi : \mathcal{O} \longrightarrow \Phi(\mathcal{O})$ une carte de V, $\mathcal{O}$ étant un ouvert de $\mathbb{R}^n$. Nous appellerons aussi $u \longrightarrow M(u)$ l'application définissant cette carte. Nous supposerons V de classe C^3, donc F de classe C^3. L'arc élémentaire sur V peut alors s'écrire, comme nous le démontrerons plus tard :

(III,11;34)
$$\begin{cases} ds = \|\overrightarrow{dM}\| = \left\| \sum_{i=1}^{n} \overrightarrow{\frac{\partial M}{\partial u_i}} du_i \right\| \\ = \sqrt{\left(\sum_{i=1}^{n} \overrightarrow{\frac{\partial M}{\partial u_i}} du_i \,\middle|\, \sum_{j=1}^{n} \overrightarrow{\frac{\partial M}{\partial u_j}} du_j \right)} = \sqrt{\sum_{i,j=1,2,\dots n} g_{i,j}(u)\, du_i\, du_j}, \text{ où} \end{cases}$$

(III,11;35)
$$g_{i,j}(u) = \left(\overrightarrow{\frac{\partial M}{\partial u_i}}(u) \,\middle|\, \overrightarrow{\frac{\partial M}{\partial u_j}}(u) \right)$$

* Il n'y a pas de maximum des longueurs des courbes joignant 2 points : la borne supérieure de ces longueurs est $+\infty$.

est une fonction **réelle** de classe C^2 sur $\mathcal{O}$. Une courbe paramétrique de classe C^1 sur V est alors représentée par des fonctions $w \longrightarrow u_i(w)$; la fonction $L = \sqrt{\sum_{i,j} g_{i,j}(u) u'_i \cdot u'_j}$ est de classe C^2, le théorème 40 est applicable, et les équations d'Euler sont alors :

(III,11;36) $$\frac{\partial}{\partial u_k}\left(\sqrt{\sum_{i,j} g_{i,j}(u) u'_i u'_j}\right) - \frac{d}{ds}\left(\frac{\partial}{\partial u'_k}\sqrt{\sum_{i,j} g_{i,j}(u) u'_i u'_j}\right) = 0$$

Dans le cas particulier d'une surface dans un espace euclidien à 3 dimensions, on a, en remplaçant u_1, u_2, par u, v :

(III,11;37) $$\begin{cases} ds = \sqrt{E\,du^2 + 2F\,du\,dv + G\,dv^2}\ , \\ E = \left\| \dfrac{\partial \vec{M}}{\partial u} \right\|^2,\ F = \left(\dfrac{\partial \vec{M}}{\partial u} \,\middle|\, \dfrac{\partial \vec{M}}{\partial v} \right),\ G = \left\| \dfrac{\partial \vec{M}}{\partial v} \right\|^2 ; \end{cases}$$

si on représente une courbe de V en représentant v en fonction de u , l'équation d'Euler s'écrit :

(III,11;38) $$\frac{\partial}{\partial v}\sqrt{E(u,v) + 2F(u,v)v' + G(u,v)v'^2} - \frac{d}{du}\left(\frac{F(u,v) + G(u,v)v'}{\sqrt{E(u,v) + 2F(u,v)v' + G(u,v)v'^2}}\right) = 0$$

C'est une équation différentielle du 2ème ordre, sa solution dépend de 2 constantes arbitraires. Si on cherche un arc de géodésique joignant 2 points A et B , on aura 2 équations pour calculer ces 2 constantes.

En fait, il peut arriver qu'il ne passe par deux points A et B aucune geodesique. Si par exemple V est, dans $\mathbb{R}^2$, l'ouvert $x^2 + y^2 > 1$, il ne passe aucune géodésique par les 2 points $A = (2,0)$, $B = (-2,0)$ (si on remplace V par son adhérence $\overline{V}$, qui est l'ensemble $x^2 + y^2 \geq 1$, il y a bien un arc de courbe paramétrique de classe C^1 joignant A à B dans $\overline{V}$ et réalisant le minimum de la longueur, et même 2 arcs symétriques par rapport à AB , donnant la même longueur minima : chacun de ces arcs est formé de la juxtaposition d'un segment de tangente AC issu de A à la circonférence $x^2 + y^2 = 1$, d'un arc CD de cette circonférence et d'un segment de tangente DB issu de B . Mais de telles courbes sont dans $\overline{V}$ et non dans V , et $\overline{V}$ n'est plus une variété; de ce fait le théorème 40 n'est plus applicable parce qu'on n'a plus affaire avec un ouvert de E_0 mais un fermé; et l'arc CD de la circonférence ne satisfait pas à l'équation d'Euler, dont les solutions sont les droites de $\mathbb{R}^2$. On peut montrer, si la variété V est compacte, ou plus généralement complète pour la métrique euclidienne, que 2 de ses points peuvent toujours être joints par au moins une géodésique, donnant le minimum absolu de la longueur; en général il y aura alors une infinité de géodésiques joignant A à B, et on montre que celles qui ne donnent pas un minimum de la longueur donnent tout au moins un minimum relatif.

Prenons, par exemple, un cylindre de révolution de $\mathbb{R}^3$ d'équation

(III,11;39) $$x^2 + y^2 = a^2 .$$

On peut le "développer" sur un plan: cela veut dire qu'on considère une représentation paramétrique du cylindre, conservant les longueurs. C'est une application H du plan $\mathbb{R}^2$ sur le cylindre, définie par $(u,v) \longrightarrow (x,y,z)$,

(III,11;40) $$\begin{cases} x = a \cos \dfrac{u}{a} \\ y = a \sin \dfrac{u}{a} \\ z = v \end{cases}$$

L'application H conserve bien les longueurs, en ce sens que $ds^2 = dx^2 + dy^2 + dz^2 = du^2 + dv^2$; alors la longueur d'un arc de courbe de classe C^1 de $\mathbb{R}^2$ est égale à celle de son image par H, et les géodésiques du cylindre sont les images par H des géodésiques de $\mathbb{R}^2$, c'est-à-dire des droites.

Mais H n'est pas une bijection. Si A est un point du cylindre, son image réciproque est un ensemble de points

$A_n = A_0 + 2\pi a n \vec{i}$, $\vec{i}$ étant le vecteur unitaire de l'axe Ox dans $\mathbb{R}^2$. Si donc A et B sont 2 points du cylindre, tout segment de droite $A_p B_q$, $p \in \mathbb{Z}$, $q \in \mathbb{Z}$ aura pour image un arc de géodésique joignant A à B. Comme $A_p B_q$ et $A_{p+k} B_{q+k}$ ont même image, on pourra se borner à prendre tous les segments de droite $A_0 B_n$, $n \in \mathbb{Z}$. Cela donne quand même une infinité d'arcs géodésiques de classe C^1 joignant A à B. Un seul d'entre eux (exceptionnellement 2 symétriques) donne le minimum de la longueur; les autres donnent tous un minimum relatif. En particulier, si $B = A$, on trouve une infinité de géodésiques d'origine et d'extrémité A. L'une a la longueur 0 et donne le minimum de la longueur. Celle qui est l'image de **A, A,** a pour longueur $2\pi a|n|$; elle est un minimum relatif, et même elle, donne le minimum absolu des longueurs des arcs de classe *C*

joignant A a lui-même "après avoir tourné algébriquement n fois autour de Oz ". On voit ici intervenir une notion de topologie algébrique, le nombre algébrique de tours d'une courbe fermée autour d'un axe.

Donnons un exemple de resolution de l'équation d'Euler des géodésiques. Si la surface est, dans $\mathbb{R}^3$, une surface de révolution autour de l'axe Oz représentée paramétriquement avec r et φ (coordonnées semi-polaires) par $Z = F(r)$, son arcs élémentaire est défini par la formule :

(III,11;41)
$$ds = \sqrt{dr^2 + r^2 d\varphi^2 + F'^2(r)\, dr^2}\ .$$

Dans ces conditions, si l'on cherche à représenter une géodésique en prenant φ comme fonction de r , la longueur d'un arc de courbe AB s'écrit sous la forme

(III,11;42)
$$\int_{r(A)}^{r(B)} \sqrt{(1 + F'^2(r)) + r^2 \varphi'^2}\, dr$$

et l'équation différentielle d'Euler s'écrit

(III,11;43)
$$\frac{r^2 \varphi'}{\sqrt{1 + F'^2(r) + r^2 \varphi'^2}} = \text{constante} = k\ ;$$

elle se résout par la quadrature :

(III,11;44)
$$\varphi = \int \frac{k\sqrt{1 + F'^2(r)}}{r\sqrt{r^2 - k^2}}\, dr + \varphi_0\ .$$

L'équation (III,11;41) s'écrit aussi

(III,11;45)
$$r^2 \frac{d\varphi}{ds} = k \text{ ou } r\left(\frac{r\, d\varphi}{ds}\right) = k \text{ ou } r \cos \Phi = k\ ,$$

Φ étant l'angle de la tangente au parallèle (orientée dans le sens des φ croissants) avec la tangente à la géodésique (orientée dans le sens des arcs croissants). On peut le comprendre géométriquement comme suit. Si un point décrit la géodésique avec la vitesse 1, le moment de sa vitesse par rapport à Oz a pour dérivée par rapport au temps le moment de son accélération; or celle-ci est dirigée suivant la normale principale à la courbe, puisque la vitesse est 1, donc suivant la normale à la surface à cause de la propriété géométrique des géodésiques que nous verrons plus loin, donc elle rencontre 1 axe Oz et son moment est nul. Ainsi le moment de la vitesse suivant Oz est constant, et c'est ce qu'exprime (III,11; 43).

Problèmes d'extrema liés

Théorème 41 - Les notations étant les mêmes qu'au théorème 40, soient M_i, $i = 1,2,\dots m$, des fonctions ayant les mêmes propriétés que L, et k_i des nombres réels donnes. Soient K_i les fonctions sur Ω_0 définies par :

(III,11,46) $$f \longrightarrow K_i(f) = \int_a^b M_i(x, f(x), \vec{f}'(x))\, dx .$$

Supposons qu'en $f_0 \in \Omega_0$ les dérivées $K_i'(f_0)$ soient indépendantes, et que $K_i(f_0) = k_i$. Pour que f_0 rende J maxima ou minima sur Ω_0, parmi toutes les fonctions f de Ω_0 pour lesquelles $K_i(f) = k_i$, i $1,2,\dots m$, il est nécessaire qu'il existe des nombres réels λ_i, $i = 1,2,\dots m$ tels que l'on ait l'équation d'Euler :

(III,11;47) $$\frac{\overleftarrow{\partial L}}{\partial y} - \frac{d}{dx}\left(\frac{\overleftarrow{\partial L}}{\partial y'}\right) = \sum_{i=1}^{m} \lambda_i \left(\frac{\overleftarrow{\partial M_i}}{\partial y} - \frac{d}{dx}\left(\frac{\overleftarrow{\partial M_i}}{\partial y'}\right)\right)$$

Il suffit d'appliquer la théorie des multiplicateurs de Lagrange (Thdorème 34).

SI F a la dimension n, ceci représente un système de-n équations différentielles du second ordre, dont la solution, si les λ_i étalent connus, dépendrait de $2n$ constantes arbitraires. Mais les λ_i ne sont pas connus; la solution générale de ce.système dépend donc de $2n + m$ constantes arbitraires. Les constantes doivent alors être déterminées de manière à satisfaire aux $2n$ conditions aux limites, et aux m conditions $K_i = k_i$, $i = 1,2,\dots m$.

Exemple Parmi tous les arcs de courbe de classe C^1 joignant A à B et de longueur donnée ℓ, trouver celui qui délimite avec le segment AB une aire maxima.

Cherchons cet arc de courbe, en exprimant y en fonction de x, et en prenant AB comme axe des x.

la longueur est alors donnée par l'intégrale :

(III,11;48) $$\ell = \int_{x(A)}^{x(B)} \sqrt{1+y'^2}\,dx$$

tandis que l'aire * est donnée par l'intégrale

(III,11;49) $$S = \int_{x(A)}^{x(B)} y\,dx \,.$$

Il doit alors exister un multiplicateur de Lagrange λ tel que l'on ait l'équation d'Euler :

(III,11;50) $$-\lambda \frac{d}{dx}\left(\frac{y'}{\sqrt{1+y'^2}}\right) = 1 \quad \text{ou} \quad -\frac{\lambda y''}{(1+y'^2)^{3/2}} = 1$$

La courbure est constante : la courbe est un arc de circonférence. Son centre et son rayon sont inconnus, elle dépend donc de 3 constantes arbitraires. En écrivant que l'arc de circonférence cherché passe par A et B, est au dessus de l'axe des x ($y \geq 0$), et admet la longueur ℓ donnée, on voit qu'on le détermine d'une manière unique. Si $\ell > \frac{\pi}{2}|AB|$, on voit que l'arc de circonférence dépasse la demi-circonférence, et que le cercle ne peut pas se représenter en exprimant y en fonction de x. Voir à ce sujet ce que nous avons dit page 367. On peut naturellement supposer que A et B sont confondus, et l'on obtient alors l'énoncé suivant:

<u>Parmi toutes les courbes de classe C^1, fermées, de longueur ℓ, celle qui embrasse l'aire maxima est une circonférence.</u>

Le rayon de cette circonférence est alors $\frac{\ell}{2\pi}$, et par conséquent son aire S est donnée par la formule $S = \frac{\ell^2}{4\pi}$, ce qui permet de dire que, <u>pour n'importe quelle courbe de classe C^1, fermée, de longueur ℓ, l'aire qu'elle embrasse satisfait à l'inégalité</u> :

(III,11;51) $$S \leq \frac{\ell^2}{4\pi} \,,$$

<u>avec l'inégalité stricte, sauf si la courbe est un cercle. Inversement, si une courbe de classe C^1, fermée, embrasse l'aire S, sa longueur satisfait à l'inégalité</u> :

(III,11;52) $$\ell \geq 2\sqrt{\pi S} \,.$$

* C'est là une "aire algébrique" et non une aire; on peut se borner aux cas où $y \geq 0$, pour avoir une aire usuelle.

avec l'inégalité stricte, sauf si la courbe est un cercle.

Encore une fois répétons que nous n'avons pas résolu le problème de façon rigoureuse . Ainsi les inégalités (III,11;49) et (III;11;50) ne sont pas vraiment justifiées.

Effet d'un changement de variables

soit F_1 un espace affine norme, h une application de classe C' de $[a,b]\times F_1$ dans F . Si alors f_1 est une fonction de classe C' sur $[a,b]$ à valeurs dans F_1 , la composée f:

(III,11;53) $$x \longrightarrow f(x) = h(x, f_1(x))$$

est une fonction de classe C^1 sur $[a,b]$ à valeurs dans F . Ainsi h définit une application $f_1 \to f$ de $E_1 = (F_1^{[a,b]})_{cb;1}$ dans $E = (F^{[a,b]})_{cb;1}$.

Nous allons voir que cette application est dérivable, si h est de classe C^2 .

Théorème 42 - L'application $f_1 \longrightarrow f$ définie par (III,11;53) de $E_1 = (F_1^{[a,b]})_{cb;1}$ dans $E = (F^{[a,b]})_{cb;1}$, où h est une application de classe C^2 de $[a,b]\times F_1$ dans F , est dérivable, et sa dérivée au point f_1 de $(F_1^{[a,b]})_{cb;1}$ est donnée par

(III,11;54) $$\begin{cases} \delta f_1 \longrightarrow \delta f \text{ , avec} \\ \overrightarrow{\delta f}(x) = \partial_2 h(x, f_1(x)) \cdot \overrightarrow{\delta f_1}(x) \end{cases}$$

Démonstration - On doit démontrer 2 choses :

I°/ La formule (III,11;54) définit (pour f_1 fixée) une application linéaire continue $\overrightarrow{\delta f_1} \longrightarrow \overrightarrow{\delta f}$ de $\vec{E_1}$ dans $\vec{E}$. Comme h est de classe C^2, $\partial_2 h$ est de classe C^1 , et le corollaire 5 du théorème 11 d'une part, le-théorème 12 d'autre part, montrent sans difficulté que, si δf_1 est dans $\vec{E_1}$, $\overrightarrow{\delta f}$ est dans $\vec{E}$. L'application $\overrightarrow{\delta f_1} \longrightarrow \overrightarrow{\delta f}$ est trivialement linéaire; sa continuité résulte des majorations

(III,11;55) $$\|\|\delta f\|\|_0 \leqslant \left(\operatorname{Sup}_{a\leqslant x\leqslant b} \|\partial_2 h(x, f_1(x))\|\right) \|\|\overrightarrow{\delta f_1}\|\|_0$$

$$\|\overrightarrow{\delta f}'\|_0 \leqslant \operatorname{Sup}_{a\leqslant x\leqslant b} \|\partial_1\partial_2 h(x,f_1(x)) + \partial_2^2 h(x, f_1(x))\cdot \vec{f_1}'(x)\| \, \|\|\overrightarrow{\delta f_1}\|\|_0$$
$$+ \operatorname{Sup}_{a\leqslant x\leqslant b} \|\partial_2 h(x, f_1(x))\| \, \|\|\overrightarrow{\delta f_1}'\|\|_0$$

qui sont de la forme :

(III,11;56) $$\left\|\left| \overrightarrow{\delta f} \right\|\right|_1 \leq \text{constante} \times \left\|\left| \overrightarrow{\delta f_1} \right\|\right|_1$$

2°/ Pour un accroissement $\overrightarrow{\delta f_1}$ de f_1, l'accroissement vrai $\overrightarrow{\Delta f}$ de $\vec{f}$ est donné par :

(III,11;57) $$\overrightarrow{\Delta f}(x) = h(x, f_1(x) + \overrightarrow{\delta f_1}(x)) - h(x, f_1(x)) ;$$

il est normal de l'approcher par $\overrightarrow{\delta f}$ donné par (III,11;54) Alors, $\overrightarrow{\delta f}$ étant défini par (III,11;54), on calculera $\overrightarrow{\Delta f} - \overrightarrow{\delta f}$, ainsi que $\overrightarrow{\Delta f}' - \overrightarrow{\delta f}'$; on utilisera, pour chacune de ces 2 expressions, la îormule des accroissements finis pour x fixé, comme dans la démonstration du théorème 38, et on utilisera également le théorème de la continuité uniforme amélioré (note (**) page 355). On montrera ainsi que, $\varepsilon > 0$ étant donné, on peut trouver $\eta > 0$ de manière que $\left\|\left| \overrightarrow{\delta f_1} \right\|\right| \leq \eta$ entraîne $\left\|\left| \overrightarrow{\Delta f} - \overrightarrow{\delta f} \right\|\right|_1 \leq \varepsilon \left\|\left| \overrightarrow{\delta f_1} \right\|\right|$; cela montrera que $\left\|\left| \overrightarrow{\Delta f} - \overrightarrow{\delta f} \right\|\right|_1$ est infiniment petit devant $\left\|\left| \overrightarrow{\delta f_1} \right\|\right|$, quand celui-ci tend vers 0, et cela signifiera bien que $\overrightarrow{\delta f}$ est la différentielle cherchée.

Ce théorème se généralise comme suit :

<u>Si h est de classe C^{m+k} sur $[a,b] \times F_1$, l'application $f_1 \to f$ est de classe C^k de $(F_1^{[a,b]})_{cb;m}$ dans $(F^{[a,b]})_{cb;m}$.</u>

Application aux géodésiques

soit V_n une variété de dimension n d'un espace affine <u>euclidien</u> E_N de dimension N. Considérons des arcs de courbe paramétriques de classe C^1 joignant 2 points A, B, tracés ou non sur V. On supposera un tel arc défini par une application $w \to M(w)$ d'un segment fixe $[a,b]$ de $\mathbb{R}$ dans E_N. Si on appelle x_i les coordonnées de M sur une base orthonormale de E_N, les x_i sont des fonctions de w, et la longueur de l'arc AB d'une courbe $\mathscr{C}$ est donnée par

(III,11;58) $$\ell = \int_a^b \sqrt{\sum x_i'^2}\; dw .$$

Définissons une courbe $\mathscr{C}_0$ par la fonction $M(w)$, et une courbe voisine $\mathscr{C}$ par la fonction $M(w) + \delta M(w)$. La différentielle $\delta\ell$ de la longueur dans les conditions du théorème 39, est donnée par :

(III,11;59) $$\delta\ell = \sum_{i=1}^{N} \int_a^b - \frac{d}{dw}\left(\frac{x_i'}{\sqrt{\sum x_i'^2}}\right) \delta x_i(w)\, dw .$$

Mais les $\dfrac{x_i'}{\sqrt{\sum_i x_i'^2}}$ sont les cosinus directeurs de la demi-tangente à $\mathscr{C}_0$ dans le sens des arcs croissants (ou w croissants). Soit $\vec{t}(w)$ le vecteur unitaire de cette demi-tangente. La formule (III,11;59) s'écrit donc :

(III,11;60) $$\delta\ell = - \int_a^b \left(\frac{d\vec{t}}{dw} \middle| \vec{\delta M}(w) \right) dw .$$

Mais on a la formule de Frénet, sur $\mathscr{C}_0$:

(III,11;61) $$\frac{d\vec{t}}{ds} = \frac{\vec{n}}{R} .$$

Cette formule, étudiée en Mathématiques Spéciales dans un espace euclidien affine à 3 dimensions, est valable pour une dimension N quelconque : l'égalité $(\vec{t}\,|\,\vec{t}) = 1$ donne, par dérivation, $\left(\frac{d\vec{t}}{ds}\middle|\vec{t}\right) = 0$, donc $\frac{d\vec{t}}{ds}$ est un vecteur normal à la courbe; on peut donc l'appeler $\frac{\vec{n}}{R}$. Alors $\vec{n}$ est un vecteur unitaire indiquantune direction, direction appelée normale principale: $R > 0$ est le rayon de courbure, par définition, et la considération de l'image sphérique montre qu'il a bien la signification habituelle $R = \lim \frac{ds}{d\theta}$; le plan $(\vec{t},\vec{n})$ (a 2 dimensions) s'appelle toujours plan osculateur.

Ainsi :

(III,11;62) $$\delta\ell = - \int_a^b \left(\frac{\vec{n}}{R} \middle| \vec{\delta M}\right) \frac{ds}{dw}\, dw = - \int_{s(a)}^{s(b)} \left(\frac{\vec{n}}{R} \middle| \vec{\delta M}\right) ds .$$

Mais Ici la courbe $\mathscr{C}$ est tracée sur la variété V . Donc $\vec{\delta M}$ n'est pas une variation arbitraire de la fonction $w \longrightarrow M(w)$. Considérons un arc de la géodésique assez petit pour se trouver dans l'image d'une carte $\mathcal{O} \xrightarrow{\Phi} \Phi(\mathcal{O})$ de V , en reprenant les notations de la page 370 . Une courbe de $\mathcal{O}$ sera définie à partir d'une application $w \to u(w)$ de $[a,b]$ dans $\mathcal{O}$, et la courbe image dans V sera définie par la fonction composée $w \longrightarrow M(u(w))$. Le théorème 42 dit que, si V est de classe C^2 , la différentielle $\overrightarrow{\delta M}$ correspondant à une variation $\overrightarrow{\delta u}$ de la courbe tracée dans $\mathcal{O}$ est donnée par :

$$\text{(III,11;63)} \qquad \overrightarrow{\delta M}(w) = M'(u(w)) \cdot \overrightarrow{\delta u}(w) = \sum_{j=1}^{n} \frac{\overrightarrow{\partial M}}{\partial u_j}(u(w))\, \delta u_j(w).$$

Mais alors le théorème des fonctions composées (théorème 11) nous dit que la différentielle $\delta\ell$ par rapport aux accroissements δu_j s'obtient en prenant sa différentielle (III,11; 62), et en y remplaçant l'accroissement $\overrightarrow{\delta M}$ par sa différentielle (III,11;63) par rapport aux δu_j . Alors la différentielle cherchée $\delta\ell$ s écrit :

$$\text{(III,11;64)} \qquad \delta\ell = -\int_{s(a)}^{s(b)} \sum_{j=1}^{n} \left(\frac{\vec{n}}{R} \,\Big|\, \frac{\overrightarrow{\partial M}}{\partial u_j}\right) \delta u_j \, ds ,$$

et cette fois les δu_j sont des fonctions scalaires arbitraires $w \longrightarrow \delta u_j(w)$ sur $[a,b]$, de classe C^1 , nulles aux extrémités a et b ; ou encore $\overrightarrow{\delta u} = (\delta u_1, \delta u_2, \ldots, \delta u_n)$ est une fonction arbitraire sur $[a.b]$, à valeurs dans $\mathbb{R}^n$, de classe C' , nulle en a et b .

Le théorème 40 (ou le lemme de Haar) montre alors qu'une géodésique de classe C^2 de V est caractérisée par les équations d Euler

$$\text{(III,11;65)} \qquad \left(\vec{n} \,\Big/\, \frac{\overrightarrow{\partial M}}{\partial u_j}\right) = 0 , \qquad j = 1,2,\ldots,n :$$

ces équations expriment que la normale principale à la géodésique est, en chacun de ses points, la normale à la variété, ou que le plan osculateur à la géodésique est normal à la variété.

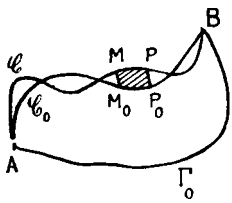
Pour une courbe $\mathcal{C}_0$ quelconque sur V, on laisse souvent $\delta\ell$ sous la forme (III,11;62) avec seulement une légère modification. Appelons γ l'angle de la normale principale à $\mathcal{C}_0$ avec la normale à la surface, compté entre 0 et π . Alors, si nous appelons $\vec{n}_\gamma$ le vecteur unitaire de la projection orthogonale de la normale principale sur l'hyperplan tangent, la projection de $\vec{n}$ est $\vec{n}_\gamma \sin\gamma$; comme $\overrightarrow{\delta M}$ est, en chaque point de la courbe $\mathscr{C}_0$ un vecteur de l'hyperplan tangent, $\sum_{j=1}^{n} \frac{\overrightarrow{\partial M}}{\partial u_j} \delta u_j$, son produit scalaire avec $\vec{n}$ est égal à son produit scalaire avec $\vec{n}_\gamma \sin\gamma$. D'où

$$\delta \ell = -\int_{s(a)}^{s(b)} \left(\frac{\vec{n}_\gamma}{R_\gamma} \middle| \overrightarrow{\delta M} \right) ds \,, \tag{III,11;66}$$

où $R_\gamma = \frac{R}{\sin\gamma}$ est le rayon de courbure géodésique de $\mathscr{C}_0$. C'est le rayon de courbure de la projection ortnogonale de la courbe sur l'hyperplan tangent. Les géodésiques sont les courbes de rayon de courbure géodésique infini, ou de courbure geodesique nulle : $\frac{1}{R_\gamma} = 0$.

Proposons-nous alors, sur une surface V_2 d'un espace affine euclidien à 3 dimensions, de trouver un arc de courbe $\mathscr{C}_0$, de longueur dnnnée ℓ , joignant 2 points donnés A, B et embrassant, avec un arc donné Γ_0 joignant A à B, une aire maxima.

Si $\mathscr{C}$ est représentée en fonction d'un paramètre w , prenant des valeurs données en A et B , nous connaissons la différentielle $\delta\ell$ quand on passe de $\mathscr{C}_0$ à une courbe, voisine $\mathscr{C}$, correspondant à l'accroissement $\overrightarrow{\delta M}(w)$; la différentielle δS de l'aire considérée est la partie principale linéaire de l'aire "algébrique" comprise entre $\mathscr{C}_0$ et $\mathscr{C}$. Faisons une évaluation intuitive de δS . Soient M_0 et P_0 deux points voisins de $\mathscr{C}_0$, de paramètres w_0 et $w_0 + dw$, l'arc M_0P_0 valant ds . Par le déplacement de la courbe, ils sont venus en M et P respectivement , $M = M_0 + \overrightarrow{\delta M}(w_0)$, $P = P_0 + \overrightarrow{\delta M}(w_0 + dw)$.



Si on assimile M_0P_0PM à un "parallélogramme", son aire est le produit de ds par la projection de $\overrightarrow{\delta M}$ sur la normale en M_0 à $\mathscr{C}_0$ située dans le plan tangent à la surface; le vecteur unitaire de cette normale est, au signe près éventuellement, ce que nous avons appelé $\vec{n}_\gamma$.

L'aire du parallèlogramme est donc $\pm\left(\vec{n}_\gamma \,/\, \overrightarrow{\delta M}\right) ds$, et la variation d'aire δS cherchée est donc

$$\delta S = \int_{s(a)}^{s(b)} \pm \left(\vec{n}_\gamma \mid \overrightarrow{\delta M}\right) ds, \tag{III,11;67}$$

le signe $\pm$ étant, en chaque point de $\mathscr{C}_0$, le signe $+$ ou le signe $-$ selon que $\mathscr{C}_0$ tourne sa convexité vers Γ_0 ou non (la normale $\vec{n}_\gamma$ étant toujours dans le sens de la concavité).

La théorie des multiplicateurs de Lagrange (théorème 41) nous indique alors qu'il doit exister une constante réelle λ telle que $\delta S - \lambda\,\delta \ell$ soit nulle en $\mathscr{C}_0$, où

$$\int_{s(a)}^{s(b)} \left(\pm 1 - \frac{\lambda}{R_\gamma}\right)\left(\vec{n}_\gamma \mid \overrightarrow{\delta M}\right) ds = 0. \tag{III,11;68}$$

Naturellement $\overrightarrow{\delta M}$ n'est pas un accroissement arbitraire; si (u,v) sont des paramètres définissant une carte de la surface, on a (formule (III,11,63))

$$\overrightarrow{\delta M} = \frac{\overrightarrow{\partial M}}{\partial u}\,\delta u + \frac{\overrightarrow{\partial M}}{\partial v}\,\delta v, \tag{III,11;69}$$

où δu et δv sont des accroissements arbitraires, nuls en A et B. Les équations d'Euler sont (lemme de Haar) :

$$\left(\pm 1 - \frac{\lambda}{R_\gamma}\right)\left(\vec{n}_\gamma \,\Big|\, \frac{\overrightarrow{\partial M}}{\partial u}\right) = 0, \qquad \left(\pm 1 - \frac{\lambda}{R_\gamma}\right)\left(\vec{n}_\gamma \,\Big|\, \frac{\overrightarrow{\partial M}}{\partial v}\right) = 0. \tag{III,11;70}$$

Comme $\vec{n}_\gamma$ ne peut pas être orthogonal à la fois aux 2 vecteurs indépendants $\frac{\overrightarrow{\partial M}}{\partial u}$, $\frac{\overrightarrow{\partial M}}{\partial v}$, ces équations sont équivalentes à

$$\pm 1 - \frac{\lambda}{R_\gamma} = 0. \tag{III,11;71}$$

Comme $R_\gamma \geqslant 0$, et que λ est constante, le signe $\pm$ est toujours le même, $\mathscr{C}_0$ est une courbe de courbure géodésique constante, ce qui généralise les résultats des pages 375 et suivantes. La courbe fermée de longueur donnée et embrassant l'aire maxima a encore une courbure géodésique constante.

Extrémités variables. Conditions de transversalité

Supposons qu'au lieu de chercher un arc de courbe de classe C' joignant A à B sur une variété V d'un espace affine euclidien E et de longueur minima (recherche des géodésiques), nous cherchions l'arc de courbe joignant un point de $\mathcal{A}$ à un point de $\mathcal{B}$ sur V, et de longueur minima, où $\mathcal{A}$ et $\mathcal{B}$ sont des sous-variétés données de V. C'est un problème plus général que les précédents, parce que les extrémités a et b de l'intervalle d'intégration, et les valeurs α et β de f en ces points, ne sont plus données, mais doivent seulement satisfaire à certaines relations ; J n'est plus seulement fonction de f mais aussi de a et b dont nous supposerons qu'ils décrivent un segment donné $\mathbb{R}_1$ compact de $\mathbb{R}$.

Théorème 43 - Soient $\mathbb{R}_1$ un segment compact de F un espace affine normé sur le corps des réels, $\mathcal{U}$ un ert de $F\times\vec{F}$, L une fonction réelle de classe C^2 sur $\mathbb{R}_1\times\mathcal{U}$. Si alors Ω est l'ouvert de $E=(F^{\mathbb{R}_1})_{c\ell;1}$ formé des f telles que $(f,\vec{f}')$ applique $\mathbb{R}_1$ dans $\mathcal{U}$, alors la fonction J, définie par (III,11;3) sur $\Omega\times\mathbb{R}_1\times\mathbb{R}_1$, est continuement dérivable, et sa dérivée en f_0, a_0, b_0, où f_0 est de classe C^2, est donnée par :

$$(\text{III},11;72)\qquad (\vec{\delta f},\delta a,\delta b)^{*}\longrightarrow \delta J=\int_a^b\left(\frac{\overleftarrow{\partial L}}{\partial y}-\frac{d}{dx}\left(\frac{\overleftarrow{\partial L}}{\partial y'}\right)\right)\cdot\vec{\delta f}\,dx$$

$$+\left(\frac{\overleftarrow{\partial L}}{\partial y'}(b,f_0(b),\vec{f}_0'(b))\cdot\vec{\delta f}(b)+L(b,f_0(b),\vec{f}_0'(b))\,\delta b\right)$$

$$-\left(\frac{\overleftarrow{\partial L}}{\partial y'}(a,f_0(a),\vec{f}_0'(a))\cdot\vec{\delta f}(a)+L(a,f_0(a),\vec{f}_0'(a))\,\delta a\right)$$

* δa et δb sont des nombres réels, on pourrait les appeler da et db. Mais puisque, pour des raisons signalées page , nous avons appelé $\vec{\delta f}$ la variation de f, et par analogie δJ la différentielle de J, nous appelons aussi δa et δb les accroissements de a et b.

<u>Démonstration</u> - J admet une dérivde partielle en f, $\frac{\overleftarrow{\partial J}}{\partial f}$, donnée par (III,11.6) . On a donc :

$$\text{(III,11;73)} \quad \frac{\overleftarrow{\partial J}}{\partial y}(f_0, a_0, b_0)\cdot\overrightarrow{\delta f} = \int_a^b \left(\frac{\overleftarrow{\partial L}}{\partial y}(x, f_0(x), \vec{f}'_0(x)) \cdot \overrightarrow{\delta f}(x) + \frac{\overleftarrow{\partial L}}{\partial y'}(x, f_0(x), \vec{f}'_0(x)) . \overrightarrow{\delta f}'(x) \right) dx$$

Les dérivées partielles de J par rapport à a et b sont bien connues : ce sont les dérivées partielles d'une intégrale définie par rapport aux bornes de l'intervalle d'intégration :

$$\text{(III,11;74)} \quad \begin{cases} \frac{\partial J}{\partial a}(f_0, a_0, b_0) = - L(a, f_0(a), \vec{f}'_0(a)) \\ \frac{\partial J}{\partial b}(f_0, a_0, b_0) = L(b, f_0(b), \vec{f}'_0(b)) \quad (*) \end{cases}$$

Pour en déduire que J est continuement dérivable sur $\Omega \times \mathbb{R}_1 \times \mathbb{R}_1$, il suffit d'appliquer le théorème 15. La continuité partielle de $\frac{\partial J}{\partial f}$ par rapport à f , pour a et b fixés, a été démontrée au théorème 38: en complétant convenablement la démonstration, on voit facilement, et nous l'admettrons, que $\frac{\partial J}{\partial f}$ est continue par rapport à f , a , b .

Les autres continuités sont évidentes. Si f, a, b tendent vers f_0 , a_0 , b_0 , dans E, $\mathbb{R}_1$, $\mathbb{R}_1$, respectivement, alors

$$\text{(III,11;75)} \quad \overrightarrow{f(b) - f_0(b_0)} = \overrightarrow{f(b) - f_0(b)} + \overrightarrow{f_0(b) - f_0(b_0)}$$

tend vers $\vec{0}$ dans $\vec{F}$, à cause de la convergence uniforme de f vers f_0 (impliquée par la convergence de f vers f_0 dans E), pour le premier terme du 2ème membre, et de la continuité de f_0 au point b_0 , pour le deuxième; de même $\vec{f}'(b) - \vec{f}'_0(b_0)$ tend vers $\vec{0}$ à cause de la convergence uniforme de $\vec{f}'$ vers $\vec{f}'_0$ et

* a, b, J (f, a, b) sont des nombres réels, donc $\frac{\partial J}{bd}$ et $\frac{\partial J}{\partial b}$ sont des dérivées usuelles d'une fonction réelle d'une variable réelle.

de la continuité de $\vec{f}'_0$ en b_0 ; donc $L(b, f(b), \vec{f}'(b)) - L(b_0, f_0(b_0), \vec{f}'_0(b_0))$ converge vers 0 à cause de la continuité de L au point $(b_0, f_0(b_0), \vec{f}'_0(b_0))$ de $(\mathbb{R}_1 \times F \times \vec{F})$; et ceci démontre la continuité de $\frac{\partial J}{\partial b}$ au point (f_0, a_0, b_0) de $\Omega \times \mathbb{R}_1 \times \mathbb{R}_1$. La continuité de $\frac{\partial J}{\partial a}$ se démontre de même.

On a donc bien une différentielle totale δJ définie, en (f_0, a_0, b_0), par :

$$\delta J = \left(\int_{a_0}^{b_0}\left(\frac{\vec{\partial} L}{\partial y} \cdot \vec{\delta f} + \frac{\overleftarrow{\partial L}}{\partial y'} \cdot \vec{\delta f'}\right) dx\right) + L(b_0, f_0(b_0), \vec{f}'_0(b_0))\,\delta b \tag{III,11;76}$$

$$- L(a_0, f_0(a_0), \vec{f}'_0(a_0))\,\delta a\,.$$

Si alors on effectue l'intégration par parties (III,11;13) dans l'intégrale du 2ème membre, comme cette fois $\vec{\delta f}$ n'est pas nécessairement nulle en a, et b_0, on obtient (III,11;72).

Nous allons modifier cette expression en faisant intervenir, au lieu de $\vec{\delta f}(a_0)$, $\vec{\delta f}(b_0)$, les différentielles $\vec{\delta \alpha}$, $\vec{\delta \beta}$, des valeurs de f en a et b.

<u>Théorème 44</u> - <u>Soient $\mathbb{R}_1$ un intervalle de $\mathbb{R}$</u>, <u>F un espace affine normé</u>, ξ <u>l'application de</u> $E = (F^{\mathbb{R}_1})_{cb;1} \times \mathbb{R}_1$ <u>dans</u> F <u>définie</u> par :

$$\xi(f, x) = f(x) \in F\,, \qquad \text{valeur de } f \text{ en } s \; * . \tag{III,11;77}$$

<u>Elle est continuement dérivable,</u> et <u>sa dérivée est donnée par</u> :

$$\xi'(f_0, x_0) \cdot (\vec{\delta f}, \vec{\delta x_0}) = \delta(f(x)) = \vec{\delta f}(x_0) + \vec{f}'(x_0) \cdot \delta x_0 \in \vec{F}\,. \tag{III,11;78}$$

* La fonction5 ressemble à l'application bilinéaire canonique $(u, \vec{X}) \longrightarrow u \cdot \vec{X}$ de $\mathcal{L}(\vec{E} \times \vec{F}) \times \vec{E}$ dans $\vec{F}$.

Démonstration - Pour x fixé, $f \longrightarrow f(x) \in F$ est une application affine continue de E dans F snn application linéaire associée étant $\vec{f} \longrightarrow \vec{f}(x)$ Donc (théorème 8 bis) ξ a une dérivée partielle par rapport à f, donnée par :

$$\frac{\partial \xi}{\partial f}(f,x)\cdot\overrightarrow{\delta f} = \overrightarrow{\delta f}(x) \in \vec{F} \tag{III;11;79}$$

Cette dérivée partielle dépend continuement de f, x, car la formule des accroissements finis (théorème 13)

$$\left\|\overrightarrow{\delta f}(x) - \overrightarrow{\delta f}(x_0)\right\| \leqslant \left\|\left|\overrightarrow{\delta f}'\right|\right\|_0 |x-x_0| \leqslant \left\|\left|\overrightarrow{\delta f}\right|\right\|_1 |x-x_0| \tag{III,11;80}$$

donne

$$\left\|\frac{\partial \xi}{\partial f}(f,x) - \frac{\partial \xi}{\partial f}(f_0,x_0)\right\| = \sup_{\left\|\left|\overrightarrow{\delta f}\right|\right\|_1 \leqslant 1}\left\|\left(\frac{\partial \xi}{\partial f}(f,x) - \frac{\partial \xi}{\partial f}(f_0,x_0)\right)\cdot\overrightarrow{\delta f}\right\| \tag{III,11;81}$$

$$= \sup_{\left\|\left|\overrightarrow{\delta f}\right|\right\|_1 \leqslant 1}\left\|\overrightarrow{\delta f}(x) - \overrightarrow{\delta f}(x_0)\right\| \leqslant |x-x_0|,$$

qui converge bien vers 0 quand x tend vers x_0. Pour f fixé, $x \longrightarrow f(x)$ n'est autre que la fonction f elle-même; elle est bien dérivable puisque f est de classe C^1, donc ξ a une dérivée partielle par rapport à x, donnée par :

$$\overrightarrow{\frac{\partial \xi}{\partial x}}(f,x) = \overrightarrow{f'}(x) \in \vec{F}\ . \ ^{*} \tag{III,11;82}$$

La fonction dérivée $\overrightarrow{\frac{\partial \xi}{\partial x}}$ dépend continuement de f et de x, car, quand f tend vers f_0 dans $(F^{\mathbb{R}_1})_{c b;1}$ et x vers 2, , $f'(x)$ tend vers $f_0'(x_0)$ (raisonnement

* Puisque x parcourt le corps des réels, nous pouvons prendre le vecteur dérivé $\overrightarrow{\frac{\partial \xi}{\partial x}} = \frac{\partial \xi}{\partial x}\cdot 1$. Voir remarque 2° après le théorèrrie 8. Ici $\overrightarrow{f'(x)}\,dx$ est le $\overrightarrow{df}$ que nous avons voulu éviter en écrivant $\overrightarrow{\delta f}$ page 353 .

déjà fait page 383 , avec b, b_0 au lieu de x, x_0). Ainsi ξ a des dérivées partielles continues, et le théorème 15 donne alors le résultat.

On pourra donc modifier la formule (III,11;72), en y remplaçant $\overrightarrow{\delta f}(b)$ par $\overrightarrow{\delta\beta} - \overrightarrow{f'_0}(b_0)\,\delta b$, $\overrightarrow{\delta f}(a)$ par $\overrightarrow{\delta\alpha} - \overrightarrow{f'_0}(a_0)\,\delta a$, d'après (III,11;78). Alors :

Corollaire 1 - Dans les conditions du théorème 43, si on appelle α, β les valeurs de f en a, b, la dérivée de J au point (f_0, a_0, b_0) est donnée par :

(III,11;84)
$$\delta J = \int_{a_0}^{b_0} \left[\frac{\overleftarrow{\partial L}}{\partial y}\left(x, f_0(x), \overrightarrow{f'_0}(x)\right) - \frac{d}{dx}\left(\frac{\overleftarrow{\partial L}}{\partial y'}\left(x, f_0(x), \overrightarrow{f'_0}(x)\right)\right)\right] \cdot \overrightarrow{\delta f}(x)\,dx$$
$$+\left[\frac{\overleftarrow{\partial L}}{\partial y'}\left(b_0, f_0(b_0), \overrightarrow{f'_0}(b_0)\right) \cdot \left(\overrightarrow{\delta\beta} - \overrightarrow{f'_0}(b_0)\,\delta b\right) + L\left(b_0, f_0(b_0), \overrightarrow{f'_0}(b_0)\right)\delta b\right]$$
$$-\left[\frac{\overleftarrow{\partial L}}{\partial y'}\left(a_0, f_0(a_0), \overrightarrow{f'_0}(a_0)\right) \cdot \left(\overrightarrow{\delta\alpha} - \overrightarrow{f'_0}(a_0)\,\delta a\right) + L\left(a_0, f_0(a_0), \overrightarrow{f'_0}(a_0)\right)\delta a\right]$$

Corollaire 2 - Supposons vérifiées les conditions du théorème 43 - Soient a, b, α, β, des fonctions données d'un paramètre t parcourant un ouvert d'un espace affine normé T , à valeurs dans $\mathbb{R}_1, \mathbb{R}_1$, F, F respectivement. Pour que f_0, t_0, avec $f_0(a(t_0)) = \alpha(t_0)$, $f_0(b(t_0)) = \beta(t_0)$, rendent maxima ou minima l'intégrale $J(f, a(t), b(t))$, parmi tous les t de T et toutes les f de Ω telles que $f(a(t)) = \alpha(t)$, $f(b(t)) = \beta(t)$, il est nécessaire que f_0 soit solution de l'équation d'Euler (III,11;16), et que t_0 vérifie en outre les conditions dites de transversalité, c'est-à-dire annule les 2 crochets de (III,11;84), dans lesquels δa, δb, $\overrightarrow{\delta\alpha}$, $\overrightarrow{\delta\beta}$, sont les différentielles des fonctions a, b, α, β, correspondant à un accroissement arbitraire δt de t à partir de t_0 .

Démonstration - Faisons un raisonnement rapide, sans rigueur. Tout d'abord f_0 doit rendre maxima ou minima $J(f, a(t_0), b(t_0))$, correspondant à des extrémités $a(t_0)$, $b(t_0)$ fixées, et cela parmi toutes les fonctions8 de Ω vérifiant $f(a(t_0)) = \alpha(t_0)$, $f(b(t_0)) = \beta(t_0)$. Donc f_0 doit vérifier l'équation d'Euler (III,11;16) (théorème 14). Alors la différentielle δJ de (III,11;84) ne comprend pas d'intégrale et se réduit aux crochets. Il faut bien remarquer que f et t ne sont pas des variables indépendantes, à cause des relations $f(a(t)) = \alpha(t)$, $f(b(t)) = \beta(t)$. Mais on peut considérer t comme une variable libre; pour t fixé, f est astreinte à certaines relations qui l'obligent à rester dans un sous-espace affine E_t de **E** (variable avec t). Comme t est une variable libre, et que la différentielle δJ de (III,11;84) ne fait intervenir que $\overrightarrow{\delta t}$ (par l'intermédiaire de $\delta a = \overleftarrow{a'}(t_0)\cdot\overrightarrow{\delta t}$, δb, $\overrightarrow{\delta\alpha}$, $\overrightarrow{\delta\beta}$), et non $\overrightarrow{\delta f}$ (parce que f_0 satisfait à l'équation d'Euler), la condition nécessaire d'extrema revient bien à écrire que $SJ = 0$ quel que soit $\overrightarrow{\delta t}$, d'où le corollaire.

Application aux géodésiques

Soit $(w, t) \longrightarrow M(w, t)$ une application de classe C^2, de $[0,1] \times \mathbf{T}$ dans une variété V de classe C^3 d'un espace euclidien affine E_N de dimension finie N; $[0,1]$ est un segment de $\mathbb{R}$, T un ouvert d'un espace affine normé (pratiquement, T pourra être $\mathbb{R}$, et $t \in T$ sera le temps). Pour tout $t \in T$, l'application partielle $w \longrightarrow M(w,t)$ définit un arc de courbe paramétrique $\mathscr{C}_t$ de classe C^2 sur V ; ses extrémites sont $A_t = M(0,t)$ et $B_t = M(1,t)$.

Comment varie sa longueur l_t en fonction de t ? La différentielle δl, correspondant à l'accroissement $\overrightarrow{\delta t}$ à partir de t_0 est donnée par (III,11:84). Mais l'intégrale se ramène'& (III,11;66); d'autre part, ici,, $a(t) \equiv 0$, $b(t) \equiv 1$, donc $\delta a = \delta b = 0$; $d(t) = A_t$, $\beta(t) = B_t$.

Posons $\mathscr{C}_0 = \mathscr{C}_{t_0}$, $A_0 = A_{t_0}$, $B_0 = B_{t_0}$. Si nous prenons une base orthonormale de E_N, on sait que les composantes de $\dfrac{\overleftarrow{\delta L}}{\delta y'}$ sont les $\dfrac{x_i'(w)}{\sqrt{\sum_{i=1}^{N} x_i'^2(w)}}$, $i = 1, 2, \ldots N$, c'est-à-dire les cosinus directeurs $\cos \Phi_i(w)$ de la demi-tangente (sens des w croissants) $\vec{t}$ à la courbe $\mathscr{C}_0$ *, au point w ; si nous appelons α_i, β_i les coordonnées de A, B, le 1er crochet de (III,11;84) vaut

$$\sum_{i=1}^{N} (\cos \Phi_i)_{w=1} \cdot \delta\beta_i = \left(\vec{t}\,(B_0) \middle| \overrightarrow{\delta B}\right) \tag{III,11;85}$$

On a un résultat analogue en A. Finalement :

$$\delta \ell = \ell'(t_0) \cdot \overrightarrow{\delta t} = -\int_{s(A)}^{s(B)} \left(\frac{\vec{n}_\gamma}{R_\gamma} \middle| \overrightarrow{\delta M} \right) ds + \left(\overrightarrow{t\,(B_0)} \middle| \overrightarrow{\delta B}\right) - \left(\overrightarrow{t(A_0)} \middle| \overrightarrow{\delta A}\right). \tag{III,11;86}$$

Dans cette formule, $\overrightarrow{\delta M}$ veut dire $\frac{\partial M}{\partial t}(w, t_0) \cdot \overrightarrow{\delta t}$, tandis que $\overrightarrow{\delta A}$, $\overrightarrow{\delta B}$ correspondent aux valeurs particulières $w = 0$, $w = 1$.

En particulier, si la courbe est une <u>géodésique de V</u> dépendant du paramètre t, ou même simplement si, pour $t = t_0$, $\mathscr{C}_0$ est une géodésique de V, on aura :

$$\delta \ell = \left(\vec{t}\,(B_0) \middle| \overrightarrow{\delta B}\right) - \left(\vec{t}\,(A_0) \middle| \overrightarrow{\delta A}\right). \tag{III,11;87}$$

C'est là un résultat remarquable :

<u>La variation de la longueur de la géodésique vérifie la même formule que la variation de la longueur d'un segment de droite dans l'espace euclidien.</u>

Si par exemple, on recherche, sur V, un arc de courbe de longueur minima joignant un point A de $\mathcal{A}$ à un point B de $\mathcal{B}$, $\mathcal{A}$ et $\mathcal{B}$ sous-variétés données de V (page 382), nous pourrons représenter paramétriquement $\mathcal{A}$ et $\mathcal{B}$, au

* Ne pas confondre le paramètre t avec le vecteur unitaire $\vec{t}$ de la tangente à $\mathscr{C}_0$.

moins localement, à l'aide de paramètres ξ, η, parcourant des ouverts d'espaces affines Ξ, H ; alors T sera $\Xi \times H$, t sera le couple (ξ, η), et le corollaire 2 nous dira que d'abord la courbe doit être une géodésique, et qu'ensuite ses extrémités A et B doivent être choisies sur $\mathcal{A}$ et $\mathcal{B}$ de manière à annuler (III,11;87) quel que soit $\overrightarrow{\delta t}$, c'est-à-dire quels que soient $\overrightarrow{\delta\xi}, \overrightarrow{\delta\eta}$; cela signifie simplement que la géodésique doit être normale en A à $\mathcal{A}$ et en B à $\mathcal{B}$.

Equations canoniques de Hamilton

Considérons un problème de calcul des variations, dans lequel F est l'espace $\mathbb{R}^m$, de sorte que la fonction f est équivalente à un système de m fonctions $f_1, f_2, \dots f_m$, de la variable x. Ecrivons les équations d'Euler, sous la forme d'équations différentielles du premier ordre au nombre de $2m$, en introduisant les fonctions auxiliaires $z_i = y'_i$. On obtient, en considérant L comme une fonction donnée des $2m+1$ variables x, y_i, z_i :

$$\text{(III,11;88)} \quad \begin{cases} y'_i = z_i \quad , \quad i = 1,2,\dots,m ; \\ \dfrac{\partial L}{\partial y_i} - \dfrac{d}{dx}\left(\dfrac{\partial L}{\partial z_i}\right) = 0 \quad , \quad i = 1,2,\dots,m . \end{cases}$$

Effectuons un changement de fonctions inconnues, A la place des fonctions inconnues y_i et z_i, nous allons prendre les nouvelles fonctions inconnues q_i et p_i, $q_i = y_i$, les p_i étant données en fonction de x, des y_i et des z_i, par la formule :

$$\text{(III,11;89)} \quad p_i = \frac{\partial L}{\partial z_i} .$$

Nous admettrons qu'il s'agit là d'un véritable changement de fonctions, c'est-à-dire que la formule (III,11;89) permet de calculer chaque z_i en fonction de x des y_i, et des p_i. D'autre part, au lieu d'utiliser la fonction donnée L de x, des y_i, et des z_i, nous introduirons la fonction H définie par la formule :

$$\text{(III,11;90)} \quad H(x, q_1, q_2, \dots q_m, p_1, p_2, \dots p_m) = \sum_{i=1}^{m} p_i z_i - L$$

H, <u>supposée exprimée en fonction de x , des q_i et des p_i , est ce qu'on appelle l'Hamiltonien</u> * . Sa différentielle s'exprime comme suit (les dérivées partielles de L étant prises par rapport aux x , y_i , z_i) :

$$dH = \sum_{i=1}^{m}\left(p_i\, dz_i + z_i\, dp_i\right) - \frac{\partial L}{\partial x}dx - \sum_{i=1}^{m}\frac{\partial L}{\partial y_i}dy_i - \sum_{i=1}^{m}\frac{\partial L}{\partial z_i}dz_i \tag{III,11;91}$$

$$= \sum_{i=1}^{m} z_i\, dp_i - \sum_{i=1}^{m}\frac{\partial L}{\partial y_i}dq_i - \frac{\partial L}{\partial x}dx ,$$

ceci étant valable aussi- bien dans l'ancien système x, y_i , z_i , que dans le nouveau x, q_i , p_i . Cela prouve que les dérivées partielles de H par rapport aux x, q_i , p_i, sont données par :

$$\frac{\partial H}{\partial x} = -\frac{\partial L}{\partial x}\,^{**} , \quad \frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial y_i} , \quad \frac{\partial H}{\partial p_i} = z_i . \tag{III,11;92}$$

Dans ces conditions, les équations différentielles d'Euler (III,11;88) s'expriment, pour les fonctions q_i et p_i de la variable x, sous la forme remarquable suivante :

$$q_i' = \frac{\partial H}{\partial p_i} , \qquad p_i' = -\frac{\partial H}{\partial q_i} . \tag{III,11;93}$$

Ainsi :

* Alors que L , exprimé en fonction des x, y_i , z_i, est le Lagranglen

** Bien noter que $\frac{\partial L}{\partial x}$ est une dérivée partielle dans le système (x , y_i , z_i) , tandis que $\frac{\partial H}{\partial x}$ est une dérivée partielle dans le système (x, q_i , p_i).

Théorème **45** - SI. dans les conditions du théorème 40, on suppose que F est l'espace $\mathbb{R}^m$, si on fait le changement de variables et de fonctions $(x, y_i, z_i) \longrightarrow (x, q_i, p_i)$ défini par (III,11;89), et si on introduit la fonction H, hamiltonien, par (III,11;90), alors, H étant supposé exprimé par rapport aux x, q_i, p_i, les équations d'Euler pour les fonctions q_i et p_i de la variable x s'écrivent sous la forme (III,11;93). Ces equations sont appelées equationsde Hamilton relatives a la fonction H.

Remaroue - Nous avons pris $F = \mathbb{R}^m$ seulement pour simplifier, mais c'est inutile. Supposons simplement F de dimension finie. On peut toujours, au lieu de y et $\vec{z} = \vec{u}y$, fonctions sur $[a,b]$ à valeurs dans F et $\vec{F}$ respectivement, prendre les nouvelles fonctions inconnues $q = y$ et $\overleftarrow{p} = \frac{\overleftarrow{\partial L}}{\partial \vec{z}}$, fonctions sur $[a,b]$ à valeurs dans F et $\overleftarrow{F}'$ respectivement. L'Hamiltonien H est une fonction réelle sur $[a,b] \times F \times \overleftarrow{F}'$, définie par :

$$(\mathrm{III},11;94) \qquad H(x, q, \overleftarrow{p}) = \langle \overleftarrow{p}, \vec{z} \rangle - L(x, y, \vec{z}),$$

$\langle \overleftarrow{p}, \vec{z} \rangle \in \mathbb{R}$ étant le produit scalaire de $\overleftarrow{p} \in \overleftarrow{F}'$ et $\vec{z} \in \vec{F}$; on suppose y et $\vec{z}$ remplacés par leurs valeurs en fonction de q et $\overleftarrow{p}$.

Les équations de Hamilton, pour les fonctions inconnues q et $\overleftarrow{p}$ de x, sont alors :

$$(\mathrm{III},11;95) \qquad \frac{\overrightarrow{dq}}{dx} = \frac{\overrightarrow{\partial H}}{\partial \overleftarrow{p}}, \quad \frac{d\overleftarrow{p}}{dx} = -\frac{\overleftarrow{\partial H}}{\partial \vec{q}}.$$

Pour $\overleftarrow{p}$ fixé, H est une fonction réelle de $\vec{q} \in \vec{F}$, donc $-\frac{\partial H}{\partial \vec{q}}(x, q, \overleftarrow{p}) \in \mathcal{L}(\vec{F}; \mathbb{R}) = \overleftarrow{F}'$, et on comprend qu'il soit égal à $\frac{d\overleftarrow{p}}{dx}$. De même $\frac{\partial H}{\partial \overleftarrow{p}}(x, q, \overleftarrow{p}) \in \mathcal{L}(\overleftarrow{F}'; \mathbb{R}) = (\overleftarrow{F}')'$ mais on sait que $(\overleftarrow{F}')'$, dual de $\overleftarrow{F}'$, n'est autre que $\vec{F}$, et on comprend encore qu'il soit égal à $\frac{\overrightarrow{dq}}{dx}$. La fonction $\overleftarrow{p}$ est appelée le "moment onjugué" de la fonction 4 ; q prend ses valeurs dans F, p dans $\overleftarrow{F}'$, d'où le nom de 'conjugué".

Supposons en particulier que L ne dépende pas explicitement de x. Dans ces conditions H non plus ne dépend pas explicitement de x, les équations précédentes montrent que la quantité :

$$\frac{dH}{dz} = \sum \frac{\partial H}{\partial q_i} q'_i + \sum \frac{\partial H}{\partial p_i} p'_i \tag{III,11;96}$$

est identiquement nulle; autrement dit, H est une intégrale première pour les courbes extrémales, H est constante le long de l'une quelconque de ces courbes. C'est ce que nous avions Indiqué à (III,21; septimo).

Applications à la Mécanique

Les **équations d'Hamilton** ont des applications extrêmement importantes, à toute la mécanique et à toute la physique thdorlque.

Considérons, pour simplifier, un problème de mécanique à liaisons fixes sans frottement, à champ de forces indépendant du temps et dérivant d'un potentiel. Dans ce problème, la position du système pourra, par exemple, se représenter à l'aide d'un nombre fini de paramètres : $q_1, q_2, \cdots, q_m$. Il existera une énergie potentielle U, qui sera une fonction connue de ces paramètres. D'autre part, il sera possible de calculer l'énergie cinétique $T = \sum \frac{1}{2} m v^2$, qui sera une forme quadratique par rapport aux dérivées premières $q'_1, q'_2, \ldots, q'_m$ $(q'_i = \frac{dq_i}{dt})$, dont les coefficients dépendront eux-mêmes de $q_1, q_2, \cdots, q_m$. Résoudre le problème de mécanique, c'est trouver les "trajectoires"; une trajectoire est définie par des fonctions de t, $t \to q_i(t)$.

On démontre alors que la trajectoire du problème de mécanique est solution d'un problème d'extremum. Si t_1 et t_2 sont deux instants déterminés, il est possible de considérer, sur la trajectoire réelle ou sur n'importe quelle trajectoire 'fictive", l'intégrale :

$$\int_{t_1}^{t_2} L(q_i, q'_i)\, dt = \int_{t_1}^{t_2} \left(T(q_i, q'_i) - U(q_i)\right) dt. \;^{*} \tag{III,11;97}$$

* Nous écrivons $L(q_i, q'_i)$ Pour $L(q_1, q_2, \cdots q_m, q'_1, q'_2, \cdots q'_m)$

On démontre que la trajectoire réelle est, parmi toutes les trajectoires fictives qui, aux instants original t_1 et final t_2 considérés, passent par les mêmes points $q_i(t_1)$, $q_i(t_2)$, celle qui rend stationnaire l'intégrale (III,11;97).

Autrement dit, chaque trajectoire est une **extrémale,** et l'équation qui donne les trajectoires du problème de mécanique **considére** est le système des équations d'Euler :

(III,11;98) $$\left\{\begin{array}{l} \dfrac{\partial L}{\partial q_i} - \dfrac{d}{dt}\left(\dfrac{\partial L}{\partial q'_i}\right) = 0 \;; \qquad i = 1,2,\ldots,m, \text{ ou} \\ \dfrac{\partial T}{\partial q_i} - \dfrac{d}{dt}\left(\dfrac{\partial T}{\partial q'_i}\right) = \dfrac{\partial U}{\partial q_i} \end{array}\right.$$

Elles s'appellent, dans ce cas là, les équations de Lagrange du problème de mécanique; $L = T-U$ s'appelle le Lagrangien. Le changement de variables de Hamilton est alors défini par la formule :

(III,11;99) $$p_i = \frac{\partial L}{\partial q'_i} = \frac{\partial T}{\partial q'_i} .$$

L'**Hamiltonien** H devient :

(III,11;100) $$H(q_i,p_i) = \sum_{i=1}^{m} q'_i \frac{\partial T}{\partial q'_i} - L .$$

Mais, si nous tenons compte de ce que T est une forme **quadratique** par rapport aux q'_i, **l'identité d'Euler** des fonctions homogènes nous donne la formule :

(III,11;101) $$\sum_{i=1}^{m} q'_i \frac{\partial T}{\partial q'_i} = 2T .$$

En exprimant alors H par :

(III,11;102) $$H = 2T - (T-U) = T+U ,$$

on voit **que** H n'est autre que l'énergie du système, (somme de son énergie potentielle et de son énergie cinétique) **exprimée** en fonction des q_i et des p_i. Les équations d Hamilton du problème de mécanique, H étant exprimé en fonction des q_i **et des** p_i, sont alors le système d'équations (III,11;93).

L'Hamiltonien H ne dépend pas du temps; et par conséquent c'est une intégrale première du système, autrement dit, le long d'une trajectoire du système, l'énergie H , somme de l'énergie cinétique et de l'énergie potentielle reste constante, propriété bien connue en mécanique élémentaire.

Calcul des variations relatif à des intégrales multiples

Posons-nous par exemple, le problème suivant. Considérons une courbe compacte $\mathscr{C}$ de classe C^1, dans un espace affine euclidien à trois dimensions sur le corps des réels, et cherchons, parmi toutes les surfaces $\mathscr{S}$ de classe C^1 bordées par cette courbe * , celle qui a l'aire minima (problème des surfaces minima!. Si nous prenons des coordonnées (x, y, z) dans l'espace, et si nous représentons cette surface en supposant z exprimée en fonction de x et y , l'aire de la surface considérée s'exprime sous la forme de l'intégrale double :

$$S = \iint_{\Sigma} \sqrt{1+p^2+q^2}\, dx\, dy\,, \quad p = \frac{\partial z}{\partial x}\,, \quad q = \frac{\partial z}{\partial y}\,; \tag{III,11;103}$$

cette intégrale est étendue à l'aire Σ , projection de $\mathscr{S}$ sur le plan des x, y , aire bordée par Γ , projection de $\mathscr{C}$. Nous avons donc à chercher $z = f(x,y)$, de classe C^1, prenant des valeurs données tout le long du contour Γ de $\mathbb{R}^2$ (pour qued passe par $\mathscr{C}$), de manière a rendre minima l'intégrale précédente.

Considérons, dans l'espace $\mathbb{R}^n$, la région ouverte $\mathcal{O}$ délimitée par une hypersurface compacte Γ de classe C^1 . Soit F un espace affine normé, $\mathcal{U}$ un ouvert de $F \times F^n$. Soit *alors* L une fonction réelle *sur* $\overline{\mathcal{O}} \times \mathcal{U}$, de classe C^2 ** . Nous la noterons $L(x_1, x_2, \dots, x_n\,;\, z\,;\, \vec{p}_1, \vec{p}_2, \dots, \vec{p}_n)$.

* Nous ne chercherons pas à préciser ici le sens de l'expression "surface $\mathscr{S}$ bordée par $\mathscr{C}$" . Nous le définirons au chapitre VI page 159 . Disons seulement que $\mathscr{S}$ et $\mathscr{C}$ sont des variétés au sens du § 9, de dimensions respectives 2 et 1, qu'elles sont sans point commun, que $\overline{\mathscr{S}} = \mathscr{S} \cup \mathscr{C}$.

** $\overline{\mathcal{O}}$ n'est pas un ouvert de $\mathbb{R}^n$, mais un compact. De même qu'on peut définir la dérivée d'une fonction sur un intervalle $[a, b]$ compact de $\mathbb{R}$ (voir début du § 2, page 184), on peut le faire aussi pour une fonction sur le compact $\overline{\mathcal{O}}$ de $\mathbb{R}^n$. Nous justifierons l'espression : "ouvert $\mathcal{O}$ délimité par une hypersurface compacte Γ de classe C^1 " au théorème 28 du chapitre VI.

Alors, si f est une application de classe C^1 de $\bar{\Sigma}$ dans F, elle admet des dérivées partielles $\vec{p}_i = \frac{\partial f}{\partial x_i}$, fonctions sur $\bar{\Sigma}$ a valeurs dans $\vec{F}$. Si l'image de $\bar{\Sigma}$ par $(f, \vec{p}_1, \cdots, \vec{p}_n)$ est dans $\mathcal{U}$, ouvert de $F \times \vec{F}^n$, on peut considérer l'intégrale multiple :

$$(\text{III},11;104) \quad J(f) = \iint\cdots\int_{\mathcal{V}} L\Big(x_1,\ldots,x_n; f(x_1,x_2,\ldots,x_n);$$

$$\vec{p}_1(x_1,\ldots,x_n), \vec{p}_2(x_1,\ldots,x_n), \quad \vec{p}_n(x_1,\ldots,x_n)\Big)\, dx_1 \ldots dx_n$$

$$= \iint\cdots\int_{\mathcal{V}} L(x_i\,;\, f\,;\, p_i)\, dx_1 \ldots dx_n .$$

Le problème proposé est le suivant : parmi toutes les fonctions f, <u>prenant des valeurs données tout le long du contour Γ</u>, quelle est celle qui rend l'intégrale J maxima ou minima ?

Pour résoudre ce problème on fait un raisonnement très analogue a celui que nous avons fait dans le cas des intégrales simples. Voici le résultat.

<u>Théorème 46</u> - <u>La fonction $J : f \longrightarrow Jf$ est de classe C^1</u>, <u>et sa différentielle est donnée par</u> :

$$(\text{III},11;105) \quad \delta J = \iint\cdots\int_{\mathcal{V}} \left(\frac{\overleftarrow{\partial L}}{\partial z}\cdot\overrightarrow{\delta f} + \sum_{i=1}^{n}\frac{\overleftarrow{\partial L}}{\partial p_i}\cdot\overrightarrow{\delta p_i}\right) dx_1\ldots dx_n \quad \left(\delta p_i = \frac{\partial}{\partial x_i}\delta f\right).$$

<u>Si L et f_0 sont de classe C^2, et si on se restreint au sous-espace des f prenant des valeurs données sur le contour Γ, cette différentielle s'écrit</u> :

$$(\text{III},11;106) \quad \delta J = \iint\cdots\int_{\mathcal{V}} \left(\frac{\overleftarrow{\partial L}}{\partial z} - \sum_{i=1}^{n}\frac{d}{dx_i}\left(\frac{\overleftarrow{\partial L}}{\partial p_i}\right)\right)\overrightarrow{\delta f}\, dx_1 \ldots dx_n$$

<u>où</u> $\frac{d}{dx_i}\left(\frac{\overleftarrow{\partial L}}{\partial p_i}\right)$ <u>veut dire la dérivée partielle en x_i de la fonction composée</u>

$$(x_1, x_2, \ldots, x_n) \rightarrow \left(\frac{\overleftarrow{\partial L}}{\partial p_i}\right)\Big(x_1, x_2, \ldots, x_n; f_0(x_1,\ldots,x_n);$$

$$\frac{\partial f}{\partial x_1}(x_1,\ldots,x_n), \ \frac{\partial f}{\partial x_2}(x_1,\ldots,x_n), \ldots, \ \frac{\partial f}{\partial x_n}(x_1,\ldots,x_n)\Big).$$

<u>Pour qu'une fonction f_0 de classe C^2 rende J maxima ou minima, parmi les fonctions f de classe C^1 prenant des valeurs données au contour, il est nécessaire que f_0 satisfasse aux équations aux dérivées partielles du 2ème ordre d'Euler</u> :

$$(\text{III},11;107) \quad \frac{\overleftarrow{\partial L}}{\partial z} - \sum_{i=1}^{n}\frac{d}{dx_i}\left(\frac{\overleftarrow{\partial L}}{\partial p_i}\right) = 0 .$$

Démonstration - Le seul point qui ne soit pas absolument analogue à ce que nous avons vu pour les intégrales simples, est le passage de (III,11; 105) à (III,11; 106).

(III,11;105) s'écrit

$$\delta J = \iint\cdots\int_{\sigma}\left(\overleftarrow{\frac{\delta L}{\delta z}}\cdot\overrightarrow{\delta f}+\sum_{i=1}^{n}\frac{d}{dx_i}\left(\overleftarrow{\frac{\delta L}{\delta p_i}}\cdot\overrightarrow{\delta f}\right)-\sum_{i=1}^{n}\frac{d}{dx_i}\left(\overleftarrow{\frac{\delta L}{\delta p_i}}\right)\cdot\overrightarrow{\delta f}\right)dx_1\ldots dx_n \qquad \text{(III,11;107 bis)}$$

ce qui vaut (III,11;106), si

$$\iint\cdots\int_{\sigma}\sum_{i=1}^{n}\frac{d}{dx_i}\left(\overleftarrow{\frac{\delta L}{\delta p_i}}\cdot\overrightarrow{\delta f}\right)dx_1,\ dx_2\ \ldots\ dx_n = 0 \qquad \text{(III,11;107 ter)}$$

Nous utiliserons une formule qui sera démontrée plus tard, la formule d'Ostrogodsky (VI,7;53). Elle permet de remplacer (III,11;107 ter) intégrale de volumes par une intégrale de surface :

$$\int\cdots\int_{\Gamma}\sum_{i=1}^{n}\left(\overleftarrow{\frac{\delta L}{\delta p_i}}\cdot\overrightarrow{\delta f}\cos\alpha_i\right)dS. \qquad \text{(III,11;107 quarto)}$$

Comme on se restreint au sous-espace des f prenant des valeurs données sur Γ, $\overrightarrow{\delta f}$ est nulle sur Γ, et cette intégrale est bien nulle.

Le passage de la formule (III,11;106) à l'équation d'Euler (III,11;107) exige un lemme de Haar à plusieures variables, qui n'est pas essentiellement plus compliqué que celui que nous avons démontré pour 1 variable.

Reprenons, par exemple, le problème de la recherche des surfaces minima. Une surface minima, où z est exprimé en fonction de x et de y dans $\mathbb{R}^3$, satisfait à l'équation aux dérivées partielles

$$\frac{d}{dx}\left(\frac{p}{\sqrt{1+p^2+q^2}}\right)+\frac{d}{dy}\left(\frac{q}{\sqrt{1+p^2+q^2}}\right)=0\ , \qquad \text{(III,11;108)}$$

qui peut encore s'écrire sous la forme :

$$\frac{r(1+p^2+q^2)-p(pr+qs)}{(1+p^2+q^2)^{\frac{3}{2}}}+\frac{t(1+p^2+q^2)-q(ps+qt)}{(1+p^2+q^2)^{\frac{3}{2}}}=0\ , \text{ ou} \qquad \text{(III,11;109)}$$

$$(r+t)(1+p^2+q^2)-(rp^2+2spq+tq^2)=0 \qquad \text{(III,11;110)}$$

En Introduisant la notion de courbure moyenne, on voit que l'équation précédente est équivalente à $R_1 + R_2 = 0$: une surface minima est une surface de courbure moyenne nulle, c'est-à-dire ayant, en chaque point, des rayons de courbure principaux opposés. Naturellement la résolution rigoureuse de ce problème de calcul des variations est soumise aux mêmes difficultés que nous avons partout vues jusqu'à présent, mais beaucoup plus considérables lorsqu'il s'agit d'un problème d'intégrales multiples. Nous savons bien que, parmi toutes les surfaces bordées par le contour $\mathscr{C}$, il existe une borne inférieure > 0 de l'aire, mais nous ignorons si cette borne inférieure est un minimum, et nous ignorons encore plus, au cas où il s'agirait bien d'un minimum, s'il est atteint par une surface où z puisse se représenter par une fonction de classe C^2 de x et de y ; ce sont là, dans les problèmes d'extréma d'intégrales multiples, des difficultés fondamentales. En admettant même qu'il en soit ainsi, et que la fonction z soit astreinte à vérifier l'équation aux dérivées partielles (III,11;110), il reste encore à trouver la solution de cette équation aux dérivées partielles pour laquelle la surface $\mathscr{S}$ passe par le contour donné $\mathscr{C}$, c'est-à-dire pour laquelle z prenne des valeurs données tout le long de la projection Γ du contour $\mathscr{C}$ sur le plan des x, y. Le problème de la recherche de la surface minima ayant pour bord une courbe donnée, s'appelle problème de Plateau. Il a été résolu par le mathématicien Douglas. La solution est beaucoup trop compliquée pour que nous puissions en parler ici.

Donnons un autre problème : $\mathcal{O}$ étant l'ouvert de $\mathbb{R}^n$ délimité par l'hypersurface Γ de classe C^1, trouver la fonction réelle f de classe C^1 sur $\overline{\mathcal{O}}$, prenant des valeurs données sur le contour Γ, et rendant minima l'intégrale multiple (dite intégrale de Dirichlet) :

$$\text{(III,11;111)} \qquad \iint\cdots\int_{\mathcal{O}} \left(\sum_{i=1}^{n} \left|\frac{\partial f}{\partial x_i}\right|^2\right) dx_1 \ldots dx_n = \iint\cdots\int_{\mathcal{O}} \left(\sum_{i=1}^{n} p_i^2\right) dx_1 \ldots dx_n$$

Ici f doit être (si elle est de classe C^2) une solution de l'équation d'Euler :

$$\text{(III,11;112)} \qquad \sum_{i=1}^{n} \frac{\partial p_i}{\partial x_i} = 0$$

c'est-à-dire de l'équation de Laplace

$$\text{(III,11;113)} \qquad \Delta f = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2} = 0 .$$

On dit encore que f doit être une fonction harmonique.

On est donc amené à résoudre ce que l'on appelle le problème de Dirichlet : trouver une fonction f de classe C^2 dans $\overline{\mathcal{O}}$, harmonique, et prenant des valeurs données sur le contour Γ . Il existe de nombreux problèmes de physique où l'on est amené à résoudre ce problème de Dirichlet; dans ces problèmes l'intégrale (III,11;111) a une interprétation simple comme représentant une énergie, et on est amené a rechercher une fonction réalisant un certain équilibre comme celle qui minimise une énergie. Au 19ème siècle, Riemann avait cru résoudre le problème de Dirichlet précisément par cette méthode, en démontrant a priori l'existence d'un minimum de l'intégrale.

Malheureusement la méthode de Riemann comportait une erreur; d'une part, rien ne démontre l'existence d'un minimum, comme nous l'avons indiqué page 361, 1°, et les remarques additionnelles de Riemann pour prouver cette existence étaient basées sur un faux théorème de compacité * ; d'autre part, comme nous l'avons dit page 361, 2°, rien ne prouve que le minimum, s'il existe, soit réalise par une fonction f de classe C^2 Il y a là en réalité des difficultes assez considérables: comme nous l'avons maintenant déjà signalé un grand nombre de fois; nous montrerons ultérieument, dans l'étude des fonctions harmoniques, comment les difficultés rencontrées dans la méthode de Riemann peuvent être soulevées, et comment on peut effectivement résoudre, par une méthode de ce type, le problème de Dirichlet.

* La méthode de Riemann revenait à supposer que la boule unité d'un Banach de dimension infinie était compacte, comme dans le cas de dimension finie; nous avons vu que ce n'est jamais vrai (théorème 45 bis du chapitre II).

IV
CALCUL INTÉGRAL

CALCUL INTEGRAL

§ 1 INTÉGRALE DE RIEMANN SUR LA DROITE

Soient? un espace de Banach sur le corps $\mathbb{K}$ des réels ou des complexes, $\vec{f}$ une fonction définie sur un intervalle $[a, b]$ de la droite réelle $\mathbb{R}$, à valeurs dans $\vec{F}$. Nous nous proposons de définir son intégrale :

$$(\text{IV}, 1;1) \qquad \int_{[a,b]} \vec{f}(x)\, dx \in \vec{F} .$$

On voit pourquoi il sera nécessaire de supposer que $\vec{f}$ prend ses valeurs dans un espace de Banach. Cn sait en effet qu'un intégrale peut être considérée comme une limite de sommes finies, du type

$$(\text{IV}, 1;2) \qquad \sum_{i=0}^{n-1} (c_{i+1} - c_i)\, \vec{f}(\xi_i) ,$$

où $c_0, c_1, c_2, \ldots, c_n$ est une suite croissante de $n+1$ points de l'intervalle $[a, b]$, avec $c_0 = a$ et $c_n = b$, $\xi_i \in [c_i, c_{i+1}]$. Or, pour pouvoir considérer une telle somme, il faut d'abord pouvoir considérer chaque terme $(c_{i+1} - c_i)\, \vec{f}(\xi_i)$, et pour cela il faut savoir faire le produit d'un élément $\vec{f}(\xi_i)$ de $\vec{F}$ par un scalaire réel $(c_{i+1} - c_i)$; ensuite il faut pouvoir considérer une somme de tels éléments de $\vec{F}$, donc $\vec{F}$ doit être normalement un espace vectoriel sur le corps des réels. D'autre part, l'intégrale n'est pas une gomme, mais une limite de sommes, il r'aut donc pouvoir dans $\vec{F}$ considérer des limites, et il est normal de supposer que $\vec{F}$ est un espace vectoriel normé sur le corps des réels. Théoriquement cela doit suffire pour pouvoir définir une intégrale. Mais il ne sera pas possible de trouver des critères pratiques et utilisables d'Intégrabilité si l'on ne suppose pas que $\vec{F}$ est complet. C'est en effet seulement dans ce cas, qu'on peut démontrer l'existence de la limite d'une suite sans connaître à l'avance cette limite. C'est pourquoi dans la suite, sauf mention expresse du contraire, nous supposerons que $\vec{F}$ est un espace de Banach. Naturellement il peut être un espace de Banach sur le corps des complexes, puisqu'alors on peut a fortiori le considérer comme espace de Banach sur le corps des réels.

Nous supposerons toujours la fonction $\vec{f}$ bornée; autrement dit, lorsque x parcourt $[a,b]$, $\|\vec{f}(x)\|$ reste majorée par un nombre fixe. Nous désignerons d'ailleurs par $\|\vec{f}\|$, comme nous l'avons déjà fait page 139, la fonction $\geqslant 0$ définie sur $[a,b] : x \Rightarrow \|\vec{f}(x)\|$, tandis que nous représenterons par $|||\vec{f}|||$ la borne supérieure de cette fonction *.

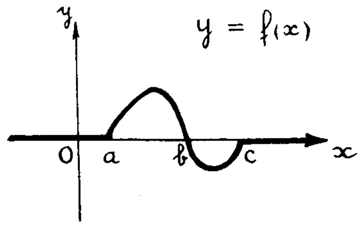
D'autre part, nous considérerons toujours des fonctions $\vec{f}$ définies sur toute la droite, et nous les intégrerons sur toute la droite, bien que ces fonctions doivent toujours être nulles en dehors d'un intervalle borné.

Nous écrirons donc des expressions :

$$(\text{IV},1;3) \qquad \int_{\mathbb{R}} \vec{f}(x)\,dx \quad \text{ou simplement} \int \vec{f}(x)\,dx \quad \text{ou} \int \vec{f} \ .$$

Par définition, le symbole (IV,1;1) est alors l'intégrale sur toute la droite de la fonction $\vec{f}$ définie sur toute la droite, égale à $\vec{f}$ dans l'intervalle $[a,b]$ et à $\vec{0}$ dans le complémentaire de cet intervalle. On appelle support d'une fonction $\vec{f}$, définie sur un espace topologique X, à valeurs dans un espace vectoriel $\vec{F}$, l'adhérence de l'ensemble des points x, où $\vec{f}(x) \neq \vec{0}$. D'après cette définition même, le support d'une fonction est toujours un ensemble fermé. Le support de $\vec{f}$ est le plus petit ensemble fermé de X, sur le complémentaire duquel $\vec{f}$ soit $\equiv \vec{0}$. Par exemple, pour la fonction réelle sur $\mathbb{R}$ dont le graphique est représenté sur la figure, l'ensemble des points où $f(x) \neq 0$ est l'intervalle $]a,c[$ ouvert, dont on retranche le point b ; mais le support est l'intervalle fermé $[a, c]$.

* Cependant, si $\vec{F}$ est le corps des scalaires lui-même, $\|\ \|$ est remplacé par $|\ |$, et on peut alors remplacer $|||\ \ |||$ par $\|\ \ \|$



Si f est la fonction réelle, égale à 0 en tous les points d'abcisse irrationnelle de $\mathbb{R}$, et à 1 en tous les points d'abcisse rationnelle, l'ensemble des points où elle est différente de 0 est l'ensemble Q des nombres rationnels et le support est la droite réelle $\mathbb{R}$ toute entière. Si $\vec{f}$ et $\vec{g}$ sont deux fonctions sur X à valeurs dans l'espace vectoriel $\vec{F}$, alors le support de $\vec{f}+\vec{g}$ est évidemment contenu dans la réunion des supports de $\vec{f}$ et $\vec{g}$; en effet, un point x où $\vec{f}(x)+\vec{g}(x)\neq 0$ appartient ou bien à l'ensemble A des points où $\vec{f}\neq\vec{0}$, ou bien à l'ensemble B des points où $\vec{g}\neq\vec{0}$, ces deux éventualités n'étant pas exclusives l'une de l'autre; il appartient donc à $A\cup B$, et par suite, le support est contenu dans $\overline{A\cup B}$, lui-même contenu dans la réunion $\bar{A}\cup\bar{B}$ des supports. Nous intégrerons donc sur $\mathbb{R}$ des fonctions $\vec{f}$, définies sur $\mathbb{R}$, mais à support compact.

Fonctions en escalier

On dit qu'une fonction f définie sur la droite réelle $\mathbb{R}$, à valeurs dans un ensemble quelconque F, est une fonction en escalier, s'il existe une suite croissante finie de points $c_0, c_1, c_2, \ldots, c_n$, de $\mathbb{R}$, tels que, dans chacun des intervalles ouverts $]-\infty, c_0[$, $]c_0, c_1[$, ..., $]c_{n-1}, c_n[$, $]c_n, +\infty[$, la fonction f soit une constante. (Aucune hypothèse supplémentaire n'est faite sur les valeurs de la fonction aux points c_i eux-mêmes) Une telle suite de points c_i s'appelle une décomposition ou subdivision A de $\mathbb{R}$, admissible pour la fonction en escalier f. Naturellement il existe une décomposition relative à f, qui est la "meilleure possible", c'est-à-dire dans laquelle le nombre des points c_i est le plus petit possible, mais on ne s'attachera pas spécialement à celle-la. Alors, pour une même fonction en escalier f, il existe une infinité de décompositions admissibles de $\mathbb{R}$.

On dit qu'une décomposition Δ' de $\mathbb{R}$ est plus fine qu'une décomposition Δ, si la suite des c'_i contient la suite des c_i ; toute décomposition plus fine qu'une décomposition admissible pour $\vec{f}$ est encore admissible pour $\vec{f}$. Etant données deux décompositions quelconques Δ' et Δ'' de $\mathbb{R}$, il existe toujours au moins une décomposition Δ plus fine que chacune d'elles; elle s'obtient en réunissant les 2 suites de points de subdivision relatifs à Δ' et à Δ'' et en les rangeant par ordre de grandeur croissante. Si on modifie en un nombre fini de points de $\mathbb{R}$ la valeur d'une fonction en escalier, elle reste une fonction en escalier. Si $\vec{F}$ est un espace vectoriel, le produit d'une fonction en escalier par un scalaire, et la somme de deux fonctions en escalier, sont encore des fonctions en escalier. Ce dernier point seul n'est pas absolument évident; si $\vec{f}$ et $\vec{g}$ sont deux fonctions en escalier, on remarque qu'il leur correspond deux décompositions Δ', Δ'', de $\mathbb{R}$, et par conséquent on peut trouver une décomposition Δ plus fine que chacune d'elles. Alors Δ est une décomposition admissible pour chacune des deux fonctions en escalier $\vec{f}$ et $\vec{g}$, et il devient évident que la somme $\vec{f}+\vec{g}$ est aussi en escalier, et admet Δ comme décomposition.

Nous définirons alors comme suit <u>l'intégrale de Riemann d'une fonction en escalier à support compact,</u> à valeurs dans un espace vectoriel $\vec{F}$ sur $\mathbb{K}$, non nécessairement normé :

$$\int \vec{f} = \sum_{i=0}^{n-1} (c_{i+1} - c_i)\, \vec{f}(\xi_i)\,, \qquad \xi_i \in]c_i\,, c_{i+1}[\ , \tag{IV,1;4}$$

où $c_0, c_1, c_2, \ldots, c_n$ est une quelconque décomposition de $\mathbb{R}$ admissible pour la fonction en escalier *

Il faut montrer que nous avons bien là une définition correcte, autrement dit que le deuxième membre de (IV,1;4) est indépendant de la décomposition choisie. Soient alors Δ' et Δ'' deux décompositions possibles pour $\vec{f}$, et appelons Δ une décomposition plus fine que chacune d'elles. Il est alors immédiat de voir que les expressions (IV,1;4) relatives à Δ' et Δ'' donnent toutes deux le même résultat que celle qui est relative à Δ.

* Comme $\vec{f}$ est supposée à support compact, elle est nécessairement nulle dans $]-\infty, c_0[$ et dans $]c_n, +\infty[$. Noter que les valeurs de $\vec{f}$ aux points c_i n'interviennent pas.

(Si en effet $]c'_i, c'_{i+1}[$ est un intervalle de la décomposition $\mathcal{A}'$ alors il est subdivisé par Δ en un certain nombre d'intervalles $]d_j, d_{j+1}[,]d_{j+1}, d_{j+2}[, \ldots,]d_{k-1}, d_k[$. Dans chacun de ces intervalles, comme dans l'intervalle $]c'_i, c'_{i+1}[$, la fonction $\vec{f}$ est une même constante $\vec{f}_i$, et par conséquent on a l'égalité :

$$(\text{IV},1;5) \qquad \sum_{\ell=j,j+1,\ldots,k-1} (d_{\ell+1} - d_\ell)\vec{f}_i = (c'_{i+1} - c'_i)\vec{f}_i .$$

En sommant ce qui correspond aux divers intervalles de $\mathcal{A}'$, on somme exactement ce qui correspond aux divers intervalles de Δ, ce qui prouve bien l'identité des sommes correspondant à $\mathcal{A}'$ et Δ).

Si λ est un scalaire, et si $\vec{f}$ et $\vec{g}$ sont des fonctions en escalier à support compact, on a manifestement la formule :

$$(\text{IV},1;6) \qquad \int \lambda \vec{f} = \lambda \int \vec{f} , \quad \int (\vec{f} + \vec{g}) = \int \vec{f} + \int \vec{g} .$$

C'est évident pour la multiplication par un scalaire; pour le cas de la somme, il suffit de choisir une décomposition Δ commune à $\vec{f}$ et $\vec{g}$, et par conséquent aussi à $\vec{f} + \vec{g}$. On voit donc :

Théorème 1 - <u>L'ensemble des fonctions en escalier à support compact, définies sur $\mathbb{R}$, à valeurs dans l'espace vectoriel $\vec{F}$ sur le corps $\mathbb{K}$ des reels ou des complexes, est lui-même un espace vectoriel sur $\mathbb{K}$, et l'intégrale $\vec{f} \to \int \vec{f}$ est une application linéaire de cet espace vectoriel dans $\vec{F}$. Si $\vec{F} = \mathbb{R}$, et si $f \geq 0$, alors $\int f \geq 0$. Si en outre $\vec{F}$ est un espace vectoriel normé, on a les deux majorations :</u>

$$(\text{IV},1;7) \qquad \left\| \int \vec{f} \right\| \leq \int \| \vec{f} \| = \int \| \vec{f}(x) \| \, dx , \quad \left\| \int_{[a,b]} \vec{f} \right\| \leq (b-a) \sup_{a \leq x \leq b} \| \vec{f}(x) \| .$$

<u>Si on modifie en un nombre fini de points de $\mathbb{R}$ la valeur d'une fonction en escalier à support compact, elle reste en escalier a support compact, et son intégrale n'est pas modifiée.</u>

Le début du théorème a été démontré avant l'énoncé. La majoration (IV,1;7) est évidente sur une décomposition Δ admissible pour $\vec{f}$: $\left\| \sum_{i=0}^{n-1} (c_{i+1} - c_i)\vec{f}_i \right\| \leq \sum_{i=0}^{n-1} (c_{i+1} - c_i) \| \vec{f}_i \|$ $\leq (b-a) \sup_i \| \vec{f}_i \|$ si $c_0 = a$, $c_n = b$.

Si maintenant $\vec{f}$ et $\vec{g}$ sont 2 fonctions en escalier à support compact qui ne diffèrent qu'en un nombre fini de points de $\mathbb{R}$, ces points sont nécessairement parmi les c_i d'une décomposition Δ admissible à la fois pour $\vec{f}$ et $\vec{g}$, et l'intégrale ne fait pas intervenir les valeurs de la fonction aux points c_i de la décomposition Δ.

C'est maintenant que, pour définir l'intégrale d'une fonction quelconque, il nous sera nécessaire d'effectuer un passage à la limite.

Intégrale supérieure de Riemann d'une fonction $f \geq 0$, bornée, à support compact

Définition. Soit f une fonction réelle, ≥ 0, bornée, à support compact, définie sur $\mathbb{R}$.

On appelle intégrale supérieure de Riemann de f, et on note $\int^* f$, la borne inférieure des intégrales des fonctions f_1 en escalier à support compact, qui majorent f; autrement dit, on peut écrire :

$$\int^* f = \operatorname*{Inf}_{\substack{f_1 \text{ en escalier} \\ f_1 \geq f}} \left(\int f_1 \right) \geq 0 . \qquad \text{(IV,1;8)}$$

On voit pourquoi nous sommes obligés de faire les restrictions précédentes, c'est-à-dire de supposer f bornée et à support compact; s'il n'en était pas ainsi, il n'existerait pas de fonction en escalier à support compact majorant f. Naturellement, si f est en escalier, on peut, pour f_1, prendre f elle-même, et $\int^* f = \int f$.

Théorème 2 - Si $f \leq g$, alors $\int^* f \leq \int^* g$; si λ est un scalaire ≥ 0, et si f et g sont deux fonctions positives, bornées, à support compact, on a les formules :

$$\int^* \lambda f = \lambda \int^* f, \quad \int^* (f+g) \leq \int^* f + \int^* g . \qquad \text{(IV,1;9)}$$

On dit encore que l'intégrale supérieure des fonctions positives possède la propriété de convexité. L'intégrale supérieure d'une fonction ne change pas quand on change sa valeur en un nombre fini de points de $\mathbb{R}$.

Démonstration - Tout est à peu près évident. Démontrons par exemple la propriété de convexité.

Soient f_1 et g_1 des fonctions en escalier, bornées, à support compact, majorant respectivement f et g, alors $f_1 + g_1$ majore $f + g$, on a donc :

$$\int^*(f+g) \leq \int(f_1+g_1) = \int f_1 + \int g_1 . \tag{IV,1;10}$$

Par suite le premier membre est aussi majoré par la borne inférieure du troisième qui n'est autre que $\int^* f + \int^* g$, et ceci démontre (IV,1;9).

Si maintenant f et g sont égales sauf en un nombre fini de points de $\mathbb{R}$, et si f_1 est en escalier à support compact et majore f, on peut trouver g_1 en escalier à support compact et majorant g, égale à f_1 sauf au plus en un nombre fini de points. Alors $\int^* g \leq \int g_1 = \int f_1$; en prenant la borne Inférieure du 3ème membre, on a $\int^* g \leq \int^* f$, d'où, par un raisonnement symétrique, l'inégalité en sens inverse et par suite l'égalité.

Remarques 1°/ On pourrait aussi définir l'intégrale inférieure de Riemann $\int_* f$ comme la borne supérieure des intégrales des fonctions en escalier majorées par f ; mais cette intégrale inférieure posséderait alors une propriété de concavité, à savoir :

$$\int_*(f+g) \geq \int_* f + \int_* g \; ; \tag{IV,1;11}$$

pratiquement elle n'est guère utilisée.

2°/ On pourrait croire que l'on a même $\int^*(f+g) = \int^* f + \int^* g$. Il n'en est rien.

2

Considérons en effet la fonction f nulle dans le complémentaire de $[0,1]$, et en tous les points irrationnels de $[0,1]$, égale à 1 en tous les points rationnels de $[0,1]$. Si f_1 est une fonction en escalier $\geq f$, dans tout intervalle $]c_i, c_{i+1}[$ d'une décomposition Δ admissible pour f_1, il existe des points rationnels, donc la valeur constante de f_1 est sûrement ≥ 1 ; alors f_1 est au moins égale à la fonction caractéristique de l'intervalle $[0,1]$. Réciproquement cette dernière est en **escalier** et majore f. Alors $\int^* f = 1$. Si maintenant nous appelons g la fonction analogue à f obtenue en échangeant les rôles des nombres rationnels et irrationnels, on a aussi $\int^* g = 1$. Cependant $f + g$ est la fonction caractéristique de $[0,1]$, et $\int^*(f+g) = \int(f+g) = 1 < 1+1$.

Dans cet exemple, on voit que, si l'on introduit les intégrales inférieures de la remarque 1°/, $\int_* f = \int_* g = 0$, et $\int_*(f+g) = \int(f+g) = 1 > 0 + 0$.

3°/ En modifiant une fonction en une infinité dénombrable de points de $\mathbb{R}$,on peut modifier son intégrale supérieure.

Par exemple la fonction f de la remarque 2°/ diffère de la fonction 0 aux points rationnels, et son intégrale supérieure est 1, et non **0** .

4°/ L'inégalité de ronvexité s'étend naturellement à une somme d'un nombre fini de fonctions; <u>mais non a la somme d'une infinité dénombrable</u>. Autrement dit, si $f_0, f_1, f_2, \ldots, f_n, \ldots$ est une suite de fonctions ≥ 0 bornées, à support dans un même intervalle $[a,b]$ de $\mathbb{R}$, et si la série $\sum_{n=0}^{\infty} f_n(x)$ converge, pour tout x , vers une limite $f(x)$, f étant elle-même bornée (et a support évidemment dans $[a, b]$) , <u>on n'a pas nécessairement</u> $\int^* f \leq \sum_{n=0}^{\infty} \int^* f_n$.

Prenons par exemple l'ensemble des nombres rationnels de $[0,1]$. Ils peuvent être rangés en une suite $a_0, a_1, a_2, \ldots, a_n, \ldots$. Appelons f_n la fonction caractéristique de $\{a_n\}$. Alors $f = \sum_{n=0}^{\infty} f_n$ n'est autre que la fonction définie dans la remarque 2°/ . Or $\int^* f = 1 > \sum_{n=0}^{\infty} \int^* f_n = 0$.

Nous introduirons, dans la théorie de l'intégrale de Lebesgue, une autre intégrale supérieure, qui, elle, possède la propriété de convexité dénombrable $\int^* \sum_{n=0}^{\infty} f_n \leq \sum_{n=0}^{\infty} \int^* f_n$, et ce sera là la source essentielle de la supériorité de l'intégrale de Lebesgue sur l'intégrale de Riemann.

Fonctions intégrables à valeurs dans un espace de Banach.

Définition - Soit $\vec{f}$ une fonction définie sur la droite réelle $\mathbb{R}$, à valeurs dans un espace de Banach $\vec{F}$. On dit que $\vec{f}$ est intégrable-Riemann, si elle est bornée, à support compact, et si, quel que soit $\varepsilon > 0$, il existe une fonction $\vec{g}$, à valeurs dans $\vec{F}$, en escalier à support compact, telle que :

$$\int^* \|\vec{f} - \vec{g}\| \leq \varepsilon . \tag{IV,1;12}$$

Il est équivalent de dire qu'il existe une suite $\vec{f}_0, \vec{f}_1, \ldots, \vec{f}_n, \ldots$ de fonctions à valeurs dans $\vec{F}$, en escalier à support compact, telles que les intégrales supérieures $\int^* \|\vec{f} - \vec{f}_n\|$ convergent vers 0 pour n tendant vers l'infini: Une telle suite de fonctions s'appelle suite d'approximation de $\vec{f}$ pour l'intégrale de Riemann *.

Si deux fonctions ne diffèrent qu'en un nombre fini de points, et si l'une est intégrable, l'autre l'est aussi (avec les mêmes suites d'approximation).

Remarques 1°/ Soit $\vec{f}$ une fonction définie sur $\mathbb{R}$, à valeurs dans l'espace de Banach $\vec{F}$ bornée et à support compact, et, supposons que, quel que soit $\varepsilon > 0$, il existe une fonction $\vec{g}$ intégrable-Riemann, telle que $\int^* \|\vec{f} - \vec{g}\| \leq \varepsilon$.

Alors $\vec{f}$ elle-même est intégrable.

* Il ne faudrait pas croire qu'une telle suite $\vec{f}_n$ converge simplement vers $\vec{f}$. Nous donnerons, page 500, un exemple, dans lequel $\vec{f}_n(x)$ ne converge vers $\vec{f}(x)$ pour aucune valeur de x.

En effet, $\varepsilon > 0$ étant donne, nous pouvons, d'abord choisir une fonction $\vec{g}$ intégrable, telle que $\int^* \|\vec{f} - \vec{g}\| \leq \frac{\varepsilon}{2}$; mais, $\vec{g}$ étant intégrable, nous pouvons choisir une fonction $\vec{h}$ en escalier, à support compact, telle que $\int^* \|\vec{g} - \vec{h}\| \leq \frac{\varepsilon}{2}$. Alors on a, d'après l'inégalité de convexité (IV,1;9), $\int^* \|\vec{f} - \vec{h}\| \leq \varepsilon$, et ceci, d'après la définition, nous montre bien que $\vec{f}$ est intégrable.

De même, si une suite de fonctions $\vec{f}_0, \vec{f}_1, \dots, \vec{f}_n, \dots$ sont toutes intégrables, et si $\int^* \|\vec{f} - \vec{f}_n\|$ tend vers 0, (en particulier si les $\vec{f}_n$ gardent toutes leur support dans un même compact, et convergent uniformément vers $\vec{f}$) alors $\vec{f}$ est aussi intégrable. On dira, ici encore, que la suite des $\vec{f}_n$ quoique non en escalier, est une suite d'approximation de $\vec{f}$ pour l'intégrale.

2°/ Si $\vec{f}$ est une fonction intégrable, alors, quel que soit $\varepsilon > 0$ il existe une fonction $\vec{g}$ en escalier, à support compact: à valeurs dans $\vec{F}$, et une fonction h en escalier à support compact, à valeurs réelles ≥ 0 , telles que $\|\vec{f} - \vec{g}\| \leq h$, et que $\int h \leq \varepsilon$; et réciproquement.

La réciproque est évidente, démontrons l'affirmation directe.

En effet, nous pouvons d'abord trouver une fonction en escalier $\vec{g}$, à support compact, telle que $\int^* \|\vec{f} - \vec{g}\| \leq \frac{\varepsilon}{2}$; mais alors, d'après la définition même de l'intégrale supérieure, nous pouvons trouver une fonction h en escalier à support compact, réelle ≥ 0 , telle que $\|\vec{f} - \vec{g}\| \leq h$ et que $\int h \leq \varepsilon$.

3°/ Si f est une fonction réelle définie sur $\mathbb{R}$, intégrable-Riemann, alors, quel que soit $\varepsilon > 0$, on peut trouver deux fonctions en escalier à support compact g_1 et g_2 , telles que $g_1 \leq f \leq g_2$, et que $\int g_2 - g_1 \leq \varepsilon$; et réciproquement.

La réciproque est encore évidente; montrons l'affirmation directe. Soit f intégrable. Déterminons, conformément à la remarque 2°/, une fonction réelle en escalier g, et une fonction en escalier $h \geq 0$, toutes deux à support compact, telles que $|f - g| \leq h$, et $\int h \leq \frac{\varepsilon}{2}$.

si alors, nous posons $g_1 = g - h$ et $g_2 = g + h$, g_1 et g_2 répondent à la question.

Intégrale d'une fonction intégrable

Théorème 3 - Soit $\vec{f}$ une fonction intégrable-Riemann à valeurs dans un espace de Banach $\vec{F}$, et soit $\vec{f}_0, \vec{f}_1, \dots, \vec{f}_n, \dots$ une suite d'approximation de $\vec{f}$ par des fonctions en escalier. Alors les quantités $\int \vec{f}_n \in \vec{F}$ ont une limite dans $\vec{F}$, et cette limite est indépendante de la suite d'approximation considérée.

Démonstration - On a la majoration :

$$\left\| \int \vec{f}_m - \int \vec{f}_n \right\| \leq \int \left\| \vec{f}_m - \vec{f}_n \right\| \leq \int^* \left\| \vec{f} - \vec{f}_m \right\| + \int^* \left\| \vec{f} - \vec{f}_n \right\| . \tag{IV,1;13}$$

Cela prouve que, lorsque m et n tendent vers l'infini, $\left\| \int \vec{f}_m - \int \vec{f}_n \right\|$ tend vers $\vec{0}$. Mais alors la suite des vecteurs $\int \vec{f}_n$ est une suite de Cauchy dans $\vec{F}$; comme $\vec{F}$ est complet *, elle a bien une limite $\vec{L}$. Reste à voir que cette limite est indépendante de la suite d'approximation considérée. Or, si nous considérons deux telles suites d'approximation $\vec{f}_0, \vec{f}_1, \dots, \vec{f}_n, \dots$ et $\vec{g}_0, \vec{g}_1, \dots, \vec{g}_n, \dots$, on pourra considérer la suite $\vec{f}_0, \vec{g}_0, \vec{f}_1, \vec{g}_1, \dots, \vec{f}_n, \vec{g}_n, \dots$, qui est encore une suite d'approximation de $\vec{f}$. La suite des quantités $\int \vec{f}_0, \int \vec{g}_0, \int \vec{f}_1, \int \vec{g}_1, \dots, \int \vec{f}_n, \int \vec{g}_n, \dots$, a donc une limite, autrement dit la suite des $\int \vec{f}_n$ et la suite des $\int \vec{g}_n$ ont bien la même limite.

* C'est essentiellement ici qu'intervient le fait que $\vec{F}$ est complet.

Définition. La valeur de la limite indiquée dans le théorème 3 s'appelle l'intégrale de Riemann de la fonction intégrable $\vec{f}$, elle se note : $\int \vec{f}(x)\,dx$, ou simplement, si aucune confusion n'est à craindre, $\int \vec{f} \in \vec{F}$.

Si deux fonctions intégrables $\vec{f}$ et $\vec{g}$ ne diffèrent qu'en un nombre fini de points de $\mathbb{R}$, elles ont meme intégrale, puisque leurs suites d'approximation sont les mêmes.

On dit que $\vec{f}$ est intégrable sur un segment $[a,b]$ de $\mathbb{R}$, si la fonction $\tilde{\vec{f}}$, égale à $\vec{f}$ sur $[a, b]$ et à $\vec{0}$ en dehors, est intégrable, et alors on pose

(IV,1;13bis)
$$\int_{[a,b]} \vec{f} = \int \tilde{\vec{f}} .$$

Remarque - Si $\vec{f}$ est une fonction en escalier, lanouvelle définition que nous venons de donner de $\int \vec{f}$ correspond à la définition initiale (IV,1;4); en effet on peut prendre alors comme suite d'approximation de $\vec{f}$, la suite $\vec{f}, \vec{f}, \vec{f}, \ldots, \vec{f}, \ldots$ et l'on trouve bien l'intégrale de $\vec{f}$ initiale.

Théoréme 4 - Si λ est un scalaire, et si $\vec{f}$ et $\vec{g}$ sont deux fonctions intégrables à valeurs dans $\vec{F}$, alors $\lambda\vec{f}$ et $\vec{f}+\vec{g}$ sont intégrables, et en outre on a la formule :

(IV,1;14)
$$\int \lambda \vec{f} = \lambda \int \vec{f}, \quad \int (\vec{f}+\vec{g}) = \int \vec{f} + \int \vec{g} .$$

Autrement dit on voit que l'ensemble des fonctions intégrables-Riemann (nécessairement toutes bornées et à support compact) est un espace vectoriel sur le corps $\mathbb{K}$, et que l'intégrale est une application linéaire de cet espace vectoriel dans $\vec{F}$.

Démonstration - Ce qui concerne la multiplication par le scalaire est évident, regardons seulement ce qui concerne la somme. Soient $\vec{f}_n$ et $\vec{g}_n$ des suites d approximation de $\vec{f}$ et $\vec{g}$. Alors on a la majoration :

(IV,1;15)
$$\int^* \left\| (\vec{f}+\vec{g}) - (\vec{f}_n+\vec{g}_n) \right\| \leq \int^* \left\| \vec{f} - \vec{f}_n \right\| + \int^* \left\| \vec{g} - \vec{g}_n \right\| .$$

Cela prouve que la suite des $\vec{f}_n + \vec{g}_n$ est une suite d'approximation de $\vec{f} + \vec{g}$, donc d'abord que $\vec{f} + \vec{g}$ est intégrable. Ensuite cela prouve que $\int (\vec{f} + \vec{g})$ est la limite des quantités $\int (\vec{f}_n + \vec{g}_n) = \int \vec{f}_n + \int \vec{g}_n$, donc c'est bien $\int \vec{f} + \int \vec{g}$.

Corollaire . Soient a, b, c, $a \leqslant b \leqslant c$, 3 points de $\mathbb{R}$. Soit $\vec{f}$ une fonction définie sur $[a,c]$, à valeurs dans?. Pour qu'elle soit intégrable sur $[a,c]$, il faut et il suffit qu'elle le soit sur $[a,b]$ et sur $[b,c]$, et on a

$$\int_{[a,c]} \vec{f} = \int_{[a,b]} \vec{f} + \int_{[b,c]} \vec{f} . \tag{IV,1;16}$$

Démonstration - Soit $\vec{\tilde{f}}$ la fonction sur $\mathbb{R}$, égale à $\vec{f}$ sur $[a,c]$ et à $\vec{0}$ en dehors. Soient $\varphi_{[a,b]}$, $\varphi_{[b,c]}$, $\varphi_{[a,c]}$, les fonctions caractéristiques des intervalles $[a,b]$, $[b,c]$, $[a,c]$.

Si $\vec{f}$ est intégrable sur $[a,b]$ et sur $[b,c]$, $\vec{f}\varphi_{[a,b]}$ et $\vec{f}\varphi_{[b,c]}$ sont intégrables, donc aussi $\vec{f}\varphi_{]b,c]}$; alors la somme $\vec{f}(\varphi_{[a,b]} + \varphi_{]b,c]}) = \vec{\tilde{f}}$ aussi, $\vec{f}$ est donc intégrable sur $[a,c]$, et (IV,1;15) donne (IV,1;16).

Inversement si $\vec{f}$ est intégrable sur $[a,c]$, et si $\vec{f}_n$ est une suite d'approximation de $\vec{f}$ pour l'intégrale sur $[a,c]$, $\int^*_{[a,c]} \|\vec{f} - \vec{f}_n\|$ tend vers 0, donc aussi $\int^*_{[a,b]} \|\vec{f} - \vec{f}_n\|$ et $\int^*_{[b,c]} \|\vec{f} - \vec{f}_n\|$, donc les $\vec{f}_n$ sont aussi une suite d'approximation de $\vec{f}$ pour l'intégrale sur $[a,b]$ et $[b,c]$, et $\vec{f}$ est intégrable sur ces intervalles.

Théorème 5 - Si la fonction $\vec{f}$, définie sur $\mathbb{R}$, à valeurs dans?, est intégrable, alors la fonction : $\|\vec{f}\| : x \to \|\vec{f}(x)\|$, est une fonction réelle $\geqslant 0$ intégrable, et on a la majoration :

$$\left\| \int \vec{f} \right\| \leqslant \int \|\vec{f}\| . \tag{IV,1;17}$$

Démonstration - Soit en effet $\vec{f}_n$ une suite d'approximation de $\vec{f}$, par des fonctions en escalier. Alors on a la majoration :

$$\left| \|\vec{f}\| - \|\vec{f}_n\| \right| \leqslant \|\vec{f} - \vec{f}_n\| , \tag{IV,1;18}$$

ce qui prouve que $\int^* \left| \|\vec{f}\| - \|\vec{f}_n\| \right|$ converge vers 0, et que par conséquent la suite des fonctions en escalier $\|\vec{f}_n\|$ est une suite d'approximation pour la fonction $\|\vec{f}\|$; donc cette fonction est bien intégrable, et $\int\|\vec{f}_n\|$ tend vers $\int\|\vec{f}\|$. Par ailleurs, on a, pour tout n, l'inégalité :

$\left\|\int \vec{f}_n\right\| \leqslant \int\|\vec{f}_n\|$; lorsque n tend vers l'infini, $\int \vec{f}_n$ tend vers $\int\vec{f}$ par définition, donc $\left\|\int\vec{f}_n\right\|$ vers $\left\|\int\vec{f}\right\|$ à cause de la continuité de la norme (théorème 9 du chapitre II?, et $\int\|\vec{f}_n\|$ tend vers $\int\|\vec{f}\|$, d'où l'inégalité (IV,1;17).

Corollaire 1 - Si f est une fonction réelle, $\geqslant 0$, intégrable, son intégrale est $\geqslant 0$

Si en effet on applique (IV,1;17), comme $|f| = f$, on trouve que $\int f \geqslant \left|\int f\right| \geqslant 0$.

Corollaire 2 - Si f et g sont deux fonctions réelles intégrables, et si $f \leqslant g$, alors $\int f \leqslant \int g$.

En effet, la fonction $g-f$ est positive, il suffit de lui appliquer le corollaire 1, et de tenir compte de ce que $\int(g-f) = \int g - \int f$ (théorème 4).

Corollaire 3 - Si $\vec{f}$ est intégrable sur $[a,b]$, on a l'inégalité

(IV,1;19) $$\left\|\int_{[a,b]}\vec{f}\right\| \leqslant (b-a)\sup_{a\leqslant x\leqslant b}\|\vec{f}(x)\| .$$

En effet on a d'abord l'inégalité (IV,1;17) relative à $\int_{[a,b]}$, soit $\left\|\int_{[a,b]}\vec{f}\right\| \leqslant \int_{[a,b]}\|\vec{f}\|$. Mais la fonction $\|\vec{f}\|$ est majorée sur $[a,b]$ par la constante $M = \sup_{a\leqslant x\leqslant b}\|\vec{f}(x)\|$, donc, d'après le corollaire 2, $\int_{[a,b]}\|\vec{f}\| \leqslant \int_{[a,b]} M = M(b-a)$.

Corollaire 3 bis - Si f est intégrable sur $[a,b]$ et réelle, et si $m \leqslant f(x) \leqslant M$, on a

(IV,1;20) $$m(b-a) \leqslant \int_{[a,b]} f \leqslant M(b-a) .$$

Conséquence immédiate du corollaire 2.

Corollaire 4 - Si f est une fonction réelle ≥ 0 , intégrable, son intégrale coïncide avec son intégrale supérieure.

Soit en effet f_n une suite d'approximation de f par des fonctions réelles en escalier à support compact. On peut supposer les $f_n \geq 0$, sans quoi on les remplace par les $|f_n|$, qui forment à fortiori une suite d'approximation. Alors on a :

$$(\text{IV},1;20\,bis)\quad \begin{cases} f \leq f_n + |f - f_n| \\ \int^* f \leq \int f_n + \int^* |f - f_n| \end{cases}$$ d'où (théorème 2) :

et de même

$$(\text{IV},1;20\,ter)\quad \begin{cases} f_n \leq f + |f_n - f| \\ \int f_n \leq \int^* f + \int^* |f - f_n| \end{cases}$$ d'où

d'où

$$(\text{IV},1;20\,quater)\quad \left| \int^* f - \int f_n \right| \leq \int^* |f - f_n| \,.$$

Le deuxième membre tend vers 0 pour n infini, et $\int f_n$ tend vers $\int f$, d'où le résultat.

Corollaire 5 - Si $\vec{f}_n$ est une suite de fonctions intégrables, si $\vec{f}$ est bornée à support compact, et si $\int^* \| \vec{f} - \vec{f}_n \|$ converge vers 0 pour n tendant vers l'infini (en Particulier si les $\vec{f}_n$ convergent uniformément vers $\vec{3}$ et gardent leur support dans un compact fixe), alors $\vec{f}$ est intégrable, et $\int \vec{f}_n$ converge vers $\int \vec{f}$.

Nous avons déja vu (remarque 1°, page 407) que $\vec{f}$ était

$$(\text{IV},1;21)\quad \left\| \int \vec{f} - \int \vec{f}_n \right\| \leq \int \| \vec{f} - \vec{f}_n \| = \int^* \| \vec{f} - \vec{f}_n \|$$

montre que $\int \vec{f}_n$ converge vers $\int \vec{f}$.

Théorème 6 (Permutabilité de l'intégrale avec les applications linéaires continues). Soient $\vec{F}$ et $\vec{G}$ deux espaces de Banach, et L une application linéaire continue de $\vec{F}$ dans $\vec{G}$. Si $\vec{f}$ est une fonction définie sur $\mathbb{R}$, à valeurs dans $\vec{F}$, et intégrable-Riemann, alors la fonction $L\circ\vec{f}$, définie sur $\mathbb{R}$, à valeurs dans $\vec{G}$, est intégrable-Riemann, et l'on a la relation :

$$\int L\circ\vec{f} = L\left(\int\vec{f}\right) \in \vec{G} \tag{IV,1;22}$$

Démonstration. Soit d'abord $\vec{g}$ une fonction en escalier à support compact, définie sur $\mathbb{R}$ à valeurs dans $\vec{F}$; alors la fonction $L\circ\vec{g}$ est une fonction en escalier à support compact, et l'on a

$$\int L\circ\vec{g} = \sum_{i=0}^{n}(c_{i+1}-c_i)\,L\circ\vec{g}_i = L\left(\sum_{i=0}^{n}(c_{i+1}-c_i)\,\vec{g}_i\right) = L\left(\int\vec{g}\right). \tag{IV,1;23}$$

Soft maintenant $\vec{f}$, intégrable-Riemann quelconque. soit $\vec{f}_0, \vec{f}_1, \dots, \vec{f}_n, \dots$ une suite d'approximation de $\vec{f}$ par des fonctions en escalier à support compact. Montrons que les fonctions $L\circ\vec{f}_n$ constituent une suite d'approximation de la fonction $L\circ\vec{f}$. On a en effet la majoration :

$$\left\{\begin{aligned} &\|(L\circ\vec{f}-L\circ\vec{f}_n)\| \leq \|L\|\,\|\vec{f}-\vec{f}_n\|, \quad \text{donc}\\ &\int^*\|(L\circ\vec{f})-(L\circ\vec{f}_n)\| \leq \|L\|\int^*\|\vec{f}-\vec{f}_n\|, \end{aligned}\right. \tag{IV,1;24}$$

ce qui prouve bien que le premier membre de cette formule converge vers 0, pour n tendant vers l'infini. Il en résulte bien que $L\circ\vec{f}$ est intégrable.

Par ailleurs, on a la relation (IV,1:23), pour $\vec{g}=\vec{f}_n$. Le premier membre tend, pour n infini, vers $\int L\circ\vec{f}$ puisque les $L\circ\vec{f}_n$ sont une suite d'approximation de $L\circ\vec{f}$. D'autre part, dans le deuxième membre, $\int\vec{f}_n$ tend vers $\int\vec{f}$, et, comme L est continue, ce deuxième membre tend vers $L\left(\int\vec{f}\right)$, et ceci démontre la relation (IV,1;23) pour $\vec{g}=\vec{f}$.

Théorème 7 - Soient $\vec{f}$ et $\vec{g}$ deux fonctions intégrables-Riemann à valeurs dans 2 espaces de Banach $\vec{F}$ et $\vec{G}$, et soit B une application bilinéaire continue de $\vec{F} \times \vec{G}$ dans un espace de Banach $\vec{H}$. Alors la fonction $B(\vec{f}, \vec{g}) : x \longrightarrow B(\vec{f}(x), \vec{g}(x))$, est intégrable-Riemann, et on a l'inégalité :

$$(IV,1;24) \qquad \left\| \int B(\vec{f}, \vec{g}) \right\| \leqslant \| B \| \; ||| \vec{f} ||| \int \| \vec{g} \| .$$

Démonstration : Soient $\vec{f}_0, \vec{f}_1, \dots, \vec{f}_n, \dots$ et $\vec{g}_0, \vec{g}_1, \dots, \vec{g}_n, \dots$ des suites d'approximation de $\vec{f}$ et de $\vec{g}$ par des fonctions en escalier à support compact. Remarquons d'abord que nous pouvons toujours supposer que l'on a l'inégalité $||| \vec{g}_n ||| \leqslant 3 ||| \vec{g} |||$. (Supposons en effet que, dans l'un des intervalles de sa décomposition, la fonction $\vec{g}_n$ soit égaie a un vecteur (constant) de $\vec{G}$, de norme $> 3 ||| \vec{g} |||$. Dans tout cet intervalle, la différence $\| \vec{g} - \vec{g}_n \|$ est alors au moins égale a la constante $2 ||| \vec{g} |||$. Si alors, dans tout intervalle de ce type, nous remplaçons le vecteur constant $\vec{g}_n$ par un vecteur constant $\vec{g}'_n$ proportionnel, de norme $||| \vec{g} |||$, on est sûr que la différence $\| \vec{g} - \vec{g}'_n \|$ y est cette fois-ci majorée par $\| \vec{g} \|$ t $\| \vec{g}'_n \| \leqslant 2 ||| \vec{g} |||$, c'est-a-dire par une quantité plus faible. La fonction $\vec{g}'_n$, remplaçant la fonction $\vec{g}_n$, vérifie alors certainement l'inégalité: $\int^* \| \vec{g} - \vec{g}'_n \| \leqslant \int^* \| \vec{g} - \vec{g}_n \|$,

et par suite les $\vec{g}'_n$ constituent encore une suite d'approximation de $\vec{g}$, mais sont cette fois toutes bornée par $3 ||| \vec{g} |||$). Dans ces conditions, remarquons que la fonction $B(\vec{f}_n, \vec{g}_n)$ est une fonction en escalier à support, et on a la majoration :

$$(IV,1;25) \qquad \| B(\vec{f}_n, \vec{g}_n) - B(\vec{f}, \vec{g}) \|$$

$$\leqslant \| B(\vec{f}_n - \vec{f}, \vec{g}_n) \| + \| B(\vec{f}, \vec{g}_n - \vec{g}) \|$$

$$\leqslant \| B \| \; \| \vec{f}_n - \vec{f} \| \; 3 ||| \vec{g} ||| + \| B \| \; ||| \vec{f} ||| \; \| \vec{g}_n - \vec{g} \|$$

Donc

$$\text{(IV,1;26)} \quad \int^{*} \| B(\vec{f}_n, \vec{g}_n) - B(\vec{f}, \vec{g}) \| \leq \|B\| \left(3 \||\vec{g}\|| \int^{*} \|\vec{f}_n - \vec{f}\| + \||\vec{f}\|| \int^{*} \|\vec{g}_n - \vec{g}\| \right).$$

Il en rdsulte que, les $B(\vec{f}_n, \vec{g}_n)$ sont une suite d'approximation de $B(\vec{f}, \vec{g})$, donc que cette fonction est intégrable-Riemann. Comme alors on a la majoration :

$$\text{(IV,1;27)} \quad \| B(\vec{f}, \vec{g}) \| \leq \|B\| \, \||\vec{f}\|| \, \|\vec{g}\| ,$$

on en déduit bien, pour l'intégrale, la majoration (IV,1;24).

Corollaire 1 - Si f et g sont deux fonctions complexes, intégrables-Riemann, alors la fonction produit fg est intégrable-Riemann, et l'on a l'inégalité :

$$\text{(IV,1;28)} \quad \left| \int fg \right| \leq \|f\| \int |g| .$$

Corollaire 2 - Si f et g sont deux fonctions réelles intégrables-Riemann, la fonction fg est intégrable-Riemann, et, si g est ≥ 0 et $m \leq f(x) \leq M$, on a l'inégalité :

$$\text{(IV,1;29)} \quad m \int g \leq \int fg \leq M \int g .$$

En effet, fg étant intégrable - Riemann, on a la majoration

$$\text{(IV,1;30)} \quad m\, g(x) \leq f(x)\, g(x) \leq M\, g(x), \text{ pour tout } x ,$$

d'où, en intégrant, l'inégalité (IV,1;29) * .

22

* Bien entendu, ces inégalités ne sont plus valables pour g non ≥ 0. Par exemple, si $f(x) = g(x) = x$, et si on intègre sur $[-1, +1]$, on a $M = 1$, mais on n'a pas $\frac{2}{3} = \int_{-1}^{+1} x^2\, dx \leq 1 \int_{-1}^{+1} x\, dx = 0$. Malgré cette mise en garde, l'expérience montre que cette erreur subsiste avec une grande ténacité !

Remarque - Toute quantité intermédiaire entre le minimum et le maximum de f peut s'écrire sous la forme :

$f(\xi)$, $\xi \in \mathbb{R}$, si f est une fonction continue (théorème des valeurs intermédiaires, corollaire du théorème 33 du Chapitre II) On a donc dans ce cas, la relation, toujours pour $g \geq 0$ donc aussi pour g de signe constant :

(IV,1;31) $$\int f(x)\, g(x)\, dx = f(\xi) \int g(x)\, dx$$

Cette relation s'appelle aussi : Théorème de la moyenne.

Exemples de fonctions intégrables-Riemann

Les principaux exemples sont tirés du corollaire 5 du théorème 5.

Théorème 8 - Pour qu'une fonction définie sur un intervalle fermé $[a,b]$ de $\mathbb{R}$, à valeurs dans un espace métrique F complet * soit réglée, il faut et il suffit qu'elle soit limite uniforme d'une suite de fonctions en escalier définies sur $[a,b]$ à valeurs dans F.

Démonstration 1°/Soit $f_0, f_1, \ldots, f_n, \ldots$ une suite de fonction& en escalier sur $[a,b]$, qui convergent uniformément vers une fonction f. Soit c un point de $[a,b]$.

Lorsque x tend vers c par valeurs strictement supérieures, chacune des fonctions f possède une limite, puisque c'est une fonction en escalier. Comme alors F est supposé complet, il résulte du théorème 66 du chapitre II que f possède la même propriété. Le même raisonnement montre que f possède une limite lorsque x tend vers c par valeurs strictement inférieures. On voit que c est une discontinuité de première espèce, et comme ceci est vrai pour tout point c, la limite f est bien réglée.

* Pour la condition nécessaire on n'a pas besoin de supposer F complet.

2°/ Inversement, soit f une fonction réglée à valeurs dans l'espace métrique F . Ici il n'est plus nécessaire de supposer F complet. Nous allons montrer que, quel que soit $\varepsilon > 0$, on peut trouver une fonction en escalier g définie sur $[a,b]$ a-valeurs dans F , et telle que $d(f,g) \leq \varepsilon$. En prenant alors sucessivement $\varepsilon = 1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots$, on formera une suite répondant à la question .

Soit x un point quelconque de $[a,b]$. Si x est distinct de a , on peut, puisque $f(y)$ a une limite quand y tend vers x par valeurs $< x$, trouver un intervalle $]x-\alpha, x]$, $\alpha > 0$, tel que, pour deux points quelconques ξ, η, de cet intervalle distincts de x , on ait : $d(f(\xi), f(\eta)) \leq \varepsilon$.

De la même manière, si x est distinct de b , on peut trouver un intervalle $[x, x+\beta[$, tel que l'on ait la même propriété. En réunissant à la fois ces deux résultats, on voit qu'on peut toujours trouver un <u>ouvert-intervalle de</u> $[a,b]$ * contenant x , tel que, pour deux points quelconques ξ, η , de cet ensemble ouvert, distincts de x et situés tous les deux du même côté de x, on ait cette même inégalité : $d(f(\xi), f(\eta)) \leq \varepsilon$. Comme alors $[a,b]$ est un compact, on peut, d'après la propriété de Heine-Borel-Lebesgue, recouvrir l intervalle $[a,b]$ à l'aide d'un nombre fini de ces ensembles ouverts.

Appelons alors $c_0 = a, c_1, c_2, \dots, c_i, \dots, c_n = b$, l'ensemble formé d'une part de toutes les origines et de toutes les extrémités de ces ouverts-intervalles, et d'autre part des points x qui ont servi à les construire.

Nous définirons alors une fonction en escalier g , comme suit : en chacun des points c_i elle est egale à la valeur de f en ce point. Dans chacun des intervalles

* Pour $x \neq a$ et $x \neq b$, c'est l'ensemble $]x-\alpha, x+\beta[$; pour $x = a$, c'est $[a, a+\beta[$, pour $x = b$, c'est $]b-\alpha, b]$. Ce sont toujours des ouverts de l'espace topologique $[a,b]$. Ce sont aussi des intervalles. C'est pourquoi nous disons : un ouvert-intervalle de $[a,b]$. Car ce n'est pas nécessairement un intervalle ouvert au sens de la page 21.

$]c_i, c_{i+1}[$, elle est égale à la constante $f(\xi_i)$, où ξ_i est un point quelconque de cet intervalle. Alors d'après la manière dont ont été choisis ces intervalles, on a bien :

$d(f(x), g(x)) \leq \varepsilon$ quel que soit x , et ceci démontre le théorème.

Corollaire 1 - Toute fonction définie sur un intervalle fermé $[a,b]$ de $\mathbb{R}$, à valeurs dans un espace métrique F et continue, est limite uniforme d'une suite de fonctions en escalier.

Remarquons que, dans le cas d'une fonction continue, la démonstration étaitbien plus simple que dans le cas d'une fonction réglée quelconque. Il suffit de remar uer que f est uniformément continue sur le compact $[a, b]$ (théorème **31** du chapitre II). Alors, $\varepsilon > 0$ étant donné, il existe $\eta > 0$ tel que $d(x',x'') \leq \eta$ entrafne $d(f(x'), f(x'')) \leq \varepsilon$. Alors, si on partage $[a, b]$ par une décomposition quelconque $c_0 = a, c_1, c_2, \ldots, c_n = b$, telle que tous les $c_{i+1} - c_i$ soient $\leq \eta$, et si on choisit g égale à $f(\xi_i)$ dans $[c_i, c_{i+1}[$, $\xi_i \in [c_i, c_{i+1}]$, et $g(b) = f(b)$,on a bien $d(f, g) \leq \varepsilon$.

Corollaire 2 - Toute fonction sur la droite $\mathbb{R}$, à support compact, à valeurs dans un espace de Banach F et réglée, est intégrable-Riemann; en particulier toute fonction continue à support compact est integrable.

Corollaire 3 - Toute fonction réelle définie sur un intervalle $[a, b]$ de $\mathbb{R}$, et monotone sur $[a, b]$, est intégrable.

Une fonction monotone est en effet réglée.

Il y a bien entendu des fonctions non réglées qui sont aussi intégrables. Par exemple la fonction, égale à $\sin \frac{1}{x}$ sur l'intervalle $]0, 1]$ et à 0 en dehors, est intégrable-Riemann, bien que sa discontinuité à l'origine ne soit pas une discontinuité de première espèce.

On a en effet :

<u>Théorème 9 - Toute fonction à valeurs dans un espace de Banach $\vec{F}$, bornée, à support compact, et partout continue, sauf au plus en un nombre fini de points, est intégrable-Riemann</u> *

<u>Démonstration</u> Nous allons appliquer la remarque 1°/ de la page 407, et montrer que, quel que soit $\varepsilon > 0$, il existe une fonction intégrable $\vec{g}$ telle que $\int^* \|\vec{f} - \vec{g}\| \leq \varepsilon$; le théorème en résultera. Soit M le maximum de $\|\vec{f}\|$, N le nombre des points de diseontinuité de $\vec{f}$. Considérons la fonction $\vec{g}$, égale à $\vec{f}$ en dehors de la réunion des intervalles $[c_i - \eta, c_i + \eta]$ où les c_i sont les points de discontinuité de $\vec{f}$, et où $\eta = \frac{\varepsilon}{2NM}$, et à $\vec{0}$ dans ces intervalles. Alors la fonction $\vec{g}$ est partout oontinue, sauf au plus en un nombre fini de points, où elle a une discontinuité de première espèce, et par conséquent elle est intégrable, d'après le corollaire 2 du théorème 8. Par ailleurs, la l'onction $\|\vec{f} - \vec{g}\|$ est majorée par la fonction en escalier égale à la constante M dans chacun des N intervalles $[c_i - \eta, c_i + \eta]$, et à 0 en dehors, de sorte que l'on a l'inégalité $\int^* \|\vec{f} - \vec{g}\| \leq M.2N\eta = \varepsilon$.

Calcul de l'intégrale d'une fonction par la méthode des sommes de Cauchy-Riemann

Soit $\vec{f}$ une fonction définie sur un intervalle $[a,b]$ de $\mathbb{R}$, à valeurs dans un espace de Banach $\vec{F}$, et intdgrable-Riemann; on se propose de calculer son Intégrale comme suit.

On choisit une décomposition Δ de $[a,b]$ en intervalles, grâce'à des points de subdivision $c_0 = a, c_1, c_2, \ldots, c_i, \ldots c_n = b$. Pour chaque indice i, on appelle $\vec{\theta}_i$ un vecteur quelconque de $\vec{F}$ adhérent à l'ensemble des valeurs de $\vec{f}$ dans l'intervalle

* Il faut se garder de croire que, si $\vec{f}$ est une fonction ayant ces propriétés, on puisse la rendre continue en modifiant sa valeur en un nombre fini de points de $\mathbb{R}$; l'exemple précédent le montre !

fermé $[c_i, c_{i+1}]$. On construit alors la somme de Cauchy-Riemann $\sum_{i=0}^{n-1} (c_{i+1} - c_i) \vec{\theta}_i$. Dans quelle mesure cette somme donne-t-elle une approximatinn de l'intégrale ? Plus précisément, nous allons montrer que, si l'on prend n'importe quelle suite de telles décompositions $\Delta_1, \Delta_2, \Delta_3, \ldots, \Delta_n, \ldots$ pourvu que la longueur du plus grand des intervalles de la décomposition Δ_n, que nous allons appeler la finesse de la décomposition Δ_n, tende vers 0 pour n infini, alors la somme précédente (où les $\vec{\theta}_i$ sont arbitraires pour chaque Δ_n) tend nécessaire - ment vers $\int \vec{f}$ pour n infini.

Théorème 10 - Soit $\vec{f}$ est une fonction définie sur $[a, b]$, intervalle borné de $\mathbb{R}$, à valeurs dans l'espace de Banach $\vec{F}$, et integrable-Riemann. Alors, quel que soit $\varepsilon > 0$, on peut trouver un nombre $\eta > 0$ ayant la propriété suivante:

pour toute décomposition Δ de l'intervalle $[a, b]$, définie par une suite finie croissante $c_0 = a, c_1, c_2, \ldots, c_n = b$, telle que le plus grand des intervalles $[c_i, c_{i+1}]$ ait une longueur $\leq \eta$, et pour tout choix, pour chaque indice i, d'un vecteur $\vec{\theta}_i$ adhérent à l'ensemble des valeurs de $\vec{f}$ sur l'intervalle $[c_i, c_{i+1}]$, on a l'inégalité :

(IV, 1; 32) $$\left\| \int_{[a,b]} \vec{f} - \sum_{i=0}^{n-1} (c_{i+1} - c_i) \vec{\theta}_i \right\| \leq \varepsilon .$$

Démonstration Tout d'abord, nous pouvons trouver une fonction en escalier $\vec{g}$, définie sur $[a, b]$ à valeur dans $\vec{F}$, et une fonction en escalier h, à valeurs réelles ≥ 0, telles que l'on ait : $\| \vec{f} - \vec{g} \| \leq h$ et $\int h \leq \frac{\varepsilon}{4}$ (remarque 2°/, page 408) Appelons $d_0 = a, d_1, d_2, \ldots, d_j, \ldots d_N = b$, les points de subdivision correspondant à une décomposition commune aux deux fonctions en escalier $\vec{g}$ et h.
Montrons alors que le nombre $\eta = \frac{\varepsilon}{8NM}$, $M = \|\vec{f}\|$, répond à la question. Soit en effet Δ une decomposition quelconque de $[a, b]$ par des points c_i, $i = 0, 1, 2, \ldots n$, de finesse $\leq \eta$. Choisissons les $\vec{\theta}_i$ comme il est indiqué, et appelons $\vec{\theta}$ la fonction en escalier, admettant les points de subdivision c_i, égale à $\vec{\theta}_i$ dans chaque intervalle

$[c_i , c_{i+1}[$, et à $\vec{f}(b)$ pour $x = b$. La somme de Riemann $\sum_{i=0}^{n-1} (c_{i+1} - c_i)\, \vec{\theta}_i$ n'est autre que $\int_{[a,b]} \vec{\theta}$.
Cherchons alors une majoration de la fonction $\| \vec{f}(x) - \vec{\theta}(x) \|$, dans chaque intervalle $[c_i , c_{i+1}[$.

1°/ Supposons d'abord que l'intervalle $[c_i , c_{i+1}[$ soit contenu dans l'un des intervalles $]d_j , d_{j+1}[$. Dans cet intervalle, on a d'abord L'inégalité :

$$\| \vec{f}(x) - \vec{\theta}(x) \| = \| \vec{f}(x) - \vec{\theta}_i \| \leqslant \| \vec{f}(x) - \vec{g}(x) \| + \| \vec{g}(x) - \vec{\theta}_i \| \tag{IV,1;33}$$

La première norme est majorée par $h(x)$. Quant à la deuxième, d'après la définition même de $\vec{\theta}_i$, elle est majorée par :

$$\sup_{c_i \leqslant \xi < c_{i+1}} \| \vec{g}(x) - \vec{f}(\xi) \| \ ; \tag{IV,1;34}$$

mais, $\vec{g}$ étant constante dans l'intervalle$]d_j , d_{j+1}[$, on peut aussi, dans cette formule, remplacer $\vec{g}(x)$ par $\vec{g}(\xi)$; donc finalement le terme $\| \vec{g}(x) - \vec{\theta}_i \|$ est majoré par $\sup_{c_i \leqslant \xi < c_{i+1}} h(\xi)$, égal aussi a $h(x)$ puisque h est une constante dans l'intervalle $]d_j , d_{j+1}[$. Finalement on a , dans l'intervalle considéré, la majoration :

$$\| \vec{f}(x) - \vec{\theta}(x) \| \leqslant 2\, h(x) . \tag{IV,1;35}$$

La somme des intégrales de la fonction $\| \vec{f} - \vec{\theta} \|$ dans les divers intervalles considérés, est finalement majorée par $2 \int h \leqslant \frac{\varepsilon}{2}$.

2°/ Supposons maintenant que l'intervalle $[c_i , c_{i+1}[$ contienne l'un des points de subdivision d_j . Comme $\vec{\theta}_i$ est adhérent à un ensemble de valeurs de $\vec{f}$, on peut affirmer que $\| \vec{f}(x) - \vec{\theta}(x) \|$ est majorée par la constante 2 M dans cet intervalle. Il y a au plus $2N$ de tels intervalles, et la longueur de chacun est majorée par η , de sorte que la somme des intégrales de $\| \vec{f} - \vec{\theta} \|$, dans ces divers intervalles, est majorée par 4 $M N \eta = \frac{\varepsilon}{2}$.

Finalement on a bien :

$$\int_{[a,b]} \| \vec{f}(x) - \vec{\theta}(x) \| dx \leqslant \varepsilon \quad , \tag{IV,1;36}$$

et le théorème est démontré.

<u>Corollaire</u> - <u>Soit f une fonction réelle définie sur $[a,b]$, intervalle borné de $\mathbb{R}$, et intégrable-Riemann. Alors, quel que soit $\varepsilon > 0$, il existe $\eta > 0$ ayant la propriété suivante : pour toute décomposition A de $[a,b]$, définie par une suite finie croissante $c_0 = a$, $c_1 , c_2 , \ldots$ $c_n = b$ telle que le plus grand des nombres $c_{i+1} - c_i$ soit $\leqslant \eta$, si l'on pose</u> $M_i = \operatorname{Sup}_{c_i \leqslant x \leqslant c_{i+1}} f(x)$, $m_i = \operatorname{Inf}_{c_i \leqslant x \leqslant c_{i+1}} f(x)$, <u>on a les inégalités</u> :

$$\sum_{i=0}^{n-1} (c_{i+1} - c_i) m_i \leqslant \int_{[a,b]} f(x) dx \tag{IV,1;37}$$

$$\leqslant \sum_{i=0}^{n-1} (c_{i+1} - c_i) M_i \quad , \text{ et}$$

$$\sum_{i=0}^{n-1} (c_{i+1} - c_i)(M_i - m_i) \leqslant \varepsilon \ . \tag{IV,1;38}$$

Cela résulte immédiatement du théorème, puisque dans $[c_i , c_{i+1}]$, on a $m_i \leqslant f(x) \leqslant M_i$, et que M_i et m_i sont adhérents à l'ensemble des valeurs de f .

Remarque - Ce corollaire précise la remarque 3°/ de la page 408 . Les fonctions g_1 et g_2 pourront être celles qui, dans $[c_i, c_{i+1}[$, prennent respectivement les valeurs m_i et M_i , et la valeur $f(b)$ en b .

Valeur moyenne d'une fonction dans un intervalle. Si $\vec{f}$ est une fonction définie sur un intervalle $[a,b]$ de $\mathbb{R}$, à valeurs dans un espace de Banach $\vec{F}$ et intégrable-Riemann, on appelle moyenne de $\vec{f}$ dans l'intervalle $[a,b]$ la quantité :

$$\mathfrak{M}(\vec{f};[a,b]) = \frac{1}{b-a}\int_{[a,b]} \vec{f}(x)\,dx \in \vec{F}\ ^{*} \tag{IV,1;39}$$

Il résulte des inégalités (IV,1;19) et (IV,1;20) que cette moyenne est majorée par $|||\vec{f}|||$, et que, d'autre part, si est réelle, elle est comprise entre la borne inférieure et la borne supérieure de f . Cette moyenne peut être calculée comme une limite par la formule suivante :

$$\mathfrak{M}(\vec{f};[a,b]) = \lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} \vec{f}\left(a + k\left(\frac{b-a}{n}\right)\right) \tag{IV,1;40}$$

Considérons, en effet, une décomposition de l'intervalle $[a,b]$ en n intervalles tous égaux à $\frac{b-a}{\text{l-b}}$, et calculons la somme de Cauchy-Riemann :

$$\sum_{i=0}^{n-1} (c_{i+1}-c_i)\,\vec{f}(c_i) = \sum_{k=0}^{n-1} \frac{b-a}{n}\,\vec{f}\left(a + k\left(\frac{b-a}{n}\right)\right) \tag{IV,1;41}$$

Il résulte du théorème 10 que, $\vec{f}$ étant intégrable-Riemann, cette somme converge nécessairement vers $\int_{[a,b]} \vec{f}$.

* Si $a = b$, on convient que la moyenne est $\vec{f}(a)$.

On obtiendra (IV,1;40) en multipliant par $\frac{n}{n+1}$ $\frac{1}{b-a}$, qui tend vers $\frac{1}{b-a}$ pour n infini, puis en ajoutant $\frac{1}{n+1}$ $\vec{f}(b)$, qui tend vers $\vec{0}$.

Si f est continue réelle, le théorème des valeurs intermédiaires (corollaire du théorème 33 du chapitre II) montre que $\mathfrak{M}(f;[a,b])$ est égale à $f(c)$, où c est un certain point de $[a,b]$.

§ 2 MESURES DE RADON SUR UN ESPACE LOCALEMENT COMPACT

Mesures de Radon sur un espace compact

Soit X un espace topologique compact. Appelons $\mathscr{C}(X)$ * l'espace, que nous avions appelé jusqu'à présent $(\mathbb{K}^X)_c$, des fonctions continues sur X à valeurs scalaires.

D'après le théorème du maximum (théorème 29 du chapitre II), une fonction continue scalaire sur le compact X est bornée , et nous avons déjà vu que la fonction $\varphi \longrightarrow |||\varphi||| = \sup_{x\in X} |\varphi(x)|$ est une norme sur $\mathscr{C}(X)$. La convergence au sens de cette norme est la convergence uniforme des fonctions φ . D'autre part, le corollaire du théorème 64 du chapitre II nous a indiqué que $\mathscr{C}(X)$ est un espace de Banach.

Le dual de cet espace, conformement à la notation généraledu dual, s'écrit $\mathscr{C}'(X)$. Un élément de ce dual est ce qu'on appelle <u>une mesure de Radon sur X</u> . Si donc μ est une telle mesure, elle définit une application qui, à tout élément φ de $\mathscr{C}(X)$, c'est-à-dire à toute fonction continue φ , définie sur X , à valeurs scalaires, fait correspondre un scalaire $\mu(\varphi)$.

* Il peut paraître étrange de changer de notation ! Mais nous étudierons plus loin le cas où X n'est pas compact, et nous introduirons alors un espace $\mathscr{C}(X)$ qui ne coïncidera plus avec $(\mathbb{K}^X)_c$. Presque toujours $\mathbb{K}$ sera le corps des complexes $\mathbb{C}$.

L'application ainsi définie est linéaire, en ce sens que l'on a

(IV,2;1) $$\mu(\varphi_1+\varphi_2)=\mu(\varphi_1)+\mu(\varphi_2)\ ,\ \mu(k\varphi)=k\,\mu(\varphi) \text{ pour } k\in\mathbb{K}\,;$$

d'autre part, elle est continue, ce qui peut s'exprimer de plusieurs manières : ou bien en disant que, lorsqu'une suite $\varphi_0, \varphi_1, \ldots \varphi_n, \ldots$ de fonctions scalaires continues converge uniformément vers 0 sur X, les nombres $\mu(\varphi_n)$ convergentvers 0 ; ou bien en disant que la mesure μ possède une norme $\|\mu\|$,telle que l'on ait

(IV,2,2) $$|\mu(\varphi)|\leq\|\mu\|\ \|\!|\varphi|\!\|\ ,\ \text{et}\ \|\mu\|=\operatorname*{Sup}_{|\!\|\varphi|\!\|\leq 1}|\mu(\varphi)|\,.$$

D'autre part $\mathcal{C}'(X)$, en tant que dual d'un espace de Banach, est lui-même un espace de Banach (théorème 50 du chapitre II). Sa structure vectorielle est définie comme suit :

Si μ et 3 sont deux mesures, la mesure μ + 3 . est définie par la formule :

(IV,2;3) $$(\mu+\nu)(\varphi)=\mu(\varphi)+\nu(\varphi)\,;$$

SI k est un scalaire,la mesure $k\mu$ est définie par la formule :

(IV,2;4) $$(k\mu)(\varphi)=k\,\mu(\varphi)\,;$$

et la norme de μ est précisément celle qui est définie à (IV,2;2).

Conformément aux notations générales, $\mu(\varphi)$ peut aussi s'écrire $\mu\cdot\varphi$ ou $<\mu,\varphi>$.

Exemples de mesures de Radon

<u>1er exemple</u> : On appelle $\delta_{(a)}$ ou <u>mesure de Dirac au point a de X ,</u> la mesure définie par la formule :

(IV,2;5) $$\delta_{(a)}(\varphi)=\varphi(a)\,.$$

On dit aussi que <u>cette mesure est constituée par la masse unité au point A de X</u> . Si X est un espace vectoriel de dimension finie, on appelle <u>mesure de Dirac</u> (sans spécifier de point a) celle qui est relative à l'origine de X .

Soient maintenant $a_0, a_1, \ldots$ une infinité dénombrable de points sur l'espace X , et $c_0, c_1, \ldots$ une infinité dénombrable de nombres complexes. On note par $\sum c_\nu \delta_{(a_\nu)}$ la mesure μ définie par la formule :

(IV,2;6) $$\left(\sum_\nu c_\nu \delta_{(a_\nu)}\right) \cdot \varphi = \sum_\nu c_\nu \varphi(a_\nu) .$$

Cette expression aura sûrement un sens, si la séries $|c_\nu|$ est convergente, ce que nous supposerons toujours. On dit que cette mesure est atomique et qu'elle est constituée par l'infinité dénombrable des masses c_ν placées aux points a_ν de X .

Montrons que la norme de cette mesure, si tous les a_ν sont distincts, est précisément la somme de la série :

(IV,2;7) $$\left\| \sum_\nu c_\nu \delta_{(a_\nu)} \right\| = \sum_\nu |c_\nu|$$

Tout d'abord, si la fonction φ est majorée en module par 1, on a certainement :

(IV,2;7 bis) $$|\mu(\varphi)| \leq \sum_\nu |c_\nu|$$

ce qui prouve que $\|\mu\| \leq \sum_\nu |c_\nu|$. D'autre part, d'après la définition de la somme d'une série à termes positifs convergente, il existe un sous-ensemble fini J de l'ensemble d'indices, tel que la somme $\sum_{\nu \in J} |c_\nu|$ soit $\geq \left(\sum_\nu |c_\nu|\right) - \frac{\varepsilon}{2}$.

SI alors nous désignons par φ une fonction continue complexe sur X , majorée en module par 1, et prenant, pour chacun des points a_ν , $\nu \in J$, la valeur $\frac{c_\nu}{|c_\nu|}$ * , on voit que l'on a exactement

* Si X est, par exemple, un intervalle $[a, b]$ de $\mathbb{R}$, il est bien Cvident qu'on peut trouver une fonction φ complexe continue, majorée en module par 1, et prenant, en un nombre fini de points de X , des valeurs données de module 1; si $d_1, d_2, \ldots d_\ell$ sont ces points, il suffit de choisir φ affine dans chaque intervalle $[a, d_1], [d_1, d_2], \ldots [d_\ell, b]$. Cette possibilité n'est plus du tout évidente si X est un compact quelconque. C'est cependant essentiel à notre démonstration. Nous verrons cela ultérieurement (corollaire 3 du théorème 11).

(IV,2;8) $$\mu(\varphi) = \sum_{\nu \in J} |c_\nu| + \sum_{\nu \notin J} c_\nu \varphi(a_\nu)$$

On a donc la minoratlon :

(IV,2;9) $$|\mu(\varphi)| \geq \left(\sum_\nu |c_\nu| - \frac{\varepsilon}{2}\right) - \sum_{\nu \notin J} |c_\nu| \geq \sum_\nu |c_\nu| - \varepsilon .$$

On a donc **nécessairement,** puisque $\|\varphi\| = 1$,

(IV,2;10) $$\|\mu\| \geq |\mu(\varphi)| \geq \sum_\nu |c_\nu| - \varepsilon \quad ;$$

comme $\varepsilon > 0$ est arbitraire, on a $\|\mu\| \geq \sum_\nu |c_\nu|$, ce qui donne le résultat cherché.

Remarquons aussi que la notation $\sum$, écrite au début sans justification, est a posteriori correcte : <u>il s'agit d'une série normalement convergente de vecteurs de l'espace de Banach</u> $\mathscr{C}'(X)$. En effet, comme $\|c_\nu \delta_{(a_\nu)}\| = |c_\nu|$, et que $\sum_\nu |c_\nu| < +\infty$, la série $\sum_\nu c_\nu \delta_{(a_\nu)}$ converge bien normalement , vers un dldment de $\mathscr{C}'(X)$; soit μ_1 cet **élément.** Comme $\nu \longrightarrow \nu(\varphi)$, pour φ fixée, est une forme linéaire continue sur $\mathscr{C}'(X)$, le théorème 60 du chapitre II nous dit que

(IV,2;11) $$\mu_1(\varphi) = \left(\sum_\nu c_\nu \delta_{(a_\nu)}\right) \cdot \varphi = \sum_\nu \left(c_\nu \delta_{(a_\nu)} \cdot \varphi\right) = \sum_\nu c_\nu \varphi(a_\nu)$$

$$= \mu(\varphi) \text{ par définition;}$$

donc $\mu = \mu_1$ et s'écrit donc correctement $\sum_3 c_\nu \delta_{(a_\nu)}$

<u>2ème exemple.</u> Prenons pour X un Intervalle compact $[a, b]$ de $\mathbb{R}$. Alors l'intégrale $\varphi \longrightarrow \int_{[a,b]} \varphi(x)dx$ définit une me sure μ de Radon sur X .

L'inégalité (IV,1;19) montre que la norme de cette mesure est $\leq b - a$. Par ailleurs, si l'on prend la valeur de cette mesure sur la fonction $\varphi_1 \equiv 1$ on voit que l'on a la minoration $\|\mu\| \geq \left|\int_{[a,b]} \varphi\right| = b - a$; <u>la norme de cette mesure est donc exactement $b - a$</u> .

Plus généralement, si f est une fonction définie sur $[a,b]$ à valeurs complexes, et lntégrable-Riemann, elle définit une mesure par la formule

(IV,2;12) $$\mu(\varphi) = \int_{[a,b]} f(x)\,\varphi(x)\,dx\,.$$

Cette mesure a visiblement une norme $\leqslant \int |f(x)|\,dx$; nous montrerons plus loin que sd norme est exactement égale a cette intégrale . Les mesures ainsi définies, contrairement aux mesures atomiques, sont appelées diffuses, en ce sens qu'elles ne comportent pas de masse ponctuelle. Gn dit aussi que cette mesure est de densité f par rapport à la mesure dx , et on l'écrit $f\,dx$ ou $f(x)\,dx$, * . C'est à cause de la notation dx pour la mesure $\varphi \longrightarrow \int_{[a,b]} \varphi$ qu'une mesure quelconque se note fréquemment $d\mu$ à la place de μ, et que l'expression $\mu(\varphi)$ s appelle aussi integrale de φ par rapport à la mesure μ ou $d\mu$, et qu'elle se note souvent par

(IV,2;13) $$\int_X \varphi(x)\,d\mu(x) \text{ ou } \int_X \varphi\,d\mu \text{ ou } \int_X \varphi\,\mu \text{ au lieu de } \mu(\varphi)$$

* Appeler $\int$ la mesure $\varphi \longrightarrow \int_{[a,b]} \varphi(x)\,dx$ est parfaitement correct; l'appeler dx est une monstruosité logique, car x figure dans 1 intégrale comme une variable muette, pouvant être remplacée par y ou t ou n'importe quel autre symbole. De même il n'est guère correct de parler de la mesure $f\,dx$ mais la notation $f(x)\,dx$ peut se justifier par l'emploi de x , écrit 2 fois, comme variable muette; alors $f(x)\,dx$ peut aussi s'écrire $f(y)\,dy$ ou $f(t)\,dt$. Ces incorrections sont les mêmes que celle qui consiste à dire : "la fonction x^2 " au lieu de : "la fonction $x \longrightarrow x^2$ " (voir page 6). Néanmoins, on se permettra souvent cet abus, comme on le fait pour les fonctions. S'il y a doute, on dira "la mesure 1 $(x)\,dx$ "au lieu de "la mesure dx ", puisque c'est la mesure $f(x)\,dx$, avec $f \equiv 1$, $f(x) = 1$ pour tout x ; on pourra aussi toujours, pour dx comme pour $f\,dx$, écrire au complet : 'la mesure $\varphi \longrightarrow \int_{[a,b]} \varphi(x)\,f(x)\,dx$".

On peut naturellement former une mesure, somme des deux mesures étudiées dans les exemples 1 et 2, c'est-à-dire définie par la formule

(IV,2;14) $$\mu(\varphi) = \sum_\nu c_\nu \varphi(a_\nu) + \int_{[a,b]} f(x)\varphi(x)\,dx \ .$$

On dit que cette mesure est la somme des masses ponctuelles c_ν placées aux points a_ν de $[a,b]$ et de la mesure diffuse de densité f par rapport à la mesure dx. Elle se note par la formule

(IV,2;15) $$\mu = \sum_\nu c_\nu \delta_{(a_\nu)} + f\,dx \ .$$

Sa norme est manifestement majorée par

(IV,2;16) $$\sum_\nu |c_\nu| + \int_{[a,b]} |f(x)|\,dx \ ;$$

montrons qu'elle est exactement égale à cette quantité, si tous les a_ν sont distincts.

Tout d'abord, $\varepsilon > 0$ étant donné, on détermine un ensemble fini J d'indices tel que $\sum_{\nu \notin J} |C_\nu| \leq \frac{\varepsilon}{6}$.

Ensuite, puisque f est intégrable-Riemann, on peut trouver une fonction en escalier q sur $[a,b]$ tel que l'on ait l'inégalité :

(IV,2;16 bis) $$\int_{[a,b]} |f(x) - q(x)|\,dx \leq \frac{\varepsilon}{6} \ .$$

Appelons alors $d_0 = a_1, d_1, \dots, d_n = b$, l'ensemble formé d'une part des points de subdivision de q et d'autre part des points a_ν, $\nu \in J$. On peut alors, pour tout $i = 0,1,2,\dots,n-1$, trouver un intervalle $[d'_i, d''_{i+1}] \subset]d_i, d_{i+1}[$ tel que l'on ait l'inégalité

(IV,2;16 ter) $$\sum_{i=0}^{n-1} |q_i|\big((d'_i - d_i) + (d_{i+1} - d''_{i+1})\big) \leq \frac{\varepsilon}{6} \ ; \ q_i = q \text{ dans }]d_i, d_{i+1}[$$

On définira alors une fonction φ de la manière suivante :

Dans chaque intervalle $[d'_i, d''_{i+1}]$ elle est égale à la constante $\frac{q_i}{|q_i|}$; en chacun des points d_i, $i = 0,1,2,\dots,n$, qui est un point a_ν, $\nu \in J$, elle est égale à la constante $\frac{c_\nu}{|c_\nu|}$

Dans les intervalles où elle n'a pas encore été définie, elle est prolongée de façon à être affine. Le module de φ est manifestement partout majoré par 1 .

D'autre part on a les inégalités suivantes :

(IV,2;16 quarto)
$$\left| \sum_\nu c_\nu \varphi(a_\nu) - \sum_\nu |c_\nu| \right| \leqslant$$
$$\left| \sum_\nu c_\nu \varphi(a_\nu) - \sum_{\nu\in J} c_\nu \varphi(a_\nu) \right|$$
$$+ \left| \sum_{\nu\in J} c_\nu \varphi(a_\nu) - \sum_{\nu\in J} |c_\nu| \right|$$
$$+ \left| \sum_{\nu\in J} |c_\nu| - \sum_\nu |c_\nu| \right|$$

Le 2^e terme est nul puisque $\varphi(a_\nu) = \frac{\overline{c_\nu}}{|c_\nu|}$ pour $\nu \in J$; le 1^{er} et le 3^e sont majorés par $\frac{\varepsilon}{6}$, d'où

(IV,2 16 quinto)
$$\left| \sum_\nu c_\nu \varphi(a_\nu) - \sum_\nu |c_\nu| \right| \leqslant \frac{2\varepsilon}{6} = \frac{\varepsilon}{3}$$

Ensuite .

(IV,2 ; 16 sexto)
$$\left| \int_{[a,b]} \mu(x)\varphi(x)\,dx - \int_{[a,b]} |\mu(x)|\,dx \right| \leqslant$$
$$\left| \int_{[a,b]} \mu(x)\varphi(x)\,dx - \int_{[a,b]} q(x)\varphi(x)\,dx \right|$$
$$+ \left| \int_{[a,b]} q(x)\varphi(x)\,dx - \int_{[a,b]} |q(x)|\,dx \right|$$
$$+ \left| \int_{[a,b]} |q(x)|\,dx - \int_{[a,b]} |\mu(x)|\,dx \right|$$

La 1ère et la 3^e intégrales sont majorées par $\frac{\varepsilon}{6}$, d'après (IV,2;16bis).

La 2^e intégrale, compte tenu de ce que φ vaut $\frac{q_i}{|q_i|}$ dans $[d'_i, d''_{i+1}]$, est majorée par $\frac{2\varepsilon}{6} = \frac{\varepsilon}{3}$ en vertu de (IV,2;16ter).

Donc

(IV,2;16 septimo)
$$\left| \int_{[a,b]} \mu(x)\varphi(x)\,dx - \int_{[a,b]} |\mu(x)|\,dx \right|$$
$$\leqslant \frac{2\varepsilon}{6} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3}$$

En ajoutant (IV,2;16 quinto) et (IV,2;16 septimo), on obtient, pour la fonction φ choisie :

(IV,2;16 octavo)
$$\left| \mu(\varphi) - \left(\int_{[a,b]} |\mu(x)|\,dx + \sum_\nu |c_\nu| \right) \right| \leqslant \varepsilon ,$$

Ce qui montre notre affirmation, à savoir que $\|\mu\|$ est exactement (IV,2;16).

Bien remarquer que la mesure de densité f ne change pas si l'on change la fonction f en un nombre fini de points. La densité& détermine la mesure, mais des densités différentes peuvent donner la même mesure.

3ème exemple Bien que nous ne connaissions pas encore les intégrales multiples, nous pouvons tout-de-même dire que, si X est un pavé fermé de $\mathbb{R}^n$, c'est-à-dire un ensemble défini par les inégalités $a_i \leqslant x_i \leqslant b_i$, $i = 1, 2, \ldots n$, alors on peut définir sur X la mesure $dx = dx_1 dx_2 \ldots . dx_n$ définie par la formule

$$(\text{IV}, 2; 17) \qquad \varphi \longrightarrow \int\!\!\int \cdots \int_{\prod_{i=1}^{n} [a_i, b_i]} \varphi(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots \ldots dx_n .$$

La norme de cette mesure n'est autre que le volume du pavé X , c'est-a-dire $\prod_{i=1}^{n} (b_i - a_i)$.

Mesures sur un espace localement compact X

Prenons par exemple $X = \mathbb{R}$, qui n'est pas compacte. Cependant les formes linéaires $\varphi \longrightarrow \varphi(a)$ et $\varphi \longrightarrow \int \varphi(x)\, dx$, continuent à avoir un sens, au moins si l'on fait sur φ certaines restrictions, par exemple la restriction d'avoir un support compact. Nous sommes donc amenés à poser la définition suivante :

Soit X un espace topologique localement compact. On **appelle** $\mathcal{C}(X)$ **l'espace** vectoriel des fonctions φ **définies sur** X, **à valeurs** scalaires, continues, et à support compact *

(ce support naturellement n'est pas précisé; chacune des fonctions φ de cet espace a un **support** K compact, mais K dépend **de** φ). Notons par ailleurs qu'il ne faut absolument pas confondre une fonction φ définie et **continue sur** X tout entier, à support compact, avec une fonction définie et continue sur un compact de X .

Par exemple, la fonction dont le graphique est donné a la figure 1 est une fonction définie et continue sur l'intervalle $[a, b]$; elle n'est pas une fonction définie et continue **sur** toute la droite $\mathbb{R}$, et de support dans $[a, b]$. **Si** on la prolonge **sur** $\mathbb{R}$ en lui **donnant** la valeur 0 en dehors **de** $[a, b]$, elle n'est pas continue **sur** $\mathbb{R}$, puisqu'elle **présente** a et b comme points de **discontinuité.**

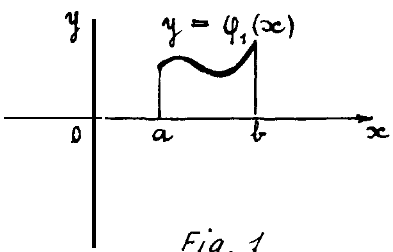


Fig. 1

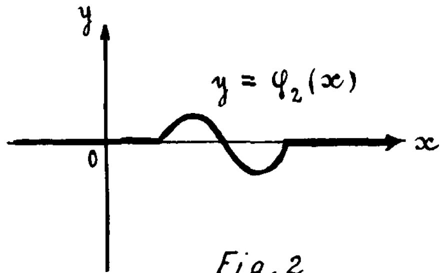


Fig. 2

Par contre la fonction dont le graphique est donné à la figure 2 est bien une fonction définie **sur** $\mathbb{R}$, continue, à support compact, par conséquent elle appartient bien à $\mathcal{C}(\mathbb{R})$ **Si** K est un compact de X , nous appellerons $\mathcal{C}_K(X)$ le **sous**-espace vectoriel de $\mathcal{C}(X)$ formé des fonctions 0 ayant leur support dans le compact K . Une fonction φ de $\mathcal{C}(X)$ est dans $\mathcal{C}_K(X)$, si et seulement si elle est nulle **sur** $\complement K$. **Mais alors elle est** nécessairement nulle, non seulement **sur** $\complement K$, **mais** sur la frontière de K, **puisqu**'un point de cette frontière est limite de points de $\complement K$ et que φ est continue. Ne pas **confondre** $\mathcal{C}_K(X)$ et $\mathcal{C}(K)$; **la fonction** φ_1 **de la figure** 1 **appartient à** $\mathcal{C}([a, b])$, **la fonction** φ_2 **de la figure** 2 **à** $\mathcal{C}_{[a,b]}(\mathbb{R})$

* $\mathcal{C}(X)$ est un sous-espace de $(\mathbb{K}^X)_{cb}$. Voir note *

page 425.

Alors $\mathcal{C}(X)$ est la réunion des espaces vectoriels $\mathcal{C}_K(X)$, lorsque K parcourt tous les sous-espaces compacts de X.

L'espace $\mathcal{C}_K(X)$ est un espace de Banach, lorsque l'on met sur cet espace la norme $|||\varphi||| = \sup_{x\in X}|\varphi(x)|$. En effet $(\mathbb{K}^X)_{cb}$ est un Banach pour cette norme (corollaire 3 du théorème 65 du chapitre II); et $\mathcal{C}_K(X)$ en est un sous-espace vectoriel fermé (si en effet $\varphi_n \in \mathcal{C}_K(X)$ converge, pour n infini, vers φ dans $(\mathbb{K}^X)_{cb}$, c'est-a-dire uniformément et a fortiori simplement, comme $\varphi_n(x) = 0$ pour $x \in \complement K$, on a aussi $\varphi(x) = 0$, donc $\varphi \in \mathcal{C}_K(X)$) ; le théorème 43 du chapitre II montre donc bien que $\mathcal{C}_K(X)$ est un Banach . Naturellement il serait également possible de mettre la même norme sur $\mathcal{C}(X)$; mais alors on peut montrer que ce ne serait pas un espace de Banach; de toute manière ce n'est pas cette norme qui va nous intéresser dans la suite.

Définition On appelle mesure de Radon sur l'espace localement compact X, une forme linéaire μ sur l'espace vectoriel $\mathcal{C}(X)$, dont la restriction à tout $\mathcal{C}_K(X)$, où K est un compact de X, est continue.

Voyons exactement ce que cela signifie.

Si μ est une mesure, alors elle définit une application qui, à toute fonction φ appartenant à $\mathcal{C}(X)$, c'est-à-dire à toute fonction scalaire φ définie sur X, continue, à support compact, fait correspondre un scalaire $\mu(\varphi)$. Cette correspondance doit être linéaire, en ce sens que l'on a (IV,2;1). D'autre part, nous n'avons pas dit que cette correspondance fût continue sur $\mathcal{C}(X)$, puisque nous n'avons pas mis de topologie ni de norme sur cet espace vectoriel. Nous avons dit seulement que sa restriction à tout $\mathcal{C}_K(X)$ était continue. Alors, quel que soit le compact K de X, et quelle que soit la suite $\varphi_0, \varphi_1, \dots, \varphi_n, \dots$ de fonctions continues ayant leur support dans K et convergeant uniformément vers 0 pour n tendant vers $+\infty$ la suite des nombres complexes $\mu(\varphi_n)$ converge vers 0. Cela revient aussi à dire qu'a tout compact K de X on peut faire correspondre une norme de la mesure relativement à K, à savoir $\|\mu\|_K$, telle que l'on ait la formule

$$|\mu(\varphi)| \leq \|\mu\|_K |||\varphi||| \text{ pour } \varphi \in \mathcal{C}_K(X); \quad \|\mu\|_K = \sup_{\substack{\varphi\in\mathcal{C}_K(X)\\ |||\varphi|||\leq 1}} |\mu(\varphi)| . \tag{IV,2;18}$$

Exemples de mesures de Radon

Exemple 1) La mesure $\delta_{(a)}$ définie à la formule (IV,2;5) est encore une mesure de Radon sur X ; de même la somme $\sum_\nu c_\nu \delta_{(a_\nu)}$ est une mesure de Radon, pourvu que la série $\sum_\nu |c_\nu|$ soit "localement convergente", c'est-à-dire que la somme partielle $\sum_{a_\nu \in K} |c_\nu|$ soit convergente, quel que soit le compact K de X. La formule (IV,2;6) définit bien en effet dans ce cas une forme linéaire sur $\mathscr{C}(X)$; d'autre part, pour tout compact K, on a la majoration $|\sum_\nu c_\nu \varphi(a_\nu)| \leq (\sum_{a_\nu \in K} |c_\nu|) \ |||\varphi|||$, pour

$\varphi \in \mathscr{C}_K(X)$, ce qui prouve que la condition de continuité est bien vérifiée; on a en outre $\|\mu\|_K \leq \sum_{a_\nu \in K} |c_\nu|$.

Mais on a en fait une estimation meilleure de la. norme. En effet, en chaque point a_ν de la frontière K de K , une fonction φ de $\mathscr{C}_K(X)$ est nécessairement nulle , (tout voisinage de ce point frontière contient des points du complémentaire de K où φ est nulle, et φ est continue (voir à ce sujet ce qui est dit à la page 431)); donc dans l'évaluation d'une somme $\sum c_\nu(a_\nu)$, les points a_ν de K n'interviennent pas, de sorte que l'on a l'inégalité $\|\mu\|_K \leq \sum_{a_\nu \in \overset{\circ}{K}} |c_\nu|$, $\overset{\circ}{K}$ étant l'intérieur de K . Il n'est alors pas difficile, en utilisant la méthode de la page 427, de montrer que la norme $\|\mu\|_K$ est exactement égale à $\sum_{a_\nu \in \overset{\circ}{K}} |c_\nu|$.

Exemple 2) Sur $\mathbb{R}$ la mesure $\int$, ou dx , se définit exactement de la même manière que dans le cas d'un intervalle compact $[a,b]$ de $\mathbb{R}$, par $\varphi \to \int \varphi$. On l'appelle la mesure canonique de $\mathbb{R}$, ou mesure de Lebesgue sur $\mathbb{R}$.

Pour tout intervalle $[a,b]$, sa norme $\|\mu\|_{[a,b]}$ est $b - a$.

Par ailleurs, si f est une fonction définie sur $\mathbb{R}$, à valeurs complexes, et localement intégrable-Riemann, c'est-à-dire intdgrable sur tout Intervalle borné, elle définit encore une mesure de Radon μ par la formule (IV,2;12), mesure appelée $f\,dx$ ou $f(x)\,dx$, dite de densité f par rapport A la mesure de Lebesgue.

La norme relative à l'intervalle compact $[a,b]$ est alors $\|\mu\|_{[a,b]} = \int_{[a,b]} |f(x)|\,dx$. Remarquons que cette affirmation ne résulte pas immédiatement de celle qui a été faite après la formule (IV,2;16). En effet les fonctions φ de $\mathscr{C}_{[a,b]}(\mathbb{R})$, contrairement à celles de $\mathscr{C}([a,b])$ (voir page 431) sont toutes nulles au point a et au point b.

Si toutefois on choisit $a', b', a < a' < b' < b$, de manière que l'on ait $\int_{[a,a']} |p(x)|\,dx + \int_{[b',b]} |\varphi(x)|\,dx \leqslant \frac{\varepsilon}{3}$,
et si φ est une fonction continue dans l'intervalle $[a',b']$ déterminée de manière que

$$\left| \int_{[a',b']} p(x)\varphi(x)\,dx - \int_{[a',b']} |p(x)|\,dx \right| \leqslant \frac{\varepsilon}{3},$$

suivant ce qui a été fait à la page 430, on pourra prolonger φ à l'intervalle $[a,b]$ de manière qu'elle soit affine dans les intervalles $[a,a']$, $[b',b]$ et qu'elle prenne la valeur 0 en a et b. On aura alors l inégalité suivante, en partageant l'intervalle d'intégration $[a,b]$ en réunion des 3 intervalles $[a,a']$, $[a',b']$, $[b',b]$:

$$\left| \int_{[a,b]} p(x)\varphi(x)\,dx - \int_{[a,b]} |p(x)|\,dx \right|$$
$$\leqslant \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

Ce qui prouve l'affirmation précédente. (On peut naturellement ensuite introduire des mesures du type $\sum_\nu c_\nu \delta_{(a_\nu)} + p\,dx$).

C'est Précisément le fait d'avoir obligé les fonctions φ à avoir toutes leur support compact, qui nous permet d'introduire des mesures du type précédent, où la fonction p peut avoir à l'infini une croissance arbitrairement rapide. Par exemple, p peut être la fonction e^{x^2} ou la fonction $e^{e^{x^2}}$, et elle définit bien une mesure de Radon $p\,dx$.

On appelle norme de la mesure μ, et on note $\|\mu\|$, la borne supérieure des $\|\mu\|_K$ pour tous les compacts K de X. C'est un nombre $\geqslant 0$, fini ou égal à $+\infty$ (et par conséquent, malgré son nom, ce n'est pas une norme). Il peut donc aussi se définir par la formule

$$\|\mu\| = \sup_{\substack{\varphi\in\mathscr{C}(X)\\ |||\varphi|||\leqslant 1}} |\mu(\varphi)|. \qquad (IV,2;19)$$

Bien qu'on ne puisse pas considérer exactement l'espace des mesures comme le dual de l'espace $\mathscr{C}(X)$, on le note quand même $\mathscr{C}'(X)$. C'est naturellement aussi un espace vectoriel, avec même définition pour la somme ou le produit par un scalaire que dans le cas où X est compact; mais ce n'est pas un espace vectoriel normé. Par contre, le sous-espace de $\mathscr{C}'(X)$ formé des μ telles que $\|\mu\|$ soit fini, est vectoriel normé. On voit même aisément que c'est un Banach.

Applications à la mécanique et à la physique

Il est courant en mécanique de considérer une distribution de masses dans l'espace, et d'autre part, en physique, de considérer une distribution de charges électriques. Les mesures de Radon, sur un espace affine E euclidien à trois dimensions, donnent de bons modèles pour ces distributions de masses et ces distributions de charges. Ainsi ce que l'on appelle en mécanique une masse ponctuellemplacée en un point a, ou en physique une charge électrique e placée en un point a, n'est autre que la mesure de Radon $m\delta_{(a)}$ ou $e\,\delta_{(a)}$.

Au contraire, si on considère en physique une distribution de masses ou de charges définie par une densité $\rho(x)$, cela veut dire que l'on considère la mesure de Radon $\rho(x)\,dx$, où dz est la mesure de volume tri-dlmensionnnelle sur E. Autrement dit, si l'on emploie trois coordonnées rectangulaires, c'est la mesure $\rho(x, y, z)\,dx\,dy\,dz$.

Considérons maintenant ce qu'on appelle, en mécanique ou en physique, la distribution de masses ou de oharges portée par une surface Σ fermée de classe C' dans l'espace euclidien à trois dimensions, et de 'densité superficielle" $\rho(x)$.

C'est la mesure de Radon définie par

(IV,2;20) $$\mu(\varphi) = \int_\Sigma \varphi(x)\,\rho(x)\,dS\;^*.$$

Il sera commode de la noter $\rho\,dS$. De la même maniére, la distribution de charges ou de masses portée par une courbe fermée L de classe C^1 et de densité linéaire ρ par rapport à l'arc ds, est définie par la formule

(IV,2;21) $$\mu(\varphi) = \int_L \varphi(x)\,\rho(x)\,ds\,,$$

et peut se noter+ ds.

* De telles "intégrales de surface" seront définies plus tard, au § 10.

Mesures vectorielles

Les mesures que nous venons de voir sont des mesures scalaires. Si maintenant $\vec{E}$ est un espace vectoriel normé, on appelle mesure $\vec{\mu}$ sur X à valeurs dans $\vec{E}$, une application linéaire de $\mathcal{C}(X)$ dans $\vec{E}$, dont la restriction à chaque $\mathcal{C}_K(X)$, K compact de X, est continue. Ici donc, pour $\varphi \in \mathcal{C}(X)$, $\vec{\mu}(\varphi)$ ou $\vec{\mu}\cdot\varphi$ est un élément de $\vec{E}$. Les normes $\|\vec{\mu}\|_K$, $\|\vec{\mu}\|$ se définissent de la même manière. Comme exemple, on pourra prendre, si $\vec{E}$ est complet, $\vec{\mu} = \sum_\nu \vec{c}_\nu\, \delta_{(a_\nu)}$, avec $\vec{c}_\nu \in \vec{E}$, ou $d\vec{\mu} = \vec{\mu}(x)\, dx$, où $\vec{\mu}$ est une fonction sur $\mathbb{R}$ à valeurs dans $\vec{E}$, localement intégrable-Riemann; on a bien alors $\sum_\nu \vec{c}_\nu\, \varphi(a_\nu) \in \vec{E}$ (en supposant, comme précédemment, que, pour tout compact K de X, $\sum_{a_\nu \in K} \|\vec{c}_\nu\| < \infty$, la série $\sum_\nu \vec{c}_\nu\, \varphi(a_\nu)$ converge d'après le théorème 55 du chapitre II) et $\int \vec{\mu}(x)\, \varphi(x)\, dx \in \vec{E}$. On a, pour les normes,

$$\Big\| \sum_\nu \vec{c}_\nu\, \delta_{(a_\nu)} \Big\|_K \leq \sum_{a_\nu \in K} \|\vec{c}_\nu\|\,, \text{ et } \Big\| \vec{\mu}\, dx \Big\|_{[a,b]} \leq \int_{[a,b]} \big\|\vec{\mu}(x)\big\|\, dx\,.^{*}$$

* Nous avons toujours écrit $\mathcal{C}(X)$ sans spécifier le corps des scalaires. Soit $\vec{E}$ un espace vectoriel normé sur le corps des complexes. Alors une $\mathbb{C}$-mesure $\vec{\mu}$ à valeurs dans $\vec{E}$, pour le corps des scalaires $\mathbb{K} = \mathbb{C}$ est une application de l'espace $\mathcal{C}(X;\mathbb{C})$ des fonctions φ complexes continues à support compact, dans $\vec{E}$, linéaire par rapport à $\mathbb{C}$. Une $\mathbb{R}$-mesure $\vec{\mu}$ à valeurs dans $\vec{E}$ pour le corps des scalaires $\mathbb{K} = \mathbb{R}$, est une application de l'espace $\mathcal{C}(X;\mathbb{R})$ des fonctions φ réelles continues à support compact, dans $\vec{E}$, linéaire par rapport à $\mathbb{R}$. Une mesure μ pour le corps $\mathbb{C}$ définit donc une mesure $\vec{\mu}_{\mathbb{R}}$ pour le corps $\mathbb{R}$. $\vec{\mu}_{\mathbb{R}}$ est simplement la restriction à $\mathcal{C}(X;\mathbb{R})$ de la fonction $\vec{\mu} : \mathcal{C}(X;\mathbb{C}) \longrightarrow \vec{E}$.

Par exemple, si $\vec{E} = \mathbb{C}$, μ est une mesure scalaire, mais $\vec{\mu}_{\mathbb{R}}$ est vectorielle, à valeurs dans l'espace vectoriel $\mathbb{C}$ de dimension 2 sur $\mathbb{R}$.

L'espace des $\mathbb{K}$-mesures scalaires (ou $\mathbb{K}$-dual de $\mathcal{C}(X;\mathbb{K})$) se notera $\mathcal{C}'(X;\mathbb{K})$

Plaçons nous toujours dans les mêmes conditions, $\vec{E}$ espace vectoriel normé sur $\mathbb{C}$ et soit $\vec{\nu}$ une mesure par rapport à $\mathbb{R}$, c'est-a-dire une application $\mathbb{R}$-liné-

(Voir suite du renvoi page suivante)

Partition de l'unité

Nous allons donner un théorème concernant les fonctions réelles continues sur un espace localement compact X, que nous aurons besoin d'utiliser de façon constante dans la théorie des mesures.

Théorème 11

Soit X un espace localement compact "dénombrable à l'infini", c'est-à-dire réunion d'une infinité dénombrable de compacts. Soit d'autre part $(\Omega_i)_{i\in I}$ ~~un~~ recouvrement (fini ou non) de X par des ouverts. Il existe un système de fonctions réelles continues α_i, dépendant du même ensemble d'indices I, tel que $0 \leq \alpha_i \leq 1$, que α_i ait son support dans Ω_i, que tout point de X ait un voisinage sur lequel un nombre fini seulement des fonctions α_i ~~ne soient pas idéntiquement nulles~~, et que la somme $\sum_{i\in I} \alpha_i$ soit identique à 1 sur X. Si X est une variété de classe C^m (m fini ou $= +\infty$) sur le corps des réels $\mathbb{R}$, ~~on peut~~ choisir les α_i de classe C^m. *

(Suite du renvoi de la page précédente)

aire de $\mathscr{C}(X.\mathbb{R})$ dans $\vec{E}$. Alors on peut, d'une manière unique, la prolonger en une mesure par rapport à $\mathbb{C}$, $\vec{\nu}_{\mathbb{C}}$ en posant $\vec{\nu}_{\mathbb{C}}(\varphi_1 + i\varphi_2) = \vec{\nu}(\varphi_1) + i\vec{\nu}(\varphi_2)$. On a $(\vec{\nu}_{\mathbb{C}})_{\mathbb{R}} = \vec{\nu}$, et $(\vec{\mu}_{\mathbb{R}})_{\mathbb{C}} = \vec{\mu}$.

Noter que les normas de $\vec{\mu}$ et $\vec{\mu}_{\mathbb{R}}$ (ou $\vec{\nu}$ et $\vec{\nu}_{\mathbb{C}}$) peuvent être distinctes. Prenons par exemple la mesure scalaire $\mu = \delta_{(a)} + i\,\delta_{(b)}$. Sa norme relativement au corps des scalaires $\mathbb{C}$ est $\operatorname{Sup}|\varphi(a) + i\varphi(b)|$, pour toutes les φ complexes majorées en module par 1 : c'est 2. Sa norme pour le corps des scalaires $\mathbb{R}$, est la même borne supérieure pour toutes les φ réelles majorées en module par 1 : c'est $\sqrt{2}$. On prendra presque toujours $\mathbb{K} = \mathbb{C}$

* ce perfectionnement (classe C^m) ne sert à rien dans le calcul intégral; mais nous en aurons besoin dans la théorie des distributions.

Nous supposons ici essentiellement qu'il s'agit d'une variété C^m par rapport au corps des réels; cesserait complètement faux par rapport au corps des complexes. On verra bien que le lemme 1 de la démonstration utilisera le corps des réels.

<u>Remarques</u> - 1°) Apparemment, la somme $\sum_{i \in I} \alpha_i$ pose un problème de convergence, d'autant plus ennuyeux que I n'est peut être pas dénombrable. Mais l'énoncé indique qu'en tout point x de X et même en tous les points d'un voisinage de x . cette somme est finie. tous les termes étant nuls sauf un nombre fini.

X étant localement compact, il revient au même de dire que, sur tout compact de X , toutes les α_i sont identiquement nulles sauf un nombre fini. On dit encore que le <u>système des α_i est localement fini</u>.

2°) On voit pourquoi les α_i définissent une partition de l'unité : on a décomposé 1 en somme de fonctions α_i , à supports dans des ouverts donnés Ω_i.

Si les $\overline{\Omega}_i$ sont compacts, les α_i sont à support compact, donc dans $\mathcal{C}(X)$. On dit que la partition de l'unité $(\alpha_i)_{i\in I}$ est <u>subordonnée</u> au recouvrement $(\Omega_i)_{i\in I}$.

3°) Si X est localement compact mais non dénombrable a l'infini, le résultat ne subsiste pas. Mais il subsiste un résultat moins fort, suffisant pour certaines applications, que nous ne donnerons pas ici (on se base sur les lemmes 1 et 2, qui sont toujours vrais pour X localement compact quelconque).

Un espace vectoriel normé de dimension finie est localement compact, et dénombrable à l'infini puisque réunion des boules $\|x\| \leq n$, $n \in \mathbb{N}$.

4°) Un espace X pour lequel le théorème de la partition de l'unité est vrai est dit <u>espace paracompact</u>. Un espace localement compact dénombrable à l'infini est paracompact. On démontre qu'un espace métrisable est paracompact.

<u>Démonstration</u> - Nous démontrerons d'abord trois lemmes.

<u>Lemme 1 - Soient Ω un ouvert d'un espace localement compact X , a un point de Ω . Il existe une fonction γ continue, vérifiant $0 \leq \gamma \leq 1$, de support compact contenu dans Ω, et telle que $\gamma(a) > 0$. Si X est une variété de classe C^m, on peut choisir γ de classe C^m . Naturellement ce lemme</u> est d'autant plus intéressant que le voisinage Ω de a est plus petit.

1°/ Soit X localement compact quelconque. Nous admettrons ce théorème dans le cas général et le démontrerons seulement si l'espace localement compact X est métrlsable. Supposons le donc muni d'une métrique. Ω étant un ouvert contenant a, et X étant localement compact, il existe une boule compacte $B(a,\eta)$ de centre a et de rayon $\eta > 0$, contenue dans Ω * Soit alors F une fonction réelle continue sur la demi-droite $\mathbb{R}_+$ des nombres réels $\geqslant 0$, vérifiant $0 \leqslant F \leqslant 1$, et telle que $F(0) > 0$ et $F(t) = 0$ pour $t \geqslant \eta$; on pourra, par exemple, prendre la fonction définie par

(IV,2;22)
$$F(t) = \begin{cases} 0 & \text{pour } t > \eta \\ \eta - t & \text{pour } 0 \leqslant t \leqslant \eta \end{cases}$$

Alors la fonction $x \longrightarrow F(d(a,x))$ répond à la question ; elle est en effet continue, comme composée de deux fonctions continues; elle prend la valeur η pour $x = a$, parce que $d(a,a) = 0$ et que $F(0) = \eta$; et, d'autre part, elle est nulle en dehors de la boule $B(a,\eta)$, puisque alors $d(a,x) > \eta$, et que $F(t) = 0$ pour $t \geqslant \eta$, donc son support est contenu dans cette boule, qui est un compact contenu dans Ω. On comprend pourquoi le fait que X soit un espace métrique est une simplification considérable du problème; en effet, la métrique donne une infinité de fonctions réelles continues sur l'espace métrique, à savoir les fonctions $x \longrightarrow d(a,x)$; au contraire, si l'on sait seulement que X est un espace topologique, on ne connait a priori aucune fonction réelle continue sur X, en dehors des fonctions constantes.

* En effet, a possède un voisinage compact V . Alors $\Omega \cap V$ est un voisinage de a, donc il existe $\eta > 0$ tel que la boule fermée $B(a;\eta)$ soit dans $\Omega \cap V$. elle est alors dans Ω, et, étant fermée et dans V compact, elle est compacte:

2°/ Supposons maintenant que X soit une variété V de dimension n sur le corps des réels, de classe C^m (éventuellement $m = +\infty$) *; Soit Φ une carte de V d'image contenant a ; Φ est un homéomorphisme de classe C^m d'un ouvert $\mathcal{O}_1$ de $\mathbb{R}^n$ sur

$$\Phi(\mathcal{O}_1) \subset V \quad ; \text{ soit } \Omega_1 = \Phi^{-1}(\Omega), \ a_1 = \Phi^{-1}(a).$$

Supposons le lemme démontré dans $\mathbb{R}^n$, avec une fonction de classe C^m. Autrement dit, supposons avoir trouvé une fonction γ_1 de classe C^m sur $\mathbb{R}^n$, à support compact K_1 contenu dans Ω_1, vérifiant $0 \leqslant \gamma_1 \leqslant 1$ et $\gamma_1(a_1) > 0$.

Alors le lemme est aussi ddmontré dans V. Définissons en effet γ sur V comme suit : $\gamma = \gamma_1 \circ \Phi^{-1}$ dans $\Phi(\Omega_1) \subset \Omega$, et $\gamma = 0$ ailleurs; Φ^{-1} est la bijection réciproque de Φ, application de $\Phi(\mathcal{O}_1)$ sur $\mathcal{O}_1$. La fonction γ est bien de classe C^m. En effet, soit $b \in V$ · ou bien $b \in \Phi(\Omega_1)$, mais alors γ est égale à $\gamma_1 \circ \Phi^{-1}$, et, comme Φ est une carte de classe C^m et γ_1 une fonction de classe C^m sur $\mathcal{O}_1$, γ est bien de classe C^m au voisinage de b (théorème 33 ter du chapitre III); ou bien $b \notin \Phi(\Omega_1)$, mais alors b est dans l'ouvert de V : $\complement_V \Phi(K_1)$, complémentaire du compact $\Phi(K_1)$, et γ est nulle dans cet ouvert, donc encore de classe C^m au voisinage de b. Nous venons de voir que γ est nulle dans le complémentaire du compact $\Phi(K_1)$, donc son support est dans le compact $\Phi(K_1) \subset \Phi(\Omega_1) \subset \Omega$, Elle vérifie bien $0 \leqslant \gamma \leqslant 1$, et $\gamma(a) = \gamma_1(a_1) > 0$.

Reste donc à résoudre le problème, avec γ_1 de classe C^m, sur $\mathbb{R}^n$; comme m peut être quelconque, nous devons trouver γ_1 de classe C^∞. Prenons sur $\mathbb{R}^n$ la métrique euclidienne habituelle; alors la fonction $x_1 \longrightarrow r = d(a_1, x_1)$ est continue, mais non dérivable au point a_1, mais son carré r^2, polynome du 2ème degré par rapport aux coordonnées. est de classe C^∞

Soit $\eta_1 > 0$ tel que la boule compacte $B(a_1; \eta_1)$ soit dans Ω_1. Si on considère, au lieu de F, la fonction réelle G sur $\mathbf{R}_+$, définie par :

$$G(s) = \begin{cases} 0 & \text{pour } s \geqslant \eta_1^2 \\ e^{-\frac{1}{\eta_1^2 - s}} & \text{pour } s < \eta_1^2, \end{cases} \tag{IV,2;23}$$

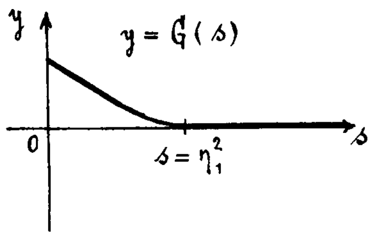
* Puisque tout point de la variété a un voisinage homéomorphe à un ouvert de $\mathbb{R}^n$, il a aussi un voisinage, plus petit, qui est compact; <u>une variété est un espace localement compact</u>.

on voit que $0 \leqslant G \leqslant 1$, que $G(s) > 0$, que $G(s)$ est nulle pour $s \geq \eta_1^2$, et que G est continue, car $\lim\limits_{\substack{s \to \eta_1^2 \\ s < \eta_1^2}} e^{-\frac{1}{\eta_1^2 - s}} = e^{-\infty} = 0$. Mais il y a plus :

<u>G est de classe C^∞</u> . En effet, toute dérivée de G , pour $s < \eta_1^2$, est, comme on le voit de proche en proche, de la forme

2;24) $$G^{(k)}(s) = P_k\left(\frac{1}{\eta_1^2 - s}\right) e^{-\frac{1}{\eta_1^2 - s}} ,$$

où P_k est un polynome a une variable. Lorsque s tend vers η_1^2 par valeurs $< \eta_1^2$, $G^{(k)}(s)$ se présente sous la forme $\infty \times e^{-\infty}$ mais il tend vers 0 , l'exponentielle l'emportant sur le polynome. Pour $A > \eta_1^2$, $G^{(k)}(s) \equiv 0$. Donc $G^{(k)}(s)$ tend vers 0 quand s tend vers η_1^2 par valeurs quelconques $\neq \eta_1^2$, et le théorème 14 du chapitre III, appliqué de proche en proche aux dérivées successives, montre bien que G est indéfiniment dérivable, avec toutes ses dérivées successives nulles pour $s = \eta_1^2$. Voici le graphique de la fonction G .



(Bien entendu, si on cherche le développement de Taylor d'ordre m de G au voisinage de $s = \eta_1^2$, il se réduit à son terme complémentaire. La <u>série de Taylor</u> de G , suivant les puissances de (A- η_1^2) , <u>est convergente</u>, puisque tous ses termes sont nuls, mais elle ne représente pas la fonction pour A < η_1^2).

Alors la fonction $\gamma_1(x_1) = G\left((d(a_1, x_1))^2\right) = G(r^2)$, est de classe C^∞, comme composée des 2 fonctions G et r^2, de classe C^∞, on a bien $\gamma_1(a_1) = G(0) > 0$; on a $0 \leqslant \gamma \leqslant 1$, puisque $0 \leqslant G \leqslant 1$; enfin γ_1 est encore nulle pour $d(a_1, x_1) \geqslant \eta_1$, puisque $G(s)$ est nulle pour $s \geqslant \eta_1^2$, et le lemme 1 est démontré.

Lemme 2 - Soit X un espace localement compact, et soit $(\Omega_i)_{i\in I}$ un recouvrement ouvert de X. Soit K un compact de X. Il existe un système de fonctions continues $\beta_i \geq 0$, dépendant du même ensemble d'indices I, toutes nulles sauf un nombre fini, telles que β_i ait un support compact et contenu dans Ω_i, et que la somme $\sum_{i\in I} \beta_i$ soit > 0 sur K. Si X est une variété de classe C^m, on peut choisir les β_i de classe C^m.

Soit ξ un point quelconque de K. Le point ξ appartient peut-être à plusieurs des ouverts Ω_i ; choisissons, n'importe comment, l'un d'entre eux, et appelons $i(\xi)$ l'indice auquel il correspond, de sorte que l'on a $\xi \in \Omega_{i(\xi)}$.

Soit alors $\gamma_\xi : x \longrightarrow \gamma_\xi(x)$, une fonction satisfaisant aux conditions du lemme 1, relativement à l'ouvert $\Omega_{i(\xi)}$ et au point ξ de cet ouvert. Appelons ω_ξ l'ensemble des points x où $\gamma_\xi(x) > 0$; alors les ω_ξ forment, lorsque ξ décrit K, un recouvrement ouvert de K. En effet, d'une part, ce sont des ouverts, et d'autre part $\xi \in \omega_\xi$, ce qui prouve bien que ces ouverts recouvrent K tout entier. Comme K est compact, il en existe un sous-recouvrement fini; autrement dit, il existe un nombre fini de points $\xi_1, \xi_2, \dots \xi_\ell$, tels que les ω_{ξ_ν}, $\nu = 1, 2, \dots \ell$, forment un recouvrement ouvert de K. Appelons alors β_i la fonction $\sum_{i(\xi_\nu)=i} \gamma_{\xi_\nu}$. Alors β_i est une fonction continue ≥ 0 ; elle est de classe C^m si X est une variété de classe C^m ; son support est contenu dans la réunion des supports des γ_{ξ_ν}, $i(\xi_\nu) = i$, c'est-à-dire compact et contenu dans Ω_i. Pour tout point x de K il existe au moins un des points ξ_ν tel que $x \in \omega_{\xi_\nu}$, puisque les ω_{ξ_ν} forment un recouvrement de K. La fonction γ_{ξ_ν} correspondante est > 0 au point x, donc la fonction β_i correspondante $(i = i(\xi_\nu))$ est > 0 au point x ; donc la fonction $\sum_{i\in I} \beta_i$ est > 0 sur K, ce qui démontre le lemme *

* Le support de $\sum_{i\in I} \beta_i = \beta$, contenu dans la réunion des supports des β_i, est un compact H, contenant l'ouvert U des points où $\beta > 0$, contenant lui-même K donc H est un voisinage compact de K. C'est là une propriété qu'on peut montrer directement : sur un espace localement compact X,

(Voir suite du renvoi page suivante)

Lemme 3. Soit X un espace localement compact, dénombrable à l' infini. Il existe une suite de compacts K_n et d'ouverts U_n d'adhérences compactes, tels que $K_n \subset \mathring{U}_n$ que les K_n forment un recouvrement de X , et sue tout compact ne rencontre qu'un nombre fini des U_n .

Puisque X est dénombrable à l'infini, il existe une suite A_n de compacts, de réunion X . Chaque A_n possède un voisinage compact B_n (note page 440); les intérieures $\mathring{B}_n \supset A_n$ ont pour réunion X . Posons

$C_n = B_0 \cup B_1 \cup B_2 \cup \ldots \cup B_n$; les C_n forment maintenant une suite croissante de compacts, et les $\mathring{C}_n \supset A_n$ ont pour réunion X . En outre, pour tout compact K de X , il existe un entier n tel que $\mathring{C}_n \supset K$; en effet la suite des $\mathring{C}_n \cap K$ est une suite croissante d'ouverts sur le compact K , de réunion K , donc l'un d'eux est déja K .

Posons ensuite $D_n = C_n - \mathring{C}_{n-1} = C_n \cap \complement \mathring{C}_{n-1}$. D_n est fermé (intersection de 2 fermé), contenu dans C_n compact, donc compact. La réunion des D_n est encore X ; si en effet $x \in X$, et si n est le premier entier tel que $x \in C_n$, on a $x \notin C_{n-1}$ donc $\notin \mathring{C}_{n-1}$ donc $x \in D_n$.

Soit alors $p(n)$ le plus petit entier p tel que, $\mathring{C}_p \supset C_n$; et soit $q(n)$ le plus grand entier q tel que $C_q \subset \mathring{C}_n$; posons $E_n = C_{p(n)} - \mathring{C}_{q(n)}$. E_n est compact pour la même raison que D_n plus haut. E_n est un voisinage de D_n ; soit en effet $x \in D_n$; alors $x \in C_n$ donc $x \in \mathring{C}_{p(n)}$;

(suite du renvoi de la page précédente)

tout compact K a un voisinage compact H , et même un système fondamental de voisinages compacts.

En effet, soit U un voisinage de K . Comme tout point x de K a un système fondamental de voisinages compacts dans X (page 74), il existe un ouvert Ω_x de X , contenant x , d'adhérence compacte $\overline{\Omega}_x$ dans U . Les Ω_x forment un recouvrement ouvert du compact K ; il existe un sous recouvrement fini $\Omega_1, \Omega_2, \ldots, \Omega_n$. Alors $\overline{\Omega}_1 \cup \overline{\Omega}_2 \cup \ldots \cup \overline{\Omega}_n$ est un voisinage compact de K contenu dans U

ensuite $x \notin \overset{\circ}{C}_{n-1}$ donc $x \in C_{q(n)}$; donc $x \in \overset{\circ}{C}_{p(n)} - C_{q(n)}$; donc $\overset{\circ}{C}_{p(n)} - C_{q(n)} = \overset{\circ}{C}_{p(n)} \cap \complement C_{q(n)}$ est un ouvert contenant D_n , et manifestement contenu dans E_n . Enfin la suite des E_n est localement finie.; car, si K est un compact, il existe un q tel que $C_q \supset K$ (et même $\overset{\circ}{C}_q \supset K$), puis un n tel que $\overset{\circ}{C}_{n-1} \supset C_q$; alors $q(n) \geq q$, donc, pour $m \geq n$, $q(m) \geq q$, $\overset{\circ}{C}_{q(m)} \supset C_q \supset K$, donc E_m ne rencontre pas K.

Alors les $K_n = D_n$ et $U_n = \overset{\circ}{E}_n$ vérifient les condition:; du lemme 3 * .

Démontrons maintenant le théorème 11, en utilisant ces K_n et U_n .

Pour tout-n , considérons le recouvrement du compact K_n de X par les ouverts $\Omega_i \cap U_n$. On peut lui appliquer le lemme 2 et trouver des fonctions $\beta_{i,n} \geq 0$ continues, de classe C^m si X est une variété de classe C^m , toutes nulles sauf un nombre fini, telles que $\beta_{i,n}$ ait son support dans $\Omega_i \cap U_n$ (donc compact puisque $\bar{U}_n$ est compact), et que la somme $\sum_{i \in I} \beta_{i,n}$ soit > 0 sur K_n . Appelons alors β_i la fonction $\sum_{n=0}^{\infty} \beta_{i,n}$. Il y a là une somme infinie, mais localement finie; si en effet K est un compact de X, il n'existe qu'un nombre fini de n tels que K rencontre U_n ; pour chacun de ces n il n'y a qu'un nombre fini de $\beta_{i,n}$ non nulles; pour tous les autres n , les fonctions $\beta_{i,n}$ de support dans U_n , sont nulles sur K . Il en résulte que β_i est une fonction continue sur X , et de classe C^m si X est une variété de classe C^m , puisque les propriétés de continuité et de différentiabilité en un point X peuvent se démontrer en considérant seulement ce qui se passe dans un voisinage de ce point, qu'on peut prendre compact, et qu'alors la somme qui donne β_i est finie dans ce voisinage. Le support de β_i est contenu dans l'adhérenc de la réunion des supports des $\beta_{i,n}$, $n = 0, 1, 2, \ldots$ mais cette réunion est localement finie, C'est-à-dire que seul un nombre fini de ces supports rencontre n'importe quel compact de X ;

* L'exemple d'un espace vectoriel normé X fera mieux comprendre cette construction. Prenons pour A, la boule $\|x\| \leq n$. Pour B_n , on pourra prendre la boule $\|x\| \leq n+1$. C_n sera B_n . $D_n = K_n$ sera la couronne $n \leq \|x\| \leq n+1$. Pour tout n , $p(n) = n+1$, $q(n) = n-2$. Alors E_n est la couronne $n-1 \leq \|x\| \leq n+2$; U_n est l'ouvert $n-1 < \|x\| < n+2$. Toute la construction dans ce cas était évidente.

d'où l'on déduit immédiatement que la réunion de ces supports est déjà fermée * et qu'elle est donc le support de β_i. Ce support est donc contenu dans Ω_i.

Si maintenant nous posons $\beta = \sum_{i\in I} \beta_i = \sum_{i,n} \beta_{i,n}$, la même démonstration prouve que β est continue, et de classe C^m si X est une variété de classe C^m. Tout point de X appartient a au moins un K_n et par conséquent en ce point $\sum_{i\in I} \beta_{i,n}$ est > 0 donc l'une au moins des fonctions $\beta_{i,n}$ est > 0 donc β est > 0 partout. Alors la fonction $\alpha_i = \beta_i/\beta$ est continue sur X, de classe C^m si X est une variété de classe C^m, et de support dans Ω_i. Le système des $\alpha_i \geq 0$ est localement fini puisqu'il en est ainsi du système des β_i, et la somme $\sum_{i\in I} \alpha_i$ vaut $\sum_{i\in I} \frac{\beta_i}{\beta} = 1$, ce qui démontre le théorème.

Corollaire 1 - Soient X un espace localement compact. dénombrable à l'infini, F un fermé de X, Ω un ouvert contenant F. Il existe une fonction réelle a sur X, continue et de classe C^m si X est une variété de classe C^m, de support contenu dans Ω, vérifiant $0 \leq \alpha \leq 1$, et égale à 1 sur un voisinage de F.

Il suffit en effet d'appliquer le théorème au recouvrement de X défini par Ω et $\complement F$. Si l'on appelle α et β les deux fonctions associées à ce recouvrement, on a $\alpha + \beta = 1$; le support de β est dans $\complement F$, son complémentaire est un ouvert contenant F, sur lequel β est nulle, donc $\alpha = 1$.

Remarquons que si F est compact on peut prendre Ω d'adhérence compacte d'après la note de la page 440; on obtient alors a à support compact. Dans ce cas d'ailleurs, on peut voir qu'il est inutile de supposer X dénombrable à l'infini, (voir remarque 3 après 1 énoncé du théorème 11).

* La réunion F d'une famille infinie F_n de fermés n'est pas fermée; mais elle l'est si cette famille est localement finie. Soit en effet $x \in X$ adhérent à la réunion F. Soit $\mathcal{V}$ un voisinage compact de x. Soit n un entier tel que $F_m \cap \mathcal{V} = \emptyset$ pour $m > n$. Alors x est déjà adhérent à $\bigcup_{m\leq n} F_m$, qui est fermé comme réunion finie de fermés, donc $x \in \bigcup_{m\leq n} F_m \subset F$, et F est bien fermé.

<u>Corollaire 2 - Soient X un espace localement compact dénombrable à l'infini, A et B deux parties fermées de X d'intersection vide. Alors il existe une fonction réelle continue α, de classe C^m si X est une variété de classe C^m, égale à 1 sur tout un voisinage de A et à 0 sur tout un voisinage de B</u>. Il suffit en effet d'appliquer le corollaire 1 à X, à l'ensemble fermé A, et à l'ouvert $\Omega = \complement B$; dire que α a son support dans $\complement B$ c'est dire qu'elle est égale à 0 sur tout un voisinage de B.

<u>Corollaire 3</u>

Soit $I = \{1, 2, \ldots n\}$

Appelons F'_i la réunion des F_j, $j \neq i$, et Ω_i le complémentaire de F'_i. L'intersection des F'_i est vide, donc la réunion des Ω_i est X, ils forment un recouvrement ouvert. Soit $(\alpha_i)_{i \in I}$ une partition de l'unité subordonnée. Alors le support de α_i est dans Ω_i, donc α_i est nulle dans un voisinage de F'_i ; comme la somme $\sum_i \alpha_i$ est 1, il en résulte qu'α_i vaut 1 dans un voisinage de F_i. Alors la fonction définie par:

$$\vec{f}(x) = \sum_{i \in I} \vec{e}_i \, \alpha_i(x), \quad \text{ou } \vec{f}_i = \sum_{i \in I} \vec{e}_i \, \alpha_i \tag{IV,2;25bis}$$

répond à la question. (On a bien l'inégalité

$$\|\vec{f}(x)\| \leqslant \left(\operatorname{Max}_{i \in I} \|\vec{e}_i\|\right) \sum_{i \in I} \alpha_i(x) = \operatorname{Max}_{i \in I} \|\vec{e}_i\| ,$$

donc

$$|||\vec{f}||| \leqslant \operatorname{Max} \|\vec{e}_i\|$$

; mais aussi $\geqslant$ puisque $\vec{f}$ vaut $\vec{e}_i$ sur F_i, donc $=$).

<u>Corollaire 4 - Soient X un espace localement compact dénombrable à l'infini, et F une partie fermée de X. Soit $\vec{g}$ une fonction définie sur $\vec{F}$, continue à valeur dans un espace de Babach $\vec{E}$. Supposons que tout point a de F possède un voisinage ouvert $\mathcal{V}_a$ dans X, tel que la restriction de $\vec{g}$ à $F \cap \mathcal{V}_a$ soit prolongeable à $\mathcal{V}_a$ en une fonc-</u>

tion continue $\vec{G}_a$ (resp. de classe C^m si X est une variété de classe C^m). Alors la fonction $\vec{g}$, définie sur F , est prolongeable a X en une fonction continue $\vec{G}$ (resp. de classe C^m) à valeur dans $\vec{E}$.

Démonstration - L'ensemble de tous les $\mathcal{V}_a$ et de $\complement F$ est un recouvrement ouvert de X . On peut lui faire correspondre une partition de l'unité subordonnée, formée par des fonctions α_a de support dans $\mathcal{V}_a$ et une fonction α_0 de support dans $\complement F$; $\left(\sum_{a\in F}\alpha_a\right) + \alpha_0 = 1$.

Considérons alors la fonction $\alpha_a \vec{G}_a$; elle est définie seulement dans $\mathcal{V}_a$, mais on peut la définir sur X tout entier en lui donnant la valeur 0 en dehors de $\mathcal{V}_a$. Elle reste alors continue (resp. de classe C^m); il suffit pour le voir de le vérifier au voisinage de tout point x de X; or, si x est dans $\mathcal{V}_a$, cette fonction coïncide avec $\alpha_a \vec{G}_a$ qui est continue (resp. de classe C^m) dans l'ouvert $\mathcal{V}_a$, et, si $x \notin \mathcal{V}_a$, elle coïncide avec 0 dans l'ouvert complémentaire du support de α_a . Si alors nous considérons la somme $\vec{G} = \sum_{a\in F} \alpha_a \vec{G}_a$, comme cette somme est localement finie, elle définit sur X une fonction continue (resp. de classe C^m). En tout point x de F, on a, pour tout $a \in F$, $\vec{G}_a(x) = \vec{g}(x)$, puisque chaque $\vec{G}_a$ prolonge $\vec{g}$; on a donc, pour tout x de F , $\sum_{a\in F}\alpha_a \vec{G}_a(x) = \vec{g}(x)\sum_{a\in F}\alpha_a(x)$; comme $\alpha_0(x)$ est nulle sur F , la somme vaut 1 , donc $\vec{G}(x) = \vec{g}(x)$, de sorte que la fonation $\vec{G}$ prolonge bien $\vec{g}$.

Remarquons que ce corollaire contient tous les précédents comme cas particuliers. Si par exemple nous considérons le corollaire 3, on voit qu'on n'a pas fait autre chose que prolonger à X la fonction $\vec{g}$ définie sur $\bigcup_i F_i$, égale à la constante $\vec{e}_i$ sur chaque F_i .

Corollaire 5 - Soient X une variété de classe C^m, $m \geq 1$, de dimension N , et V une sous-variété fermée de classe C^m de X , de dimension n . Soit $\vec{g}$ une fonction définie sur V, à valeur dans un Banach $\vec{E}$, et de classe C^m . Elle est prolongeable à X , en une fonction $\vec{G}$ de classeleur C^m à va dans $\vec{E}$.

Démonstration - 11 suffit de nous ramener aux conditions du précédent corollaire. Cr, si nous considérons un point a de V, on peut trouver un voisinage ouvert $\mathcal{V}_a$ de a dans X, et une carte Φ_a, C^m-difféomorphisme d'un ouvert $\mathcal{O}_a$ de $\mathbb{R}^N$ sur $\mathcal{V}_a$. On peut en outre supposer, si on écrit $\mathbb{R}^N$ sous forme à'un produit $\mathbb{R}^n \times \mathbb{R}^{N-n}$ et qu'on représente chacun de ses points x comme un couple (y,z), $y \in \mathbb{R}^n$, $z \in \mathbb{R}^{N-n}$, que l'application Φ_a amène l'intersection $\mathcal{O}_a \cap \mathbb{R}^n$ sur l'intersection $\mathcal{V}_a \cap V$ (corollaire 2bis du théorème 32 du chapitre III *)

Le problème de prolongement de $\vec{g}$ de $V \cap \mathcal{V}_a$ a $\mathcal{V}_a$ est alors ramené au problème de prolongement de la fonction $a = \vec{f} \circ \Phi_a$, définie sur $\mathcal{O}_a \cap \mathbb{R}^n$, à l'ouvert $\mathcal{O}_a$ ou tout au moins à un voisinage de $\alpha = \Phi_a^{-1}(a)$ dans cet ouvert. Or un tel prolongement est évident dans $\mathbb{R}^N$: il est défini par la fonction $(y, z) \to \vec{h}(y)$, c.q.f.d.

Corollaire 6 - Soit X un espace localement compact dénombrable à l'infini, et soient $F_1, F_2, \ldots, F_n$, des sous-ensembles fermés 2 a 2 disjoints. 11 existe des voisinages de $F_1, F_2, \ldots, F_n$, qui sont encore 2 a 2 disjoints.

Construisons en effet la fonction G du corollaire 3 à valeurs réelles, correspondant à des valeurs $e_i = i$, $i = 1, 2, \ldots, n$. Les voisinages ouverts $\mathcal{V}_i$ des F_i définies par $\mathcal{V}_i = \{x \in X ; i - \frac{1}{2} < G(x) < i + \frac{1}{2}\}$ répondent a la question.

On dit qu'un espace X ayant cette propriété est normal. Un espace localement compact dénombrable à l'infini est normal; plus généralement, un espace paracompact est normal; la réciproque n'est pas vraie.

* Ce corollaire a été démontré pour une sous-variété V de dimension n d'un espace affine de dimension N, disons $\mathbb{R}^N$. Ici X est une variété abstraite de dimension N V une sous-variété de dimension n de X Mais par une carte d'un voisinage de a dans X, on se ramène aussitôt à un ouvert de $\mathbb{R}^N$ et à une sous-variété de cet ouvert; c'est a cette situation qu'on applique le corollaire.

Corollaire 7 - Soit X un espace localement compact dénombrable à l'infini. Soit $(\Omega_i)_{i\in I}$ un recouvrement ouvert de X . Il existe un autre recouvrement $(\Omega'_i)_{i\in I}$ localement fini, dependant du même ensemble d'indices, et tel que $\bar{\Omega}'_i \subset \Omega_i$.

Appliquons en effet le théorème il lui-même, il suffit de prendre pour Ω'_i l'ensemble des x tels que $\alpha_i(x) > 0$. Ω'_i est bien ouvert; $\bar{\Omega}'_i$ est le support de α_i , donc il il est contenu dans Ω_i ; en tout point x de X , la somme $\sum\limits_{i\in I} \alpha_i(x)$ vaut 1 , donc l'une des $\alpha_i(x)$ est >0, et x est dans 1' Ω'_i correspondant : les Ω'_i forment bien encore un recouvrement, et il est localement fini puisque le système des α_i est localement fini.

Indiquons enfin sans démonstration le résultat suivant qui généralise partiellement le corollaire 5 (mais qui est relatif aux fonctions continues, non de classe C^m) :

Avant de donner le corollaire suivant, introduisons une notion nouvelle. Soit $\vec{F}$ un Danach. Soit $\vec{f}$ une fonction continue à support compact sur X , à valeurs dans $\vec{F}$. On dit qu'elle est décomposable si elle s'exprime sous la forme

$$\vec{f}(x) = \sum_{i=1}^{N} \vec{g}_i \, \varphi_i(x) \quad , \quad \text{ou} \quad \vec{f} = \sum_{i=1}^{N} \vec{g}_i \, \varphi_i \; ,$$

où les $\vec{g}_i$ sont des vecteurs constants de $\vec{F}$, et les φ_i des fonctions scalaires continues à support compact. L'ensemble des valeurs d'une telle fonction est un sous-espace vectoriel de dimension finie de $\vec{F}$, contenu dans celui qui est engendré par les $\vec{g}_i$. Réciproquement, soit $\vec{f}$ une fonction continue à support compact, prenant ses valeurs dans un sous-espace vectoriel $\vec{G}$ de dimension finie de dimension finie de $\vec{F}$. Soit $\vec{g}_1, \vec{g}_2, \dots, \vec{g}_N$, une base de $\vec{G}$. Pour tout $x \in X$, $\vec{f}(x)$ est un vecteur de $\vec{G}$, donc s'écrit d'une manière unique $\overrightarrow{f(x)} = \sum\limits_{i=1}^{N} \vec{g}_i \, \varphi_i(x)$; donc la fonction $\vec{f}$ est la somme $\sum\limits_{i=1}^{N} \vec{g}_i \, \varphi_i$; la fonction φ_i : X - @ . est composée de $\vec{f} : X \longrightarrow \vec{G}$ et de l'application continue $\vec{G} \longrightarrow \mathbb{C}$, qui, à chaque vecteur de $\mathcal{G}$, fait correspondre sa i-ième coordonnée suivant la base choisie, donc elle est continue; son support est contenu dans celui de $\vec{f}$, donc compact, et $\vec{f}$ est décomposable. On peut donc aussi appeler les fonctions continues à support compact décomposables: fonctions continues à support compact de <u>rang fini</u>. Leur ensemble est un sous-espace de l'espace vectoriel $\mathcal{C}(X;\vec{F})$ des fonctions continues à support compact à valeurs dans $\vec{F}$; on peut montrer qu'il est isomorphe au produit tensorlel de $\mathcal{C}(X)$ par $\vec{F}$, et il est noté couramment $\mathcal{C}(X)\otimes\vec{F}$. Il est identique à $\mathcal{C}(X;\vec{F})$ si $\vec{F}$ est de dimension finie.

<u>Corollaire 8</u> - <u>Soit K un compact de X, $\vec{f}$ une fonction continue sur K, à valeurs dans un Banach $\vec{F}$. Quel que soit $\varepsilon > 0$, et quel que soit le voisinage $\mathcal{V}$ de K, il sxiste une fonction continue $\vec{f}_\varepsilon$ sur X, à support compact $\subset \mathcal{V}$ décomposable, telle que $|||\vec{f}-\vec{f}_\varepsilon||| \leq \varepsilon$. On peut même la choisir de la forme $\vec{f}_\varepsilon = \sum_{i=1}^{N} \vec{g}_i\,\varphi_i$, où les φ_i sont $\geqslant 0$, et telle que l'on ait aussi l'inégalité</u>

$$\Big\| \,\|\vec{f}\| - \sum_{i=1}^{N} \|\vec{g}_i\|\,\varphi_i \Big\| \leq \varepsilon .$$

<u>Démonstration</u> - Pour tout a de K, il existe un voisinage ouvert $\mathcal{V}_a \subset \mathcal{V}$ de a dans X tel que, pour tout x de $\mathcal{V}_a$, on ait : $\|\vec{f}(x)-\vec{f}(a)\| \leq \varepsilon$.

Un nombre fini des $\mathcal{V}_a$ recouvre K, soit $\mathcal{V}_{a_1}, \mathcal{V}_{a_2},\dots,\mathcal{V}_{a_n}$. Soit $\alpha_1, \alpha_2,\dots,\alpha_N, \alpha_0$, une partition de l'unité de X relative au recouvrement par les $\mathcal{V}_{a_i}$ et $\complement K$. Alors la fonction décomposable $\vec{f}_\varepsilon = \sum_{i=1}^{N} \overrightarrow{f(a_i)}\,\alpha_i$ répond à la question. En effet $\vec{f}(x)\,\alpha_0(x)$ est nul quel que soit x, car $\alpha_0(x)$ n'est $\neq 0$ que dans $\complement K$ où $\vec{f}(x)$ est nul; donc on a $\vec{f}(x) = \sum_{i=1}^{N} \alpha_i(x)\,\overrightarrow{f(x)}$.

On a alors

$$\left\| \overrightarrow{f(x)} - \sum_{i=1}^{N} \vec{f}(a_i)\,\alpha_i(x) \right\| = \left\| \sum_{i=1}^{N} \left(\overrightarrow{f(x)} - \overrightarrow{f(a_i)} \right) \alpha_i(x) \right\|$$

$\leq \sum_{i=1}^{N} \varepsilon\,\alpha_i(x)^{*} \leq \varepsilon$. De même

$$\left| \left\| \overrightarrow{f(x)} \right\| - \sum \left\| \overrightarrow{f(a_i)} \right\| \alpha_i(x) \right| \leq \left| \sum_{i=1}^{N} \left(\left\| \overrightarrow{f(x)} \right\| - \left\| \overrightarrow{f(a_i)} \right\| \right) \alpha_i(x) \right|$$

$\leq \sum_{i=1}^{N} \varepsilon\,\alpha_i(x) \leq \varepsilon$.

Théorème 12' Si X est un espace métrique quelconque, Y une partie fermée de X , $\vec{g}$ une fonction continue sur Y à valeurs dans un espace vectoriel normé de dimension finie, il existe une fonction continue $\vec{G}$ sur X à valeurs dans $\vec{E}$, qui prolonge $\vec{g}$ (c'est-à-dire coïncide avec $\vec{g}$ sur Y), et vérifie :

(IV, 2;26) $$\sup_{x\in X} \| \vec{G}(x) \| = \sup_{y\in Y} \| \vec{g}(y) \| \leq \ldots$$

Si g est réelle, on peut choisir G réelle, avec en outre

(IV, 2, 27) $$\sup_{x\in X} G(x) = \sup_{y\in Y} g(y), \quad \inf_{x\in X} G(x) = \inf_{y\in Y} g(y)$$

(ces quantités pouvant être égales à $\pm\infty$) .

* $\left\| \overrightarrow{f(x)} - \overrightarrow{f(a_i)} \right\| \leq \varepsilon$ pour $x \in \mathcal{V}_{a_i}$, et $\alpha_i(x) = 0$ pour $x \notin \mathcal{V}_{a_i}$

Si X est localement compact et rdunion d'une infinité dénombrable de parties compactes, on a le même résultat avec $\vec{E}$ espace de Banach de dimension infinie.

Support d'une mesure de Radon

Intuitivement, le support d'une distribution de masses en mécanique, ou d'une distribution de charges électrinues en physique, est le plus petit ensemble fermé contenanttoute la masse ou toute la charge.

Définition. Soient X un espace localement compact, $\vec{\mu}$ une mesure de radon sur X, à valeurs dans un espace vectoriel normé $\vec{E}$, Ω un ouvert de X. On dit que $\vec{\mu}$ est nulle dans Ω, si, pour toute fonction φ de $\mathcal{C}(X)$ de support dans Ω, $\vec{\mu}(\varphi)$ est nulle * .

Intuitivement, cela signifie que, dans l'ouvert Ω, il n'y a pas de masse ou de charge.

De la même manière, on dit que deux mesures $\vec{\mu}_1$ et $\vec{\mu}_2$ sur X co'incident ou sont égales dans Ω, si la différence $\vec{\mu}_1 - \vec{\mu}_2$ est nulle dans Ω.

Theorème 13 - Soient X un espace localement compact, $(\Omega_i)_{i\in I}$ une famille quelconque (finie ou non) d'ouverts de cet espace, de réunion Ω. Si une mesure de radon $\vec{\mu}$ sur X est nulle sur chacun des Ω_i, alors elle est aussi nulle dans leur réunion Ω.

Démonstration - Soit φ une fonction de $\mathcal{C}(X)$, de support compact K dans Ω ; nous voulons démontrer que $\vec{\mu}(\varphi)$ est nulle; Comme K est compact, il suffit d'un nombre fini des Ω_i pour le recouvrir, soit par exemple $(\Omega_i)_{i\in J}$. Soit $(\alpha_i)_{i\in J}$ une partition de l'unité relative à cet ensemble fini d'ouverts et au compact K.

* Bien faire attention : le support de φ est un compact, Ω est un ouvert. On verra par la suite pourquoi il y a lieu de prendre Ω ouvert.

Puisque $\sum_{i\in J}\alpha_i = 1$ sur K, on a, sur X tout entier, l'identité $\varphi = \sum_{i\in J}\alpha_i\varphi$ (*). Alors, $\vec{\mu}$ étant linéaire, on a $\vec{\mu}(\varphi) = \sum_{i\in J}\vec{\mu}(\alpha_i\varphi)$. Or chacun des termes de la somme du deuxième membre est nul, car $\alpha_i\varphi$ a son support dans Ω_i et $\vec{\mu}$ est supposée nulle dans Ω_i, et par conséquent $\vec{\mu}(\varphi)$ est bien nulle aussi. On voit à quoi a servi la partition de l'unité : à décomposer φ, de support dans la réunion Ω des Ω_i, en somme d'un nombre fini de fonctions $\varphi_i = \alpha_i\varphi$, ayant leurs supports dans les Ω_i.

Corollaire - Soit $\vec{\mu}$ une mesure de Radon sur un espace localement compact X. Il existe un ouvert Ω plus grand que tous les autres, dans lequel $\vec{\mu}$ soit nulle.

En effet, considérons tous les ouverts dans lesquels $\vec{\mu}$ est nulle; leur réunion est encore un ouvert, et μ est nulle dans cet ouvert, d'après le théorème; c'est donc bien le plus grand ouvert dans lequel $\vec{\mu}$ soit nulle.

Définition - On appelle support d'une mesure de Radon $\vec{\mu}$ sur X l'ensemble fermé F complémentaire du plus grand ouvert Ω dans lequel $\vec{\mu}$ soit nulle. La seule mesure de support vide est la mesure nulle ($\vec{\mu}(\varphi) = \vec{0} \in \vec{E}$ pour toute φ).

Il résulte de cette définition que, si φ est une fonction de $\mathscr{C}(X)$ et $\vec{\mu}$ une mesure de Radon, et si le support de φ et le support de $\vec{\mu}$ sont sans point commun, c'est-à-dire si φ s'annule sur un voisinage du support de $\vec{\mu}$, alors $\vec{\mu}(\varphi) = \vec{0}$. En effet le support de φ est alors situé dans le complémentaire du support de $\vec{\mu}$, c'est-à-dire, par définition du support, dans un ouvert dans lequel $\vec{\mu}$ est nulle.

On en déduit aussi que, si deux fonctions φ_1 et φ_2 de $\mathscr{C}(X)$ sont égales sur tout un voisinage du support de $\vec{\mu}$, $\vec{\mu}(\varphi_1)$ et $\vec{\mu}(\varphi_2)$ sont égales; en effet $\varphi_1 - \varphi_2$ est nulle sur un ouvert contenant le support de $\vec{\mu}$, donc le support de $\varphi_1 - \varphi_2$ et le support de $\vec{\mu}$ ont une intersection vide, d'où résulte notre assertion. Mais on peut démontrer le résultat plus précis suivant :

(*) C'est vrai pour tout point x de K, puisqu'alors $\sum_{i\in J}\alpha_i(x) = 1$; et pour $x \notin K$, car alors les 2 membres sont nuls.

<u>Théorème 13 bis</u> - <u>Si φ s'annule sur le support de $\vec{\mu}$, alors $\vec{\mu}(\varphi) = 0$; si φ_1 et φ_2 sont égales sur le support de $\vec{\mu}$, alors $\vec{\mu}(\varphi_1) = \vec{\mu}(\varphi_2)$.</u>

<u>Démonstration</u> - Soient F le support de $\vec{\mu}$ et K le support de φ. soit $\varepsilon > 0$; comme $\vec{\mu}$ est continue sur $\mathcal{C}_K(X)$, il existe $\eta > 0$ tel que $\psi \in \mathcal{C}_K(X)$, $\|\psi\| \leq \eta$, entraîne $\|\vec{\mu}(\psi)\| \leq \varepsilon$. Comme alors φ est continue et nulle sur F, il existe un voisinage $\mathcal{U}$ de F dans lequel $|\varphi| \leq \eta$.

Soit α une fonction continue réelle sur X, $0 \leq \alpha \leq 1$, continue à support compact $\subset \mathcal{U}$, égale à 1 sur un voisinage du compact $K \cap F$ (corollaire 1 du théorème 11).

Considérons alors le support de la fonction $\alpha\varphi - \varphi = (\alpha-1)\varphi$. Tout d'abord il est contenu dans le support K de φ. Comme d'autre part $\alpha - 1$ est nulle sur un voisinage de $K \cap F$, il est contenu dans $\complement(K \cap F) = \complement K \cup \complement F$; étant dans K, il est donc dans $\complement F$ c'est-à-dire dans le complémentaire du support de $\vec{\mu}$. On a donc $\vec{\mu}(\alpha\varphi - \varphi) = \vec{0}$, ou $\vec{\mu}(\alpha\varphi) = \vec{\mu}(\varphi)$.

Mais maintenant on a $\alpha\varphi \in \mathcal{C}_K(X)$, et $\|\alpha\varphi\| \leq \eta$ (car $|\varphi| \leq \eta$ sur $\mathcal{U}$, et α a son support dans $\mathcal{U}$) ; donc $\|\vec{\mu}(\alpha\varphi)\| \leq \varepsilon$ et par suite $\|\vec{\mu}(\varphi)\| \leq \varepsilon$.

Comme ε est arbitraire, on en déduit bien $\vec{\mu}(\varphi) = \vec{0}$.

<u>Corollaire</u> - <u>Pour qu'une mesure $\vec{\mu}$ sur X ait son support contenu dans l'ensemble A formé d'un nombre fini de points $a_1, a_2, \dots, a_n$, il faut et il suffit qu'elle soit une combinaison linéaire de mesures de Dirac</u>

(IV, 2 ; 27 bis) $$\vec{\mu} = \vec{c}_1\,\delta_{(a_1)} + \vec{c}_2\,\delta_{(a_2)} + \cdots + \vec{c}_n\,\delta_{(a_n)}\,.$$

<u>Démonstration</u> - Tout d'abord il est évident que, si $\vec{\mu}$ a la forme indiquée, elle a certainement son support dans A car elle est nulle sur le complémentaire de A. Il nous faut montrer la réciproque. Soit donc $\vec{\mu}$ une mesure ayant son support dans A. Pour tout i, soit θ_i une fonction de $\mathcal{C}(X)$, égale à 1 au point a_i et à 0 aux points a_j, $j \neq i$. (corollaire 3 du théorème 11). Posons alors une fois pour

toutes $\vec{\mu}(\theta_i) = \vec{c}_i$. Alors, pour toute fonction φ , on peut écrire

(IV,2;27 ter) $$\varphi = \sum_{i=1}^{n} \varphi(a_i)\,\theta_i + \chi \;,$$

ou χ est une fonction nulle sur **A** . On déduit du théorème que $\vec{\mu}(\chi) = \vec{0}$, de sorte que l'on a

(IV,2;27quarto) $$\vec{\mu}(\varphi) = \sum_{i=1}^{n} \varphi(a_i)\,\vec{\mu}(\theta_i) = \sum_{i=1}^{n} \vec{c}_i(a_i) \;,$$

ce qui démontre le corollaire.

Théorème 14 - Soit $\vec{\mu}$ une mesure de Radon sur un espace localement compact X . Un point a de X appartient au support F de $\vec{\mu}$, si et seulement si, quel que soit l'ouvert ω de X contenant a , la mesure $\vec{\mu}$ n'est a s

Le point a n'appartient pas au support F , si et seulement s'il existe un ouvert ω contenant a dans lequel $\vec{\mu}$ est nulle.

Démonstration - L'équivalence de $A \Rightarrow B$ et $\not{B} \Rightarrow \not{A}$ (théorème 11 du chapitre 1) montre que les 2 parties du théorème sont équivalentes. Montrons la deuxième. Le point a n'appartient pas à F , si et seulement s'il est dans $\Omega = \complement F$, et, comme Ω est le plus grand ouvert où $\vec{\mu} = \vec{0}$, cela revient à dire : si et seulement s'il existe un ouvert ω contenant a, où $\vec{\mu} = \vec{0}$

Exemple 1/ - Soit $\vec{\mu} = \sum_\nu \vec{c}_\nu \,\delta_{(a_\nu)}$ une mesure de Radon atomique; les a_ν sont supposés distincts, et, pour tout compact K de X la somme $\sum_{a_\nu \in K} \|\vec{c}_\nu\|$ est supposee finie (voir page 433). Le support de cette mesure de Radon est l'adhérence de la réunion des a_ν .

Démonstration - Appelons en effet **A** la réunion des a_ν, et $\bar{A}$ son adhérence. Si Ω est le complémentaire de $\bar{A}$, il est bien évident que, si φ a son support dans Ω, on a $\vec{\mu}(\varphi) = \vec{0}$; donc nécessairement le support F de $\vec{\mu}$ est contenu dans $\bar{A}$. Il nous reste à démontrer que chaque point a_j de A appartient effectivement au support F ; on aura donc $F \supset A$, et par suite, comme F est fermé, $F \supset \bar{A}$, ce qui montrera bien finalement que $F = \bar{A}$. Soit donc $a_j \in A$, soit $\mathcal{V}$ un voisinage compact de a_j. Alors la série $\sum_{a_\nu \in \mathcal{V}} \| \vec{c}_\nu \|$ est convergente; il existe donc un sous-ensemble fini J de l'ensemble d'indices I, tel que $\sum_{\substack{a_\nu \in \mathcal{V} \\ \nu \notin J}} \|\vec{c}_\nu\| \leq \frac{1}{2} \|\vec{c}_\nu\|$. Comme J est fini, il existe un ouvert ω contenant a_j, contenu dans $\mathcal{V}$, et ne contenant aucun autre des a_k pour $k \in J$; alors on a la majoration suivante :

$$\sum_{\substack{a_\nu \in \omega \\ \nu \neq j}} \|\vec{c}_\nu\| \leq \frac{1}{2} \|\vec{c}_j\| . \tag{IV,2;28}$$

Soit maintenant Ω un voisinage ouvert quelconque de a_j.

Montrons qu'il est impossible que $\vec{\mu}$ soit nulle dans Ω, et cela montrera bien que a_j appartient au support F de $\vec{\mu}$. Or il existe, d'après le lemme 1 du théorème 11, une fonction réelle φ continue, de support compact dans $\omega \cap \Omega$, vérifiant $0 \leq \varphi \leq 1$ et telle que $\varphi(a_j) = 1$.

Alors on a les relations

$$\vec{\mu}(\varphi) = \vec{c}_j + \sum_{\nu \neq j} \vec{c}_\nu \ \varphi(a_\nu) \quad , \text{donc} \tag{IV,2;29}$$

$$\|\vec{\mu}(\varphi)\| \geq \|\vec{c}_j\| - \sum_{\substack{a_\nu \in \omega \\ \nu \neq j}} \|\vec{c}_\nu\| \geq \frac{1}{2} \|\vec{c}_j\| ,$$

donc $\vec{\mu}(\varphi) \neq \vec{0}$, ce qui démontre la propriété.

Exemple 2/

Si $\vec{f}$ est une fonction intégrable Riemann sur $\mathbb{R}$ a valeurs dans l'espace de Banach $\vec{E}$, le support de la mesure $\vec{f}\,dx$, de densité $\vec{f}$ par rapport à la mesure dx, est contenu dans le support de la fonction $\vec{f}$ et il est exactement le Support de $\vec{f}$ si $\vec{f}$ est continue.

Cette propriété établit une liaison entre la notion de support d'une fonction, et de support d'une mesure.

Tout d'abord, comme dans le theorème précédent, il est évident que le support de $\vec{\mu}$ est contenu dans le support de $\vec{f}$: en effet, si φ a son support dans le complémentaire du support de $\vec{f}$, l'intégrale $\int \varphi\, \vec{f}\, dx$ est manifestement nulle. Démontrons qu'il est exactement le support de $\vec{f}$, si $\vec{f}$ est continue. Soit donc a un point tel que $\vec{f}(a) \neq 0$. Nous allons montrer que a appartient nécessairement au support de $\vec{\mu}$. Il en résultera que le support de $\vec{\mu}$ contient tous les points x où $\vec{f}(x) \neq 0$, et par conséquent aussi l'adhérence de l'ensemble de ces points, c'est-à-dire précisément le support de $\vec{f}$, et par conséquent le support de $\vec{\mu}$ sera identique au support de $\vec{f}$. Comme $\vec{f}$ est continue au point a, il existe un ouvert ω contenant a, tel que l'on ait, pour tout x de ω, l'inégalité

$$\left\|\vec{f}(x) - \vec{f}(a)\right\| \leq \frac{1}{2}\left\|\vec{f}(a)\right\| . \tag{IV,2;30}$$

On voit alors que, quel que soit l'ouvert Ω contenant a, il est impossible que $\vec{\mu}$ soit nulle dans Ω. Choisissons en effet une fonction quelconque $\varphi \geq 0$ de $\mathcal{C}(X)$, de support dans $\omega \cap \Omega$. On a alors les relations

$$\left\|\int\left(\vec{f}(x) - \vec{f}(a)\right)\varphi(x)\,dx\right\| \leq \frac{1}{2}\left\|\vec{f}(a)\right\|\int\varphi \quad . \tag{IV,2;31}$$ Donc

$$\left\|\int\vec{f}(x)\,\varphi(x)\,dx\right\| \geq \left\|\int\vec{f}(a)\,\varphi(x)\,dx\right\| - \frac{1}{2}\left\|\vec{f}(a)\right\|\int\varphi = \frac{1}{2}\left\|\vec{f}(a)\right\|\int\varphi > 0, \tag{IV,2;32}$$

qui prouve bien que $\vec{\mu}(\varphi) \neq 0$, et par conséquent $\vec{\mu}$ n'est pas nulle dans Ω.

Remarque Si $\vec{f}$ n'est pas continue, le support de $\vec{f}\,dx$ n'est plus nécessairement celui de $\vec{f}$. Par exemple, si $\vec{f}$ est partout nulle, sauf au point a de $\mathbb{R}$, le support de $\vec{f}$ est $\{a\}$; mais $\vec{f}\,dx$ est nulle et son support est vide.

Exemple 3/ Si $\vec{\mu}$ est la mesure de Radon définie sur $\mathbb{R}$ par la formule $\vec{\mu} = \sum_\nu \vec{c}_\nu \,\delta_{(a_\nu)} + \vec{f}\,dx$, où les points a_ν sont tous distincts, la somme $\sum_\nu \|\vec{c}_\nu\|$ localement convergente, et la fonction $\vec{f}$ continue, alors le support de $\vec{\mu}$ est la réunion A du support de $\vec{f}$ et de l'adhérence de l'ensemble des a_ν .

Ce résultat se démontre très simplement à partir de ceux qui précèdent. Comme il est évident que le support de $\vec{\mu}$ est contenu dans l'ensemble indiqué, il suffit successivement de montrer que ce support contient nécessairement chacun des points a,, et d autre part chaque point a tel que $\vec{f}(a) \neq \vec{0}$ et n'appartenant pas à l'adhérence des a_ν. Considérons donc un point a_j , on répètera alors exactement la construction de l'exemple 1, en choisissant pour ω un intervalle ouvert vérifiant en outre

$$\int_\omega \|\vec{f}(x)\|\,dx \leq \frac{1}{4}\|c_j\| .$$

On construira ensuite, pour tout voisinage ouvert Ω de a_j , la même fonction φ que dans l'exemple 1, mais on aura cette fois

$$\vec{\mu}(\varphi) = \vec{c}_j + \sum_{\nu\neq j} \vec{c}_\nu\,\varphi(a_\nu) + \int_\omega \vec{f}(x)\,\varphi(x)\,dx ,$$

donc

$$\|\vec{\mu}(\varphi)\| \geq \|\vec{c}_j\| - \frac{1}{2}\|\vec{c}_j\| - \frac{1}{4}\|\vec{c}_j\| = \frac{1}{4}\|\vec{c}_j\| ,$$

ce qui démontre la première affirmation.

Considérons ensuite un point a tel que $\vec{f}(a) \neq \vec{0}$, et que a n'appartienne pas à l'adhérence de l'ensemble des a_ν . On construira alors le même ouvert ω que dans l'exemple 2, mais en outre de manière qu'il ne contienne aucun des points a_ν . La démonstration de l'exemple 2 se poursuit alors sans modification.

En réunissant les résultats indiqués page 430 aux exemples 1,2,3 précédents, on a donc le théorème suivant :

Théorème 15 -**Soient** a_ν des points distincts de X , $\vec{c}_\nu$ des vecteurs de l'espace de Banach $\vec{E}$, tels que, pour tout compact K de X , la somme $\sum_{a_\nu \in K} \|\vec{c}_\nu\|$ soit finie. Alors on peut **définir** une mesure $\vec{\mu} = \sum_\nu \vec{c}_\nu \, \delta_{(a_\nu)}$ **sur** X , à valeurs dans $\vec{E}$, par

(IV,2;32 bis)
$$\Big(\sum_\nu \vec{c}_\nu \, \delta_{(a_\nu)}\Big) \cdot \varphi = \sum_\nu \vec{c}_\nu \, \varphi(a_\nu) .$$

Pour tout compact K, on a l'inégalité

(IV,2;32 ter)
$$\Big\| \sum_\nu \vec{c}_\nu \, \delta_{(a_\nu)} \Big\|_K \leq \sum_{a_\nu \in K} \| \vec{c}_\nu \| .$$

En outre, on a

(IV,2;32 quarto)
$$\Big\| \sum_\nu \vec{c}_\nu \, \delta_{(a_\nu)} \Big\| \leq \sum_\nu \| \vec{c}_\nu \| ;$$

ces inégalités deviennent des égalités si $\vec{E}$ est le corps des scalaires. *

* Par contre, il n'en est **rien si** $\vec{E}$ **est** un espace vectoriel de dimension ≥ 2 . Prenons par exemple $\vec{E} = \mathbb{R}^n$, muni de la norme de Minkowski

$$\|(x_1 , x_2 , . . . , x_n)\|_p = \Big(\sum |x_i|^p\Big)^{\frac{1}{p}} , \qquad 1 < p \leq +\infty .$$

Soient $a_1 , a_2 , \dots , a_n$, n points de X compact, et soit $\vec{c}_\nu$ le point de $\mathbb{R}^n$ dont la ν- ième coordonnée est 1 toutes les autres nulles (les $\vec{c}_\nu$ sont la base canonique de $\mathbb{R}^n$). Alors $\|\vec{\mu}(\varphi)\| = \| \sum \varphi(a_\nu) \vec{c}_\nu \| = \Big(\sum_{\nu=1}^{n} |\varphi(a_\nu)|^p\Big)^{\frac{1}{p}}$; la borne supérieure de cette **expression** pour $\|\varphi\| \leq 1$ est obtenuepour $\varphi = 1$, c'est $n^{\frac{1}{p}}$; or $\sum_{\nu=1}^{n} \|\vec{c}_\nu\| = n$; on a $n^{\frac{1}{p}} < n$ si $n \geq 2$.

Si X est la droite $\mathbb{R}$, et si $\vec{f}$ est une fonction sur $\mathbb{R}$ à valeurs dans $\vec{E}$ localement - - ; $\vec{f}\,dx$ à valeurs dans E, par

$$\vec{f}\,dx \cdot \varphi = \int \vec{f}(x)\,\varphi(x)\,dx \tag{IV,2;32 quinto}$$

Pour tout intervalle $[a,b]$ de $\mathbb{R}$, on a :

$$\left\| \vec{f}\,dx \right\|_{[a,b]} \leq \int_{[a,b]} \left\| \vec{f}\,dx \right\| dx \;, \tag{IV,2;32 sexto}$$

l'inégalité étant une égalité si $\vec{E}$ est le corps des scalaires.

Pour la mesure somme, on a, pour tout intervalle compact $[a,b]$ de $\mathbb{R}$:

$$\left\| \sum_{\nu} \vec{c}_{\nu}\,\delta_{(a_{\nu})} + \vec{f}\,dx \right\|_{[a,b]} \tag{IV,2;32 septimo}$$

$$\leq \sum_{a_{\nu}\in[a,b]} \left\| \vec{c}_{\nu} \right\| + \int_{[a,b]} \left\| \vec{f}(x) \right\| dx \;,$$

l'inégalité devenant une égalité si $\vec{E}$ est le corps des scalaires.

Le support de la mesure $\sum \vec{c}_{\nu}\,\delta_{(a_{\nu})}$ est l'adhérence de la réunion des a_{ν} ; le support de la mesure $\sum_{\nu} \vec{c}_{\nu}\,\delta_{(a_{\nu})} + \vec{f}\,dx$, si $\vec{f}$ est continue, est la réunion de l'adhérence de l'ensemble des a_{ν} et du support de $\vec{f}$.

Prolongement d'une mesure à des fonctions continues φ de support non compact

On a l'intuition qu'on peut donner un sens à $\vec{\mu}(\varphi)$, si $\vec{\mu}$ est une mesure de Radon de support A, φ une fonction scalaire continue de support B, même si B n'est pas compact, pourvu que A et B aient une intersection compacte K.

Soit en effet α une fonction de $\mathscr{C}(X)$, égale à 1 sur tout un voisinage de K (de telles fonctions existent d'après le corollaire 1 du théorème 11). Alors $\alpha\varphi$ a son support dans l'intersection des supports de α et de φ, donc compact.

Nous allons poser, par définition :

(IV,2;33) $$\vec{\mu}(\varphi) = \vec{\mu}(\alpha\varphi),$$

qui possède un sens parce que $\alpha\varphi \in \mathscr{C}(X)$. Pour justifier cette définition, nous devons seulement montrer que le 2ème membre est indépendant du choix de α. Or, si α_1 et α_2 sont deux fonctions égales à 1 sur un voisinage de **K**, $(\alpha_1\varphi - \alpha_2\varphi)$ et $\vec{\mu}$ ont des supports d'intersection vide (si un point appartient au support de $\vec{\mu}$, il est dans A ; s'il appartient au support de $(\alpha_1 - \alpha_2)\varphi$, il est dans le support de φ, donc dans B, et dans celui de $\alpha_1 - \alpha_2$, donc dans $\complement K$; l'intersection est donc dans $A \cap B \cap \complement K = K \cap \complement K = \emptyset$) ; alors on a bien $\vec{\mu}(\alpha_1\varphi - \alpha_2\varphi) = \vec{0}$, ou $\vec{\mu}(\alpha_1\varphi) = \vec{\mu}(\alpha_2\varphi)$.

En particulier, si $\vec{\mu}$ est à support compact, on peut définir $\vec{\mu}(\varphi)$ pour φ continue à support quelconque, en posant $\vec{\mu}(\varphi) = \vec{\mu}(\alpha\varphi)$, où $\alpha \in \mathscr{C}(X)$ est égale à 1 sur un voisinage du support de $\vec{\mu}$.

Par exemple, si $\mu = \delta_{(a)}$, on trouve que, pour φ continue scalaire à support quelconque, on a toujours (IV, 2;5). On conçoit même que, dans le cas de $\mu = \delta_{(a)}$, on puisse calculer $\mu(\varphi)$ pour des fonctions φ discontinues; c'est ce que nous ferons au §3.

Si $\vec{\mu}$ est à support compact, on appelle "masse totale" de $\vec{\mu}$ le vecteur $\vec{\mu}(1)$ de **E**. Par exemple, si $\vec{E}$ est complet, $\vec{\mu} = \sum_\nu \vec{c}_\nu \, \delta_{(a_\nu)}$, où les a_ν sont contenus dans un même compact et où $\sum_\nu \|c_\nu\| < \infty$, la masse totale est $\sum_\nu \vec{c}_\nu \in \vec{E}$. Si $X = \mathbb{R}$, si $d\vec{\mu} = \vec{\mu}\,dx$, où $\vec{\mu}$ a un support compact, la masse totale est $\int \vec{\mu}(x)\,dx \in \vec{E}$.
On voit d'où vient l'expression "masse totale"; si μ est à valeurs réelles et représente une distribution de charges en physique, c'est bien la "charge totale" (charge algébrique, naturellement : la masse totale de $\delta_{(a)} - \delta_{(b)}$ est nulle).

Le prolongement de $\vec{\mu}$ que nous venons de définir a les propriétés évidentes suivantes :

Thkorème **16** - Soient A et **B** des ensembles fermés de X, d'intersection compacte K. Pour une mesure de Radon $\vec{\mu}$ sur X à valeurs dans un space vectoriel normé $\vec{E}$, de support dans **A**, et une fonction scalaire φ continue de support dans B, on peut définir $\vec{\mu}(\varphi) \in \vec{E}$ par (IV,2; 33) où α est une fonction appartenant à $\mathscr{C}(X)$, égale à 1 sur un voisinage de K ; le résultat est indépendant du choix de α. Si φ_1 et φ_2 ont les mêmes propriétés que φ, et si k est un scalaire, on a $\vec{\mu}(\varphi_1+\varphi_2)=\vec{\mu}(\varphi_1)+\vec{\mu}(\varphi_2)$, $\vec{\mu}(k\varphi)=k\vec{\mu}(\varphi)$. Si des fonctions continues φ_n, de support dans B, convergent uniformement vers 0 pour n tendant vers l'infini, les $\vec{\mu}(\varphi_n) \in \vec{E}$ convergent vers $\vec{0}$. On a toujours la majoration

(IV, 2 ; 33bis)
$$\|\vec{\mu}(\varphi)\| \leqslant \|\vec{\mu}\| \sup_{x\in X} |\varphi(x)|$$

Remarque Dans l'énoncé, il suffit même qu'il existe un voisinage de K sur lequel les φ_n convergent uniformément vers 0 pour n infini, pour que les $\vec{\mu}(\varphi_n)$ convergent vers $\vec{0}$.

Principe du recollement des morceaux de mesures

Théorème 17 - Soit X un espace localement compact, $(\Omega_i)_{i\in I}$ une famille quelconque d'ouverts formant un recouvrement de X. Soit $(\vec{\mu}_i)_{i\in I}$ une famille de mesures, dépendant du même ensemble d'indices; on suppose que $\vec{\mu}_i$ est une mesure définie sur l'espace localement compact Ω_i * à valeurs dans un même espace vectoriel normé $\vec{F}$; et que, dans toute intersection $\Omega_i \cap \Omega_j$ de deux des ouverts, les mesures $\vec{\mu}_i$ et $\vec{\mu}_j$ sont égales. Alors il existe sur X une mesure $\vec{\mu}$ et une seule, à valeurs dans $\vec{E}$, qui, dans chaque ouvert Ω_i, soit égale à la mesure $\vec{\mu}_i$.

Démonstration - L'unicité de $\vec{\mu}$ est évidente : si $\vec{\mu}_1$ et $\vec{\mu}_2$ sont deux mesures avant la propriété indiquée, $\vec{\mu}_1 - \vec{\mu}_2$ est nulle dans chaque Ω_i, donc, d'après le théorème 13, dans leur réunion, qui est X ; donc $\vec{\mu}_1=\vec{\mu}_2$.

* Un ouvert Ω d'un espace localement compact X est aussi localement compact.

Pour démontrer l'existence, bien qu'il soit possible de le faire en utilisant un théorème 11 modifié (voir remarque **3")** sans aucune hypothèse de dénombrabilité sur **X** , nous supposerons, pour simplifier, que X est réunion d'une infinité dénombrable de compacts, et appliquerons la partition de l'unité.

Soit donc $(\alpha_i)_{i\in I}$ une partition continue de l'unité subordonnée au recouvrement $(\Omega_i)_{i\in I}$ de X. Soit φ une fonction de $\mathscr{C}(X)$. La fonction $\alpha_i\varphi$ appartient à $\mathscr{C}(\Omega_i)$; on peut donc lui appliquer la mesure $\vec{\mu}_i$ et considérer $\vec{\mu}_i(\alpha_i\varphi)$. Définissons alors la mesure $\vec{\mu}$ par la formule

$$\vec{\mu}(\varphi) = \sum_{i\in I} \vec{\mu}_i(\alpha_i\varphi) . \tag{IV,2;33 ter}$$

Tout d'abord cette formule possède un sens; en effet le support de φ est un compact, et sur ce compact toutes les fonctions α_i sont nulles sauf un nombre fini, de sorte que la somme précédente est finie. Elle définit évidemment une application linéaire de $\mathscr{C}(X)$ dans le Banach $\vec{F}$. Cette application est trivialement continue sur $\mathscr{C}_K(X)$; on a en effet l'inégalité :

$$\|\vec{\mu}(\varphi)\| \leqslant \sum_{i\in I} \|\vec{\mu}_i\|_K \|\alpha_i\varphi\| \leqslant \left(\sum_{i\in J}\|\vec{\mu}_i\|_K\right)\|\varphi\| , \tag{IV,2;33 quarto}$$

où J est l'ensemble fini des i tels que α_i ne soit pas identiquement nulle sur **K.**

Elle définit donc une mesure sur X à valeur dans $\vec{E}$. Il nous reste à démontrer que l'on a $\vec{\mu} = \vec{\mu}_j$ dans Ω_j . Or, si φ a son support compact dans Ω_j , celui de $\alpha_i\varphi$ est, pour tout i , contenu dans $\Omega_i \cap \Omega_j$; comme, dans cet ouvert, on a $\vec{\mu}_j = \vec{\mu}_i$, on peut remplacer, dans la formule (IV,2;33 ter)

$$\vec{\mu}_i(\alpha_i\varphi) \quad \text{par} \quad \vec{\mu}_j(\alpha_i\varphi) ,$$

et on obtient

$$\vec{\mu}(\varphi) = \sum_{i\in I}\vec{\mu}_j(\alpha_i\varphi) = \vec{\mu}_j\left(\sum_{i\in I}\alpha_i\varphi\right) = \vec{\mu}_j(\varphi) , \tag{IV,2;33 quinto}$$

ce qui prouve ce théorème.

Mesures complexes, mesures réelles

Définition Nous n'avons pas, jusqu'à présent, fait spécialement attention au corps des scalaires $\mathbb{K}$, qui pouvait être indifféremment $\mathbb{R}$ ou $\mathbb{C}$. Nous appelions $\mathscr{C}(X)$ l'espace des fonctions scalaires sur X, continues à support compact; nous traitions par là les 2 cas à la fois, mais il est évident que, dans certains problémes, on doit fixer de quel corps il s'agit.

Soit alors $\mathscr{C}(X;\mathbb{C}) = \mathscr{C}(X)$ l'espace des fonctions φ complexes, et soit μ une mesure complexe sur X, $\mu \in \mathscr{C}'(X)$.

un dit que μ est réelle, si $\mu(\varphi)$ est réel , pour toute fonction φ réelle. *

Pour qu'une mesure de Radon μ soit réelle, il faut et il suffit que., pour toute fonctinn complexe φ de $\mathscr{C}(X)$, on ait l'égalité

(IV,2;34) $$\mu(\overline{\varphi}) = \overline{\mu(\varphi)},$$

autrement dit que $\mu(\varphi)$ et $\mu(\overline{\varphi})$ soient complexes conjugués.

Soit en effet μ une mesure réelle; toute fonction φ admet une décomposition unique $\varphi = \varphi_1 + i\varphi_2$, où φ_1 et φ_2 sont des fonctions réelles. Elles sont alors nécessairement continues à support compact. On a alors $\mu(\varphi) = \mu(\varphi_1) + i\mu(\varphi_2)$, et d'autre part $\mu(\overline{\varphi}) = \mu(\varphi_1) - i\mu(\varphi_2)$, on a donc finalement la relation (IV,2;34), puisque $\mu(\varphi_1)$ et $\mu(\varphi_2)$ sont réels. Inversement si cette relation est vérifiée, cela prouve que,pour toute φ réelle, $\mu(\varphi)$ estscnpropre complexe conjugué , donc est réel.

Soit maintenant μ une mesure de Radon quelconque; notons $\overline{\mu}$ et appelons mesure complexe conjuguée la mesure de Radon définie par la formule

(IV,2;35) $$\overline{\mu}(\varphi) = \overline{\mu(\overline{\varphi})}, \text{ pour } \varphi \in \mathscr{C}(X) \; ; \text{ ou } \overline{\mu(\varphi)} = \overline{\mu}(\overline{\varphi}).$$

* Voici une autre interprétation. Soit μ une mesure complexe, application c-linéaire de $\mathscr{C}(X;\mathbb{C})$ dans le corps des scalaires $\mathbb{C}$. Elle définit a fortiori (note page 435) une mesure $\vec{\mu}_{\mathbb{R}}$ relativement à $\mathbb{R}$, à valeurs dans $\mathbb{C}$, c'est-à-dire une application $\mathbb{R}$-linéaire de $\mathscr{C}(X;\mathbb{R})$ dans l'espace vectoriel $\mathbb{C}$ de dimension 2 sur $\mathbb{R}$. Alors μ est réelle si et seulement si $\vec{\mu}_{\mathbb{R}}$ prend ses valeurs dans le sous-espace $\mathbb{R}$ de $\mathbb{C}$.

Tout d'abord il s'agit bien là a une mesure de Radon. En effet, c'est bien une fonction complexe sur $\mathscr{C}$ (X) ; cette fonction est linéaire, car

$$(IV,2;36)\quad \begin{cases} \bar{\mu}(\varphi_1+\varphi_2) = \overline{\mu(\overline{\varphi_1+\varphi_2})} = \overline{\mu(\bar{\varphi}_1)+\mu(\bar{\varphi}_2)} = \overline{\mu(\bar{\varphi}_1)}+\overline{\mu(\bar{\varphi}_2)} = \bar{\mu}(\varphi_1)+\bar{\mu}(\varphi_2) \\ \bar{\mu}(k\varphi) = \overline{\mu(\overline{k\varphi})} = \overline{\bar{k}\mu(\bar{\varphi})} = k\,\overline{\mu(\bar{\varphi})} = k\bar{\mu}(\varphi)^{(*)} \end{cases}$$

et elle est manifestement continue sur tout $\mathscr{C}_K(X)$, K compact.

* Regarder avec soin cette succession d'égalités; il est essentiel que, dans (IV,2;35), le 2ème membre contienne 2 fois le symbole de conjugaison. La fonction $\varphi \longrightarrow \overline{\mu(\varphi)}$ serait <u>semi-linéaire</u> et non linéaire, de même que $\varphi \longrightarrow \mu(\bar{\varphi})$.

Bien entendu, la mesure complexe conjuguée de $\overline{\mu}$ est μ, autrement dit il y a réciprocité, et l'on peut dire que μ et $\overline{\mu}$ sont deux mesures complexes conjuguées l'une de l'autre. Deux mesures μ et $\overline{\mu}$ sont complexes conjuguées, si et seulement si leurs valeurs respectives sur 2 fonctions φ, $\overline{\varphi}$, complexes conjuguees, sont toujours complexes conjuguées (deuxième formule (IV,2;35)); ou encore si, pour toute fonction φ réelle, $\mu(\varphi)$ et $\overline{\mu}(\varphi)$ sont complexes conjugués. L'application $\mu \longrightarrow \overline{\mu}$ est une bijection semi-linéaire de $\mathscr{C}'(X)$ sur lui-même. Autrement dit :

$$\text{(IV,2;37)} \quad \left\{ \begin{aligned} \overline{\mu_1 + \mu_2} &= \overline{\mu}_1 + \overline{\mu}_2 \\ \overline{k\mu} &= \overline{k}\,\overline{\mu} \end{aligned} \right.$$

Une mesure est réelle, si et seulement si elle est sa propre complexe conjuguée (formule (IV,2;34)). La relation entre μ et $\overline{\mu}$ peut encore s'exprimer de la façon suivante :

Toute mesure de Radon complexe μ peut s'écrire, d'une manière et d'une seule, sous la forme $\mu = \mu_1 + i\,\mu_2$, où μ_1 et μ_2 sont des mesures de Radon réelles. Alors $\overline{\mu} = \mu_1 - i\mu_2$.

Si en effet une telle décomposition est possible, on a nécessairement, en tenant compte de (IV,2;37) et de ce qu'une mesure réelle est sa propre complexe conjuguée, $\overline{\mu} = \mu_1 - i\mu_2$ et par conséquent, on a nécessairement la solution unique

$$\text{(IV,2;38)} \quad \mu_1 = \frac{\mu + \overline{\mu}}{2}, \quad \mu_2 = \frac{\mu - \overline{\mu}}{2i}.$$

Réciproquement les deux mesures définies par cette formule sont bien leurs propres complexes conjuguées, et sont par conséquent réelles, et on a bien $\mu = \mu_1 + i\mu_2$.

Il n'est pas Inutile de rassembler tous ces résultats en un seul tableau :

$$\text{(IV,2;39)} \quad \left\{ \begin{aligned} &\varphi = \varphi_1 + i\varphi_2, \quad \overline{\varphi} = \varphi_1 - i\varphi_2 \;; \quad \varphi_1 = \frac{\varphi + \overline{\varphi}}{2}, \quad \varphi_2 = \frac{\varphi - \overline{\varphi}}{2i} \\ &\overline{\mu}(\varphi) = \overline{\mu(\overline{\varphi})}, \quad \overline{\mu(\varphi)} = \overline{\mu}(\overline{\varphi}) \;; \\ &\mu = \mu_1 + i\mu_2, \quad \overline{\mu} = \mu_1 - i\mu_2 \;; \quad \mu_1 = \frac{\mu + \overline{\mu}}{2}, \quad \mu_2 = \frac{\mu - \overline{\mu}}{2i} \end{aligned} \right.$$

Mesures réelles positives

On dit qu'une $\mathbb{C}$- mesure de Radon μ sur X est ≥ 0, si, pour toute fonction φ de $\mathscr{C}(X)$ qui est ≥ 0, on a $\mu(\varphi) \geq 0$; on écrit $\mu \geq 0$. Alors $\varphi_1 \leq \varphi_2$ entraîne $\mu(\varphi_1) \leq \mu(\varphi_2)$.

Il résulte évidemment de cette définition qu'une mesure positive est nécessairement réelle.

Une mesure de Radon μ est dite négative, $\mu \leq 0$, si, pour toute $\varphi \geq 0$, on a $\mu(\varphi) \leq 0$.

Pour qu'une mesure soit négative, il faut et il suffit qu'elle soit l'opposée d'une mesure positive.

si μ et ν sont 2 mesures réelles sur X, on dit que $\mu \geq \nu$, si $\mu - \nu \geq 0$. Cela revient à dire que, pour toute $\varphi \geq 0$ de $\mathscr{C}(X)$, on a $\mu(\varphi) \geq \nu(\varphi)$.

si $\mu \geq \nu \geq 0$, et $\varphi \geq \Psi \geq 0$, on a $\mu(\varphi) \geq \nu(\Psi) \geq 0$

Remarque. Soit μ une mesure ≥ 0. Pour toute φ réelle, on a $-|\varphi| \leq \varphi \leq |\varphi|$, donc $-\mu(|\varphi|) \leq \mu(\varphi) \leq \mu(|\varphi|)$, donc

(IV, 2; 39 bis) $$|\mu(\varphi)| \leq \mu(|\varphi|).$$

Cette relation subsiste pour φ complexe, mais elle est considérablement plus difficile à montrer, Soit K le support de φ, $\mathcal{V}$ un voisinage compact de K. Soit Ψ une fonction associée à φ suivant le corollaire 8 du théorème 11, telle que $\Psi = \sum_{i=1}^{N} g_i \Psi_i$, $g_i \in \mathbb{C}$, $\Psi_i \in \mathscr{C}(X)$ mais ≥ 0, à support dans $\mathcal{V}$:

$$\left| \varphi(x) - \sum_{i=1}^{N} g_i \Psi_i(x) \right| \leq \frac{\varepsilon}{2\|\mu\|_{\mathcal{V}}}$$

$$\left| |\varphi(x)| - \sum_{i=1}^{N} |g_i| \Psi_i(x) \right| \leq \frac{\varepsilon}{2\|\mu\|_{\mathcal{V}}}$$

Alors la 1ère et la 2ème inégalité donnent successivement :

$$\left|\mu(\varphi)\right| \leqslant \sum_{i=1}^{N} |g_i|\,\mu(\Psi_i) + \frac{\varepsilon}{2} \leqslant \mu(\varphi) + \varepsilon\ ;$$

ε étant arbitraire, celà prouve (IV,2;39 bis). Ouf !

On donnera plus loin une autre méthode, utilisant la théorie du prolongement de Lebesgue. En considérant μ comme une $\mathbb{R}$ - mesure (application $\mathbb{R}$ - linéaire de $\mathscr{C}(X;\mathbb{R})$ dans $\mathbb{R}$), une fonction $\vec{\varphi} \in \mathscr{C}(X;\mathbb{C})$ est à valeurs vectorielles dans $\mathbb{C}$, espace de dimension 2 sur $\mathbb{R}$; et alors (IV,2;39bis) est la formule (IV,3;27).

Il résulte de (IV,2;39bis) que pour le calcul de $\|\mu\|_K$, on peut se borner a considérer des fonctions φ réelles $\geqslant 0$, et que l'on a

(IV,2;39 septimo)
$$\|\mu\|_K = \sup_{\substack{\varphi \in \mathscr{C}_K(X) \\ 0 \leqslant \varphi \leqslant 1}} \mu(\varphi)\ , \quad \text{et}$$

$$\|\mu\| = \sup_{\substack{\varphi \in \mathscr{C}(X) \\ 0 \leqslant \varphi \leqslant 1}} \mu(\varphi)$$

En d'autres termes, en reprenant les notations de la note page 435, μ et $\mu_{\mathbb{R}}$ ont même norme si μ est une $\mathbb{C}$ - mesure réelle ≥ 0. Nous verrons (remarque page 456, 7) que ceci reste vrai si μ est une $\mathbb{C}$ - mesure réelle quelconque; mais naturellement pas si μ est un6 $\mathbb{C}$ - mesure complexe quelconque, comme le montre l'exemple de la note page 435. On peut encore dire :

dans la formule $\|\mu\|_K = \sup\limits_{|\varphi| \leq 1} |\mu(\varphi)|$, on doit prendre toutes les φ complexes; on peut se borner aux φ réelles si μ est réelle, aux $\varphi \geq 0$ si $\mu \geq 0$.

La relation binaire $\mu \leq \nu$ définit une relation d'ordre sur l'ensemble des mesures réelles sur X. La seule chose non triviale a vérifier pour le justifier est que $\mu \leq 3$ et $\mu \geq 3$ entrafne $\mu = \nu$. Or ces inégalités entraîner t $\mu(\varphi) = \nu(\varphi)$ pour $\varphi \geq 0$; mais toute fonction φ réelle de $\mathscr{C}(X)$ est différence de 2 fonctions ≥ 0 de $\mathscr{C}(X)$ (par exemple $\varphi = |\varphi| - (|\varphi| - \varphi)$), donc c'est aussi vrai pour φ réelle; donc Pour $\varphi = \varphi_1 + i\varphi_2$ complexe et par suite $\mu = \nu$. La relation d'ordre est "compatible avec structure $\mathbb{R}$- vectorielle de l'espace des mesures réelles', en ce sens que

$$\text{(IV,2;39 Octavo)} \left\{ \begin{array}{l} \mu_1 \leq \nu_1 \quad , \quad \mu_2 \leq \nu_2 \text{, implique } \mu_1 + \mu_2 \leq \nu_1 + \nu_2 \text{,} \\ \text{et } \mu \leq \nu \text{ implique } k\mu \leq k\nu \quad \text{pour } k \geq 0 \text{.} \end{array} \right.$$

Un espace vectoriel sur $\mathbb{R}$ muni d'une structure d'ordre ayant ces propriétés, est appelé <u>espace vectoriel ordonné</u>. La première relation exprime encore que $\mu \geq \nu$ est équivalent à $\mu - \nu \geq 0$, et que $\mu \geq 0$, $\nu \geq 0$, implique $\mu + \nu \geq 0$; la deuxième que $\mu \geq 0$ implique $k\mu \geq 0$ pour $k \geq 0$. Remarquons que dans ce cas, $\mu \leq \nu$ implique $-\nu \leq -\mu$, donc $k\mu \geq k\nu$ pour $k \leq 0$.

Exemple Considérons la mesure définie par (IV,2;15) sur $\mathbb{R}$. Sa complexe conjuguée est :

(IV,2;40) $$\sum_\nu \bar{c}_\nu \delta_{(a_\nu)} + \bar{f}\, dx \;.$$

Alors la mesure μ est réelle, si et seulement si ces deux mesures sont égales, c'est-à-dire si et seulement si la mesure μ définie par $\sum_\nu (c_\nu - \bar{c}_\nu)\delta_{(a_\nu)} + (f - \bar{f})\,dx$ est nulle. Or elle est sûrement nulle si tous les c_ν sont réels et si la fonction f est réelle. Réciproquement, si cette mesure est nulle, son support est vide, et par conséquent, d'après le théorème 15 bis, on peut en déduire, si la fonction f est continue, que toutes les quantités $c_\nu - \bar{c}_\nu$ sont nulles et que la fonction $f - \bar{f}$ est nulle, autrement dit que les c_ν sont réels et que la fonction* est réelle. Ce résultat naturellement ne subsiste pas nécessairement si la fonction f n'est pas continue; il est bien toujours vrai que les c_ν sont réels, mais la fonction+ n'est pas nécessairement réelle; ce qu'on peut dire, et c'est l'essentiel, c'est que f et $\mathcal{R}f$, quoique non nécessairement égales, définissent la même mesure $f\,dx = \mathcal{R}f\,dx$ (puisque $(f - \mathcal{R}f)\,dx = \left(\frac{f-\bar{f}}{2}\right)dx = 0$), et qu'on peut donc, si f n'est pas réelle, la remplacer par sa partie réelle $\mathcal{R}f$ sans changer la mesure. La mesure μ définie par (IV,2; 15) est sûrement positive si tous les c_ν sont $\geqslant 0$ et si la fonction f est $\geqslant 0$. On démontre que la réciproque est vraie si la fonction f est continue. *

Si en effet, tous les points a_ν étant supposés distincts, un seul des coefficients c_ν est < 0 , ou si en un point a la fonction f , supposée continue, prend une valeur $f(a) < 0$, alors on démontre sans difficultés que la fonction $\varphi \geqslant 0$, construite aux exemples 1, 2, 3, des pages 446 à 449, vérifierait $\varphi < 0$, ce qui serait contradictoire.

Ensembles ordonnés réticulés

Un ensemble ordonné E est dit réticulé si toute partie formée de 2 éléments admet une borne supérieure (c'est a-dire, rappelons le, un plus petit majorant) et une borne inférieure. Il en est alors de même pour toute partie finie non vide de E . Par exemple la droite réelle $\mathbb{R}$ est réticulée; l'ensemble ordonné $\mathbb{R}^X$ des fonctions

* Le théorème 32 - 53 - 54 répondra complètement aux mêmes questions si f n'est pas continue.

sur X à valeurs réelles (voir page 20) est réticulé, de même que l'ensemble $(\mathbb{R}^X)_c$ des fonctions réelles continues sur X topologique; par contre, par exemple, l'ensemble des polynomes à coefficients réels sur $\mathbb{R}$, pour la relation d'ordre habituelle, n'est évidemment pas réticulé. Si E est un espace vectoriel ordonné (ce qui est le cas des exemples précédents) il suffit, pour qu'il soit réticulé que, pour tout $e \in E$, $e^+ = \mathrm{Sup}(e,0)$ existe. Si alors en effet e et f sont 2 éléments quelconques de E, l'élément $e+(f-e)^+$ majore $e+0=e$ et $e+(f-e)=f$; et si g majore e et f, $g-e$ majore $e-e=0$ et $f-e$ donc majore $(f-e)^+$, donc g majore $e+(f-e)^+$, qui est ainsi la borne supérieure de e et f. D'autre part e et f ont une borne inférieure, à savoir $-\mathrm{Sup}(-e,-f)$. Enfin on a toujours $\mathrm{Sup}(e,f)+\mathrm{Inf}(e,f)=e+f$. En effet il est évident que $\mathrm{Sup}(e-h,f-h)=\mathrm{Sup}(e,f)-h$; en prenant $h=e+f$, on a la formule voulue. On pose aussi $e^-=\mathrm{Sup}(-e,0)=(-e)^+$; alors $e=e+0=\mathrm{Sup}(e,0)+\mathrm{Inf}(e,0)=e^+-e^-$. On appelle e^+ et e^- la partie positive et la partie négative de e ; malgré son nom, e^- est ≥ 0 et non ≤ 0, et c'est plutôt $-e^-$ qui mériterait le nom de partie négative * . On voit alors que tout élément e de E est différence de 2 éléments ≥ 0, e^+ et e^-.

Il l'est en général d'une infinité de manières car on peut poser $e=e_1-e_2$ avec $e_1=e^++a$, $e_2=e^-+a$, $a\geq 0$, mais e^+ et e^- sont les plus petits éléments ≥ 0 dont e soit différence. Si en effet $e=e_1-e_2$, e_1 et $e_2\geq 0$, e_1 majore e et 0 donc e^+, e_2 majore $-e$ et 0 donc e^-; et on retrouve la forme précédemment signalée, $e_1=e^++a$, $e_2=e^-+a$, $a\geq 0$. Les éléments e^+ et e^- sont ≥ 0 "étrangers", en ce sens que le seul élément ≥ 0 qui les minore tous les deux est 0 ; et même tout élément g qui minore à la fois e^+ et e^- est ≤ 0. En effet, celà revient à dire que $\inf(e^+,e^-)=0$; or c'est $e^++\inf(0,e^--e^+)=e^++\inf(0,-e)=e^+-\sup(0,e)=e^+-e^+=0$. Enfin on appelle valeur absolue de e la borne supérieure de e et $-e$: $|e|=\sup(e,-e)$;

Alors $|e|+e=\sup(2e,0)=2e^+$, donc $|e|=2e^+-e=2e^+-(e^+-e^-)=e^++e^-$; comme d'ailleurs $\inf(e^+,e^-)=0$, e^++e^- vaut aussi $\sup(e^+,e^-)$.

* De même, a et b sont la partie réelle et la partie imaginaire du nombre complexe $a+ib$; c'est plutôt ib qui mériterait le nom de partie imaginaire.

On dit qu'un ensemble ordonné est complètement réticulé si toute partie non vide majorée a une borne supérieure, et toute partie non vide minorée une borne inférieure (la première condition entraîne d'ailleurs la deuxième, car si une partie A non vide est minorée, l'ensemble de ses minorants est majoré non vide, donc a une borne supérieure, qui est la borne inférieure de A). La droite réelle $\mathbb{R}$ est complètement réticulée, ainsi que l'ensemble $\mathbb{R}^X$ des fonctions sur X à valeurs réelles. Mais, si X est topologique, l'ensemble $(\mathbb{R}^X)_c$ des fonctions réelles continues sur X qui est réticulé, n'est pas complètement réticulé. Par exemple :

Si X est la droite $\mathbb{R}$, et si A est le sous-ensemble de $(\mathbb{R}^{\mathbb{R}})_c$ formé des fonctions continues majorées par la fonction caractéristique de $[-1, +1]$ la borne supérieure de A dans l'ensemble ordonné $\mathbb{R}^{\mathbb{R}}$ de toutes les fonctions réelles sur $\mathbb{R}$ est cette fonction caractéristique, qui est discontinue, et on voit sans peine que A n'a pas de borne supérieure dans l'ensemble $(\mathbb{R}^{\mathbb{R}})_c$ des fonctions continues.

Théorème 18 - L'espace vectoriel ordonné des mesures de Radon réelles sur l'espace localement compact X est complètement réticulé. Si μ est une mesure réelle, $\mu^+ = \mathrm{Sup}(\mu, 0)$, $\mu^- = \mathrm{Sup}(-\mu, 0)$, $|\mu| = \mathrm{Sup}(\mu, -\mu) = \mathrm{Sup}(\mu^+, \mu^-) = \mu^+ + \mu^-$, on a les formules suivantes :

(IV,2;41) $$\mu^+(\varphi) = \sup_{\substack{0 \leq \psi \leq \varphi \\ \psi \in \mathcal{C}(X)}} \mu(\psi), \quad \text{pour } \varphi \geq 0$$

(IV,2;41 bis) $$\|\mu\| = \|\mu^+\| + \|\mu^-\|$$

les normes étant prises sur X ou sur un compact K de X ;

(IV,2;41 ter) $$\begin{cases} |\mu|(\varphi) = \sup\limits_{\substack{|\psi| \leq \varphi \\ \psi \in \mathcal{C}(X)}} \mu(\psi), \quad \text{pour } \varphi \geq 0 \quad \text{et} \\ \| |\mu| \| = \|\mu\| = \|\mu^+\| + \|\mu^-\|, \end{cases}$$

sur X ou sur un compact K de X ,

On a donc, pour $\varphi \geq 0$, $\mu(\varphi) \leq \mu^+(\varphi)$, $-\mu(\varphi) \leq \mu^-(\varphi)$, et pour φ quelconque $|\mu(\varphi)| \leq |\mu|(|\varphi|)$.

Démonstration - Désignons par $\mathcal{C}^+(X)$ l'ensemble des $\varphi \geq 0$ de $\mathcal{C}(X)$. Définissons alors la fonction μ^+ sur $\mathcal{C}^+(X)$, à valeur dans $\mathbb{R}_+$, par la formule (IV,2;41).

Elle donne bien un résultat fini, parce que l'on a l'inégalité $|\mu(\Psi)| \leq \|\mu\|_K \|\Psi\| \leq \|\mu\|_K \|\varphi\|$, si K est le support de φ. On obtient bien un résultat ≥ 0, parce que, parmi les fonctions Ψ, figure la fonction 0. Montrons que cette fonction μ satisfait sur $\mathcal{C}^+(X)$ aux propriétés de "linéarité" suivantes :

(IV,2;41 quarto)
$$\mu^+(\varphi_1 + \varphi_2) = \mu^+(\varphi_1) + \mu^+(\varphi_2)$$
$$\mu^+(\lambda\varphi) = \lambda\,\mu^+(\varphi), \text{ pour } \lambda \geq 0$$

La deuxième propriété est évidente, il suffit de montrer la première. Quelles que soient les fonctions Ψ_1 et Ψ_2, $0 \leq \Psi_1 \leq \varphi_1$ et $0 \leq \Psi_2 \leq \varphi_2$, on a $0 \leq \Psi_1 + \Psi_2 \leq \varphi_1 + \varphi_2$, d'où $\mu(\Psi_1) + \mu(\Psi_2) = \mu(\Psi_1 + \Psi_2) \leq \mu^+(\varphi_1 + \varphi_2)$; on en déduit aussitôt, en prenant la borne supérieure par rapport à Ψ_1 et Ψ_2 :

$$\mu^+(\varphi_1) + \mu^+(\varphi_2) \leq \mu^+(\varphi_1 + \varphi_2)$$

Par ailleurs si $\Psi \in \mathcal{C}^+(X)$, $0 \leq \Psi \leq \varphi_1$ t φ_2, on peut trouver deux fonctions Ψ_1 et Ψ_2 vérifiant

$$0 \leq \Psi_1 \leq \varphi_1,\ 0 \leq \varphi_2 \leq \varphi_2, \text{ et } \Psi_1 + \Psi_2 = \Psi.$$

Il suffit en effet de prendre

$$\Psi_1 = \operatorname{Inf}(\Psi, \varphi_1), \quad \Psi_2 = \Psi - \Psi_1 ;$$

d'après la définition même, on a bien $0 \leq \Psi_1 \leq \varphi_1$, $0 \leq \Psi_2$ et $\Psi_1 + \Psi_2 = \Psi$; soit d'autre part $x \in X$; si $\Psi(x) \leq \varphi_1(x)$, on a $\Psi_1(x) = Y(x)$, et par suite $\Psi_2(x) = 0 \leq \varphi_2(x)$; si au contraire on a $\Psi(x) > \varphi_1(x)$, on a $\Psi_1(x) = \varphi_1(x)$ d'où $\Psi_2(x) = \Psi(x) - \varphi_1(x) \leq \varphi_1(x) + \varphi_2(x) - \varphi_1(x) = \varphi_2(x)$; on a donc bien dans tous les cas $0 \leq \Psi_2 \leq \varphi_2$.

On aura alors l'inégalité

$$\mu(\Psi) = \mu(\Psi_1) + \mu(\Psi_2) \leq \mu^+(\varphi_1) + \mu^+(\varphi_2) ;$$

d'ou en prenant la borne supérieure par rapport à Ψ, l'inégalité $\mu^+(\varphi_1+\varphi_2)\leq \mu^+(\varphi_1)+\mu^+(\varphi_2)$,

qui achève de démontrer (IV,2;41 quarto).

Soit maintenant φ une fonction réelle de signe quelconque de $\mathcal{C}(X)$. Elle peut d'une infinité de manières s'exprimer comme différence $\varphi=\varphi_1-\varphi_2$ de deux fonctions ≥ 0 de $\mathcal{C}(X)$; il suffit par exemple de prendre $\varphi=\varphi^+-\varphi^-$, car φ^+ et φ^- sont tous deux continues à support compact.

Définissons alors $\mu^+(\varphi)$ par la formule

$$\mu^+(\varphi)=\mu^+(\varphi_1)-\mu^+(\varphi_2)$$

Ceci n'est justifié que si nous montrons l'indépendance de $\mu^+(\varphi)$ par rapport à la décomposition $\varphi=\varphi_1-\varphi_2$ *. Or si $\varphi=\varphi_3-\varphi_4$ est une autre décomposition, on a $\varphi_1+\varphi_4=\varphi_2+\varphi_3$, d'où, d'après (IV,2;41 quarto), $\mu^+(\varphi_1)+\mu^+(\varphi_4)=\mu^+(\varphi_2)+\mu^+(\varphi_3)$, et par suite $\mu^+(\varphi_1)-\mu^+(\varphi_2)=\mu^+(\varphi_3)-\mu^+(\varphi_4)$, ce qui prouve bien cette indépendance.

μ^+ est maintenant définie comme une fonction sur $\mathcal{C}(X;\mathbb{R})$ à valeurs dans $\mathbb{R}$. Il est en outre évident que cette fonction μ^+ est linéaire sur cette espace vectoriel. La formule $\mu^+(\lambda\varphi)=\lambda\mu^+(\varphi)$ est évidente; il nous faut seulement montrer la formule d'additivité. Or, si φ et Ψ sont deux fonctions réelles de $\mathcal{C}(X)$, on écrira $\varphi=\varphi_1-\varphi_2$ et $\Psi=\Psi_1-\Psi_2$, avec $\varphi_1,\varphi_2,\Psi_1,\Psi_2\geq 0$. On aura donc la décomposition $\varphi+\Psi=(\varphi_1+\Psi_1)-(\varphi_2+\Psi_2)$, d'où la formule suivante (en utilisant encore (IV,2;41 quarto)

$$\mu^+(\varphi+\Psi)=\mu^+(\varphi_1+\Psi_1)-\mu^+(\varphi_2+\Psi_2)=\mu^+(\varphi_1)+\mu^+(\Psi_1)-\mu^+(\varphi_2)-\mu^+(\Psi_2)=\mu^+(\varphi)+\mu^+(\Psi),$$

qui achève de montrer la linéarité de μ^+.

Vérifions maintenant la propriété de continuité qui permettra d'affirmer que μ^+ est une $\mathbb{R}$ – mesure scalaire.

* En utilisant φ^+ et φ^- il serait inutile de démontrer ce point. Mais alors on ne pourrait plus démontrer ensuite la linéarité de μ^+, car $(\varphi+\Psi)^+\neq\varphi^++\Psi^+$

Soit K un compact et soit α une fonction $\geqslant 0$ de $\mathscr{C}(X)$ égale à 1 sur un voisinage de K (corollaire du théorème'11). Posons $\mu^+(\alpha) = k$. On a alors les inégalités, pour φ réelle' :

(IV,2;41 quinto) $$-\|\varphi\|\alpha \leqslant \varphi \leqslant \|\varphi\|\alpha \quad \text{donc}$$

$$-\|\varphi\|\mu^+(\alpha) \leqslant \mu^+(\varphi) \leqslant \|\varphi\|\mu^+(\alpha) \quad \text{ou}$$

$$|\mu^+(\varphi)| \leqslant k\|\varphi\|,$$

qui prouvent bien que μ^+ est continue sur $\mathscr{C}(X;\mathbb{R})$, et que l'on a l'inégalité $\|\mu^+\|_K \leqslant k = \mu^+(\alpha)$ *.

Ceci achève de prouver que μ^+ est une $\mathbb{R}$ – mesure, évidemment $\geqslant 0$. D'après sa définition, $\mu^+ \geqslant \mu$. En outre toute mesure $\nu \geqslant 0$ majorant μ , c'est-à-dire majorant μ et 0 , vérifie pour toute $\varphi \geqslant 0$, et $0 \leqslant \Psi \leqslant \varphi$: $\nu(\varphi) \geqslant \nu(\Psi) \geqslant \mu(\Psi)$, donc, en prenant la borne supérieure par rapport à Ψ , $\nu(\varphi) \geqslant \mu^+(\varphi)$, donc $\nu \geqslant \mu^+$. Cela prouve que μ^+ est la borne supérieure de μ et 0 dans l'espace vectoriel ordonné des mesures réelles sur X . Donc cet espace est réticulé.

Les normes étant additives, il est bien évident que l'on a l'inégalité $\|\mu\|_K \leqslant \|\mu^+\|_K + \|\mu^-\|_K$, K compact quelconque. Montrons que l'on a en réalité l'égalité des deux membres. Quel que soit $\varepsilon > 0$, il existe une fonction $\varphi_1 \in \mathscr{C}_K(X)$ telle que l'on ait l'inégalité

$$\|\mu^+\| \geqslant \mu^+(\varphi_1) \geqslant \|\mu^+\|_K - \frac{\varepsilon}{4}, \quad 0 \leqslant \varphi_1 \leqslant 1$$

(On peut prendre $\varphi_1 \geqslant 0$ parce que $\mu^+ \geqslant 0$; voir formule (IV,2;39 septimo))

* Nous venons de démontrer en passant une propriété générale : toute forme linéaire sur $\mathscr{C}(X) \geqslant 0$ sur $\mathscr{C}^+(X)$, est une mesure, c'est-à-dire est continue sur tout $\mathscr{C}_K(X)$, K compact $\subset X$.

Il existe alors $\Psi_1 \in \mathscr{C}_K(X)$, $0 \leqslant \Psi_1 \leqslant \varphi_1 \leqslant 1$, telle que

$$\mu(\Psi_1) \geqslant \mu^+(\varphi_1) - \frac{\varepsilon}{4} \geqslant \|\mu^+\|_K - \frac{\varepsilon}{2} .$$

De même il existe $\Psi_2 \in \mathscr{C}_K(X)$, 0 a $\Psi_2 \leqslant 1$, telle que

$$-\mu(\Psi_2) \geqslant \|\mu^-\|_K - \frac{\varepsilon}{2}$$

Alors $\theta = \Psi_1 - \Psi_2$ est dans $\mathscr{C}_K(X)$, on a $|\theta| \leqslant 1$, et $\mu(\theta) \geqslant \|\mu^+\|_K + \|\mu^-\|_K - \varepsilon$; ε étant arbitraire, cela prouve bien notre affirmation, Le résultat est le même pour les normes sur X .

Considérons maintenant la mesure $|\mu| = \mu^+ + \mu^-$

Montrons que $|\mu|$ peut ètre uéfinie par la formule (IV,2;41 ter). Tout d'abord le 1er membre est certainement au moins égal au deuxième, car, si $\Psi = \Psi^+ - \Psi^-$, $\mu(\Psi) = \mu(\Psi^+) - \mu(\Psi^-) \leqslant \mu^+(\Psi^+) + \mu^-(\Psi^-) \leqslant \mu^+(\varphi) + \mu^-(\varphi) = |\mu|(\varphi)$. Mais il existe des fonctions $\Psi_1, \Psi_2, 0 \leqslant \Psi_1 \leqslant \varphi, 0 \leqslant \psi_2 \leqslant \varphi$, telles que l'on ait

$$\mu(\Psi_1) \geqslant \mu^+(\varphi) - \frac{\varepsilon}{2} \quad , \quad -\mu(\Psi_2) \geqslant \mu^-(\varphi) - \frac{\varepsilon}{2} .$$

La fonction $\theta = \Psi_1 - \Psi_2$ vérifie alors $|\theta| \leqslant \varphi$, et $\mu(\theta) \geqslant \mu^+(\varphi) + \mu^-(\varphi) - \varepsilon = |\mu|(\varphi) - \varepsilon$ $\varepsilon > 0$ étant arbitraire, cela donne bien (IV,2;41 ter).

(IV,2;41 ter) montre que $\||\mu|\|_K$

$$= \sup_{\substack{0 \leqslant \varphi \leqslant 1 \\ \varphi \in \mathscr{C}_K(X)}} |\mu|(\varphi) = \sup_{\substack{|\Psi| \leqslant 1 \\ \Psi \in \mathscr{C}_K(X)}} |\mu(\Psi)| = \|\mu\|_K ,$$

donc on a bien $\||\mu|\|_K = \|\mu^+\|_K + \|\mu^-\|_K$.

Nous avons défini $\mu^+, \mu^-, |\mu|$ comme $\mathbb{R}$ - mesures scalaires (applications $\mathbb{R}$ - linéaires de $\mathscr{C}(X;\mathbb{R})$ dans $\mathbb{R}$). En posant $\mu^+(\varphi_1 + i\varphi_2) = \mu^+(\varphi_1) + i\,\mu^+(\varphi_2)$, etc... on les définit comme $\mathbb{C}$ - mesures scalaires, mais réelles $\geqslant 0$ au sens des pages 453-455.

On a vu (formule (IV,2;39 septimo)) que, pour une $\mathbb{C}$ - mesure $\geqslant 0$, la norme se calcule à partir de ses valeurs sur des fonctions $\varphi \geq 0$; on en déduit que les formules $\|\mu\| = \||\mu|\| = \|\mu^+\| + \|\mu^-\|$ restent valables pour ces $\mathbb{C}$ - mesures réelles. En outre la formule (IV,2;41 ter) reste valable pour des Ψ complexes (et toujours $\varphi \geqslant 0$), ainsi que sa conséquence $|\mu(\varphi)| \leqslant |\mu|(|\varphi|)$ pour φ complexe (car

$$|\mu(\varphi)| \leqslant |\mu^+(\varphi)| + |\mu^-(\varphi)| \leqslant \mu^+(|\varphi|) + \mu^-(|\varphi|) = |\mu|(|\varphi|)).$$

Il reste enfin à montrer que l'espace des mesures réelles est complètement réticulé (c'est assez remarquable; rappelons que l'espace $\mathscr{C}(X;\mathbb{R})$, lui, n'est pas complètement réticulé). Soit donc $(\mu_i)_{i\in I}$ une famille quelconque de mesures réelles, majorée; il existe donc 3 telle que $\mu_i \leqslant 3$ pour tout i . Pour tout sous-ensemble fini non vide J de 1 , appelons μ_J la borne supérieur6 des μ_i , $i \in J$ on a $\mu_K \geq \mu_J$ si $K \supset J$. Il nous faut montrer que 1'ensemble des μ_J admet une borne supérieure. Soit $\varphi \geqslant 0$. Posons

(IV,2;41 sexto) $$\mu_0(\varphi) = \operatorname*{Sup}_J \mu_J(\varphi)$$

(Attention, ce n'est pas du tout la même chose que $\operatorname*{Sup}_{i\in I} \mu_i(\varphi)$, qui donnerait un résultat inexact, car, pour 2 mesures μ_1 et μ_2 , on n'a pas $(\operatorname{Sup}(\mu_1,\mu_2))(\varphi) = \operatorname{Sup}(\mu_1(\varphi),\mu_2(\varphi))$; par exemple, pour μ et 0 , on n'a pas $\mu^+(\varphi) = \operatorname{Sup}(\mu(\varphi),0) = (\mu(\varphi))^+$).
Cette quantité est finie, car elle majorée par $\nu(\varphi)$.
On définit'ainsi une fonction sur $\mathscr{C}^+(X)$. Montrons qu'elle est "linéaire", au sens vu à la formule (IV,2;41 ouarto). Pour la multiplication par les scalaires, c'est évident ; -voyons l'addition. On a

$$\mu_J(\varphi_1+\varphi_2) = \mu_J(\varphi_1) + \mu_J(\varphi_2) \leq \mu_0(\varphi_1) + \mu_0(\varphi_2) ;$$

en prenant la borne supérieure par rapport à J , on a $\mu_0(\varphi_1+\varphi_2) \leq \mu_0(\varphi_1) + \mu_0(\varphi_2)$. Mais il existe $J \subset I$, fini, tel que $\mu_J(\varphi_1) \geqslant \mu_0(\varphi_1) - \frac{\varepsilon}{2}$, et K fini tel que $\mu_K(\varphi_2) \geqslant \mu_0(\varphi_2) - \frac{\varepsilon}{2}$. Alors, si $L = J \cup K$, on aura

$$\mu_0(\varphi_1+\varphi_2) \geqslant \mu_L(\varphi_1+\varphi_2) = \mu_L(\varphi_1) + \mu_L(\varphi_2) \geqslant \mu_J(\varphi_1) + \mu_K(\varphi_2) \geqslant \mu_0(\varphi_1) + \mu_0(\varphi_2) - \varepsilon,$$

d'où le résultat cherché. On définira alors μ_0 sur $\mathscr{C}(X;\mathbb{R})$ exactement par le même procédé que pour la definition de μ^+ à partir de μ au début de la démonstration du théorème. Alors μ_0 devient une forme $\mathbb{R}$-linéaire sur $\mathscr{C}(X;\mathbb{R})$. Il faut montrer sa continuité sur $\mathscr{C}_K(X;\mathbb{R})$. Utilisant le même procédé que ~~pour~~ μ^+ antérieurement, on aura : en choisissant un indice i quelconque, et pour $\iota \in J$:

$$\mu_J \leqslant \nu \quad \text{donc} \quad \mu_J^+ \leqslant \nu^+ ; \; -\mu_J \leqslant -\mu_i \text{ donc } \mu_J^- \leqslant \mu_i^- ;$$

pour toute φ réelle, on a donc

$$\mu_J(\varphi) = \mu_J(\varphi^+) - \mu_J(\varphi^-) \leqslant \nu^+(\varphi^+) + \mu_i^-(\varphi^-) \leqslant \nu^+(|\varphi|) + \mu_i^-(|\varphi|).$$

De même

$$-\mu_J(\varphi) = -\mu_J(\varphi^+) + \mu_J(\varphi^-) \leqslant \mu_i^-(\varphi^+) + \nu^+(\varphi^-) \leqslant \nu^+(|\varphi|) + \mu_i^-(|\varphi|).$$

Finalement $|\mu_J(\varphi)| \leqslant (\nu^+ + \mu_i^-)(|\varphi|)$. Donc, en passant à la borne supérieure pour J, $|\mu_0(\varphi)| \leqslant (\nu^+ + \mu_i^-)(|\varphi|)$. On en déduit bien la continuité de μ_0, et en outre

$$\|\mu_0\|_K \leqslant \|\nu^+\|_K + \|\mu_i^-\|_K .$$

Maintenant μ_0 est une $\mathbb{R}$-mesure (on peut la prolonger de façon évidente à $\mathscr{C}(X;\mathbb{C})$ en une $\mathbb{C}$-mesure réelle; elle majore évidemment toutes les μ_J, et elle est la plus petite mesure a posséder cette propriété, d'après la définition même (IV,2;41 sexto), donc c'est la borne supérieure des μ_J, et l'ensemble des mesures réelles sur X est bien complètement réticulé.

<u>Remarque</u> - Nous avons vu (formule (IV,2;39 septimo)) que la norme d'une $\mathbb{C}$-mesure $\mu \geqslant 0$ pouvait se calculer par des $\mu(\varphi)$, φ réelles $\geqslant 0$. Maintenant nous pouvons dire que, pour une $\mathbb{C}$-mesure μ réelle, la norme peut se calculer par des $\mu(\varphi)$, φ réelles (et qu'elle est donc encore égale à la norme de $\mu_{\mathbb{R}}$: voir note page 4351; c'est ce que nous avons vu, au cours de la démonstration du théorème. Naturellement la formule (IV,2;39 septimo) n'est pas valable, mals on a toujours, par (IV,2;41 ter), $|\mu(\varphi)| \leqslant |\mu|(|\varphi|)$, pour φ réelle ou complexe.

On dit qu'une mesure vectorielle $\vec{\mu}$ sur x est absolument mdjorée par une mesure λ réelle $\geqslant 0$, si, pour Soute $\varphi \in \mathscr{C}(X)$ réelle $\geqslant 0$, on a

(IV,2;42) $$\|\vec{\mu}(\varphi)\| \leqslant \lambda(\varphi)$$

Si alors φ est réelle de signe quelconque on a aussi

(IV,2;42 bis) $$\|\vec{\mu}(\varphi)\| = \|\vec{\mu}(\varphi^+) - \vec{\mu}(\varphi^-)\| \leqslant \lambda(\varphi^+) + \lambda(\varphi^-) = \lambda(|\varphi|) \cdot$$

On peut démontrer que cette propriété subsiste pour φ complexe, en utilisant une partition de l'unité (voir la démonstration de (IV,2;39 bis)). Si μ est une $\mathbb{C}$ - mesure réelle, la formule (IV,2;41 ter) montre que la plus petite mesure $\geqslant 0$ majorant absolument μ est $|\mu|$.

Soit $\vec{E}$ un espace vectoriel de dimension finie, et soit $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, une base de Ë comme espace vectoriel sur le corps $\mathbb{R}$. Pour toute φ réelle, on peut écrire

$$\vec{\mu}(\varphi) = \sum_{i=1}^{n} \mu_i(\varphi)\, \vec{e}_i \,,$$

et les $\mu_i : \varphi \longrightarrow \mu_i(\varphi)$ sont des formes $\mathbb{R}$ -linéaires sur $\mathscr{C}(X;\mathbb{R})$, continues sur les $\mathscr{C}_K(X;\mathbb{R})$, K compact de X , c'est-à-dire des $\mathbb{R}$ mesures scalaires sur X . Pour toute $\varphi \geqslant 0$, on a

$$\|\vec{\mu}(\varphi)\| \leqslant \underset{i=1,2,\ldots,n}{\mathrm{Max}} \|\vec{e}_i\| \sum_{i=1}^{n} |\mu_i|(\varphi)$$

donc $\vec{\mu}$ est absolument majorée par la mesure $\geqslant 0$: $\underset{i=1,2,\ldots n}{\mathrm{Max}} \|\vec{e}_i\| \sum_{i=1}^{n} |\mu_i|$ Elle admet alors une plus petite majorante absolue; plus généralement toute mesure à valeurs dans un Banach $\vec{E}$, admettant une majorante absolue, admet une plus petite majorante absolue, à Savoir la borne inférieure de l'ensemble minoré (par 0) non vide des majorantes absolues (l'ensemble des mesures réelles est complètement réticulé).

Mais une mesure à valeurs dans un Banach $\vec{E}$ de dimension infinie n'admet en général pas de majorante absolue. Considérons par exemple le cas de X compact, et soit $\vec{E} = \mathscr{C}(X)$, avec sa norme habituelle. L'application identique $\varphi \longrightarrow \varphi$ définit une mesure $\vec{\mu}$ sur X , à valeurs dans $\vec{E} = \mathscr{C}(X)$. Cette mesure n'a pas de majorante absolue (sauf si X est un ensemble fini). Soit en effet λ une majorante absolue. On devrait avoir, Pour toute $\varphi \geq 0$, $\|\varphi\| = \|\mu(Q)\| \leq \lambda(\varphi)$. Mais, si X est infini, on peut, pour tout n, trouver n ouverts de X, 2 à 2 disjoints (corollaire 6 du théorème 11, appliqué à n ensembles fermés réduits a des points distincts). Pour chacun d'eux on peut trouver une α_i continue, $0 \leq \alpha_i \leq 1$, de support dans cet ouvert, et prenant la valeur 1 en au moins un point (lemme 1 du théorème 11). On aurait alors $1 = \|\alpha_i\| \leq \lambda(\alpha_i)$ d'où $n \leq \sum_{i=1}^{n} \lambda(\alpha_i) = \lambda\left(\sum_{i=1}^{n} \alpha_i\right) \leq \lambda(1)$, ce qui serait absurüe puisque $\lambda(1)$ est fini. Donc $\vec{\mu}$ n'a pas de majorante absolue.

<u>Remarque</u> - Nous avons défini $|\mu|$ pour une $\mathbb{C}$ - mesure réelle μ , mais pas pour une $\mathbb{C}$ - mesure complexe. On peut maintenant le faire. Plus généralement, si $\vec{\mu}$ est Une mesure à valeurs dans un vectoriel $\vec{E}$ de dimension finie sur $\mathbb{R}$ ou $\mathbb{C}$ on peut appeler $|\vec{\mu}|$ sa plus petite majorante absolue; c'est une mesure ≥ 0 et

$$\|\vec{\mu}(\varphi)\| \leq |\mu|(|\varphi|) \text{ . On a } \|\vec{\mu}\| \leq \||\vec{\mu}|\| ;$$

on a l'égalité pour une mesure scalaire ($\mathbb{R}$ - mesure réelle ou $\mathbb{C}$ - mesure complexe), mais pas en général pûur une mesure vectorielle (voir note (*) au théorème 15), donc en particulier pas généralement pour une $\mathbb{R}$ - mesure à valeurs dans $\mathbb{C}$.

§ 3 PROLONGEMENT D'UNE MESURE POSITIVE. THÉORIE DE LEBESGUE

Soit μ une mesure $\geqslant 0$ sur un espace localement compact X. Nous nous proposons de définir $\mu(\varphi)$ pour d'autres fonctions φ que les fonctions continues à support compact. C'est ainsi que si $\mu = dx$, on a pu, par l'intermédiaire de la théorie de Riemann, définir $\int \varphi\, dx$, pour bien d'autres fonctions que les fonctions continues à support compact. Si par ailleurs $\mu = \delta_{(a)}$, on peut définir tout naturellement $\mu(\varphi) = \delta_{(a)}(\varphi) = \varphi(a)$, pour toute fonction complexe φ sur X.

La théorie que nous allons donner ici ou théorie de Lebesgue est considérablement plus générale que la théorie de Riemann; même dans le cas de la mesure $\mu = dx$, nous trouverons bien d'autres fonctions intégrables-Lebesgue que les fonctions intégrables-Riemann antérieurement trouvées. Par contre cette théorie est longue, délicate et souvent fastidieuse, et nous nous permettrons d'admettre un certain nombre de théorèmes *

Nous supposerons une fois pour toutes, sans le répéter constamment **, que X est un espace localement compact, dénombrable à l'infini, c'est-à-dire qui est la réunion d'une infinité dénombrable de compacts. C'est le cas par exemple des espaces vectoriels de dimension finie, puisqu'ils peuvent toujours être considérés comme réunion dénombrable de boules fermées c'est-à-dire de parties co'mpactes. Tous les espaces localement compacts rencontrés usuellement en analyse possèdent cette propriété, c'est donc là une restriction sans grande importance; elle n'est pas indispensable pour toute la théorie, mais pour plusieurs des théorèmes fondamentaux; D'autre part, nous supposerons toujours donnée une fois pour toutes sur X une mesure de Radon $\mu \geqslant 0$. On dit que X est un espace mesuré, si c'est un espace localement compact, dénombrable à l'infini, muni d'une mesure $\mu \geqslant 0$ donnée. Le fait que μ soit $\geqslant 0$ est essentiel.

* L'exposé que nous donnons ici de la théorie de Lebesgue n'est ni le plus court ni le meilleur. Mais : 1°/ il introduit le minimum de notions nouvelles, notamment une étude des fonctions semi-continues inférieurement n'est pas nécessaire; 2°/ il traite d'un seul coup l'intégrale des fonctions vectorielles; 3°/ Tout ce qui est introduit ici est,de toute façon, d'un usage constant en analyse.

** Nous le répéterons si c'est une hypothèse essentielle pour la validité d'un théorème. C'est une restriction qui est déja intervenue au théorème 11,

Mesures extérieures des ensembles ouverts

Soit $\mathcal{O}$ une partie ouverte de X . On appelle mesure extérieure de l'ensemble ouvert $\mathcal{O}$, relativement à μ, la norme $\|\mu\|_{\mathcal{O}}$ de la restriction de μ à cet ouvert, c'est-à-dire la borne supérieure des $\mu(\varphi)$ pour toutes les fonctions φ appartenant à $\mathcal{C}(X)$ de support dêns $\mathcal{O}$ et vérifiant $0 \leq Q \leq 1$. Cette mesure'se notera $\mu^*(\mathcal{O})$.

Le nombre $\mu^*(\mathcal{O})$ est ≥ 0 , fini ou infini.

On démontre sur les mesures des ensembles ouverts un certain nombre de propriétés.

Certaines de ces propriétés sont tout-à-fait évidentes.

Par exemple, si $\mathcal{O}_1$ et $\mathcal{O}_2$ sont deux ouverts, et si l'on a $\mathcal{O}_1 \subset \mathcal{O}_2$, on a bien évidemment $\mu^*(\mathcal{O}_1) \leq \mu^*(\mathcal{O}_2)$, à cause de la positivité de μ .

D'autre part $\mu^*(\emptyset) = 0$, et $\mu^*(X) = \|\mu\|$.

Voici maintenant deux exemples de propriétés moins élémentaires.

1°) Soit $\mathcal{O}_0, \mathcal{O}_1, \mathcal{O}_2, \ldots \mathcal{O}_n, \ldots$ une suite croissante d'ensembles ouverts de X , leur réunion $\mathcal{O}$ est alors encore un ensemble ouvert; on a la formule

$$\mu^*(\mathcal{O}) = \mu^*\left(\bigcup_n \mathcal{O}_n\right) = \lim_{n\to\infty} \mu^*(\mathcal{O}_n) \tag{IV,3;1}$$

Démontrons cette formule.

Tout d'abord, on a toujours $\mu^*(\mathcal{O}) \geq \mu^*(\mathcal{O}_n)$.

Soit alors M un nombre fini quelconque tel que $M < \mu^*(\mathcal{O})$; nous devons simplement montrer qu'il existe un entier n tel que $\mu^*(\mathcal{O}_n) \geq M$. D'après la définition de $\mu^*(\mathcal{O})$, il existe une fonction Q de $\mathcal{C}(X)$, de support dans $\mathcal{O}$, telle que $0 \leq \varphi$ 6 1 , et $\mu(\varphi) \geq M$. Mais les ouverts $\mathcal{O}_n$ coupent le support compact K de φ suivant des ouverts de ce compact, et définissent ainsi une suite croissante d'ouverts sur ce compact. Comme la réunion $\mathcal{O}$ des $\mathcal{O}_n$ recouvre K, il existe déjà un entier n tel que $\mathcal{O}_n$ recouvre le compact K ; alors φ est une fonction appartenant à $\mathcal{C}$ (X) et de support dans $\mathcal{O}_n$ vérifiant $0 \leq \varphi \leq 1$, de sorte que, d'après la définition même 'de $\mu^*(\mathcal{O}_n)$, on a nécessairement $\mu(\varphi) \leq \mu^*(\mathcal{O}_n)$ On a donc bien $\mu^*(\mathcal{O}_n) \geq M$, et ceci démontre notre affirmation.

2°) Si $\mathcal{O}_0, \mathcal{O}_1, \mathcal{O}_2, \dots \mathcal{O}_n, \dots$ est une suite d'ouverts, de réunion $\mathcal{O}$, on a :

$$\mu^*(\mathcal{O}) \leq \sum_{n=0}^{\infty} \mu^*(\mathcal{O}_n) . \tag{IV,3;2}$$

Soit en effet $\varphi \in \mathcal{C}(X)$, de support K dans $\mathcal{O}$, $0 \leq \varphi \leq 1$. Les $\mathcal{O}_n$ forment un recouvrement ouvert de **K ; comme K est** un compact, un nombre fini d'entre eux, $(\mathcal{O}_i)_{i\in I}$, suffit à recouvrir **K** . Soit $(\alpha_i)_{i\in I}$ une partition de l'unité associée à K et aux $\mathcal{O}_i$ (théorème 11). Alors $\alpha_i \varphi$ est dans $\mathcal{C}(X)$, de support dans $\mathcal{O}_i$, et $0 \leq \alpha_i \varphi \leq 1$, donc $\mu(\alpha_i \varphi) \leq \mu^*(\mathcal{O}_i)$.

Donc

$$\mu(\varphi) = \sum_{i\in I} \mu(\alpha_i \varphi) \leq \sum_{i\in I} \mu^*(\mathcal{O}_i) \leq \sum_{n=0}^{\infty} \mu^*(\mathcal{O}_n) . \tag{IV,3;3}$$

En prenant la borne supérieure du ler membre pour toutes les φ considérées, on obtient bien (Iv,3;2)

Remarques 1°/ Supposons que μ soit la mesure atomique $\sum_\nu c_\nu \delta_{(a_\nu)}$, tous les c_ν étant ≥ 0 . Il résulte alors du théorème 15 que l'on a :

$$\mu^*(\mathcal{O}) = \|\mu\|_{\mathcal{O}} = \sum_{a_\nu \in \mathcal{O}} c_\nu \leq +\infty . \tag{IV,3;4}$$

2°/ Si $X = \mathbb{R}$, et si μ est la mesure $f\,dx$, $f \geq 0$ localement intégrable Riemann, si d'autre part $\mathcal{O}$ est l'intervalle ouvert borne $]a,b[$, on a :

$$\mu^*(]a,b[) = \int_{[a,b]} f(x)\,dx \tag{IV,3;5}$$

On a en effet trivialement $\mu^*(]a,b[) \leq \int_{[a,b]} f(x)\,dx$.

Mais on peut trouver a', b',

$a < a' < b' < b$, tels que $\int_{[a,a']} f(x)\,dx$

$$+ \int_{[b',b]} f(x)\,dx \leq ((a'-a)+(b'-b)) \sup_{a\leq x\leq b} |f(x)| \leq \varepsilon .$$

Si alors φ vaut 1 dans $[a', b']$, 0 dans $\complement\,]a'', b''[$, $a'' = \dfrac{a + a'}{2}$, $b'' = \dfrac{b + b'}{2}$, et est affine là ou elle n'a pas encore été définie, on a

$$\mu(\varphi) \geq \int_{[a',b']} f(x)\,dx \geq \int_{[a,b]} f(x)\,dx - \varepsilon \; ;$$

$\varepsilon > 0$ étant arbitraire, cela donne le résultat.

3°/ Si la mesure μ est nulle dans l'ouvert $\mathcal{O}$ (voir définition de la page 444), alors $\mu^*(\mathcal{O}) = 0$, et réciproquement. La mesure du complémentaire du support de μ est nulle.

4°/ Si l'adhérence de l'ouvert $\mathcal{O}$ est compacte, alors la mesure $\mu^*(\mathcal{O})$ est finie. En effet, d'après le corollaire 1 du théorème 11, on peut trouver une fonction θ appartenant à $\mathcal{C}(X)$, ≥ 0 , et $\equiv 1$ sur $\mathcal{O}$. Alors, pour toutes les fonctions φ servant à définir $\mu^*(\mathcal{O})$, on a nécessairement $\mu(\varphi) \leq \mu(\theta)$, et par suite $\mu^*(\mathcal{O}) \leq \mu(\theta)$.

Mesure intérieure d'un compact

Soit K un compact de X . On appelle mesure intérieure de K relativement à μ, et on note $\mu_*(K)$, la borne inférieure des $\mu(\varphi)$, pour toutes les fonctions φ de $\mathscr{C}(X)$, ≥ 0 , et ≥ 1 sur des voisinages de K . Cette mesure est toujours finie, puisque, d'après le corollaire 1 du théorème 11, il existe toujours au moins une fonction φ_0 répondant à la question et par conséquent la borne inférieure cherchée est $\leq \mu(\varphi_0)$,

supposons que K soit un compact contenu dans l'ouvert $\mathcal{O}$. Montrons que $\mu_*(K) \leq \mu^*(\mathcal{O})$.

D'après le corollaire 1 du théorème 11, il existe une fonction φ de $\mathscr{C}(X)$, vérifiant $0 \leq \varphi \leq 1$, égale à 1 sur un voisinage de K et de support dans $\mathcal{O}$. Alors, d'après les définitions mêmes de $\mu_*(K)$ et de $\mu^*(\mathcal{O})$, on a les inégalités $\mu_*(K) \leq \mu(\varphi)$ et $\mu(\varphi) \leq \mu^*(\mathcal{O})$, ce qui démontre notre affirmation.

Exemples - 1°) Si μ est la mesure atomique $\sum_\nu c_\nu \delta_{(a_\nu)}$, où les c_ν sont ≥ 0 , alors

(IV,3;5bis) $$\mu_*(K) = \sum_{a_\nu \in K} c_\nu .$$

En effet, pour toute $\varphi \geq 0$, et ≥ 1 sur un voisinage de K , on a sûrement $\sum_\nu c_\nu \varphi(a_\nu) \geq \sum_{a_\nu \in K} c_\nu$, donc $\mu_*(K) \geq \sum_{a_\nu \in K} c_\nu$. Soit d'autre part $\mathcal{O}$ un voisinage ouvert de K , d'adhérence compacte (note page 440). Alors $\sum_{a_\nu \in \mathcal{O}} c_\nu < +\infty$. Donc il existe un ensemble fini J tel que $\sum_{\substack{\nu \notin J \\ a_\nu \in \mathcal{O}}} c_\nu \leq \varepsilon$. Appelons $\mathcal{O}'$ l'ouvert obtenu en retirant de $\mathcal{O}$ les points a_ν , $\nu \in J$, $a_\nu \notin K$.

Alors

(IV,3;5ter) $$\mu_*(K) \leq \mu^*(\mathcal{O}') = \sum_{a_\nu \in \mathcal{O}'} c_\nu \leq \sum_{a_\nu \in K} c_\nu + \varepsilon \quad ;$$

comme ε est arbitraire, on a bien $\mu_*(K) \leq \sum_{a_\nu \in K} c_\nu$ donc $= \sum_{a_\nu \in K} c_\nu$.

2") Si μ est la mesure $p\,dx$ sur $\mathbb{R}$, avec $p \geqslant 0$, et si K est un intervalle borné $[a,b]$, on a

$$\mu_*([a,b]) = \int_{[a,b]} p(x)\,dx . \tag{IV,3;5 quarto}$$

On procède exactement comme dans 1°). On voit d'abord immédiatement que $\mu_*([a,b] \geqslant \int_{[a,b]} p(x)\,dx$. On choisit ensuite un intervalle ouvert $]a',b'[$ $a'<a<b<b'$, tel que $\int_{[a,a']} p(x)\,dx + \int_{[b,b']} p(x)\,dx \leqslant \varepsilon$. Alors

$$\mu_*([a,b]) \leqslant \mu^*(]a',b'[) = \int_{[a',b']} p(x)\,dx$$
$$\leqslant \int_{[a,b]} p(x)\,dx + \varepsilon ,$$

d'où le résultat,

Ensembles mesurables, mesure des ensembles

Définition - On appelle mesure extérieure d'une partie A de X, relativement à μ, la borne inférieure des mesures extérieures des ouverts contenant A. On la note $\mu^*(A)$. On appelle mesure intérieure de A relativement à μ la borne supérieure des mesures intérieures des compacts contenus dans A. On la note $\mu_*(A)$. On a bien évidemment $\mu_*(A) \leqslant \mu^*(A)$, puisque si un compact K est contenu dans un ouvert $\mathcal{O}$, on a $\mu_*(K) \leqslant \mu^*(\mathcal{O})$.

Si $A \subset B$, on a bien entendu $\mu^*(A) \leqslant \mu^*(B)$, et $\mu_*(A) \leqslant \mu_*(B)$.

Si A est réunion d'un nombre fini ou d'une infinité dénombrable d'ensemble A_n, $n = 0,1,2,\ldots$, on a bien évidemment

$$\mu^*(A) \leqslant \sum_{n\geqslant 0} \mu^*(A_n) . \tag{IV,3;6}$$

C'est évident si l'un des $\mu^*(A_n)$ vaut $+\infty$. Supposons les donc tous finis.

Pour chaque n, on peut trouver un ouvert $\mathcal{O}_n \supset A_n$ tel que $\mu^*(\mathcal{O}_n) \leqslant \mu^*(A_n) + \frac{\varepsilon}{2^{n+1}}$; on a alors, si $\mathcal{O} = \bigcup_n \mathcal{O}_n$:

(IV,3;6 bis)
$$\mu^*(A) \leqslant \mu^*(\mathcal{O}) \leqslant \sum_n \mu^*(\mathcal{O}_n) \text{ (par IV,3;2)}$$
$$\leqslant \sum_n \mu^*(A_n) + \varepsilon ,$$

et comme ε est arbitraire, cela donne bien (IV,3;6).

supposons les A_n disjoints. ~~L'inégalité (IV,3;6) ne devient pas du tout nécessairement une égalité (elle le deviendra si les A_n sont mesurables, voir formule (IV,3; 7 ter)).~~ [Signalons toutefois qu'elle devient une égalité si, non seulement les A_n, mais les $\overline{A}_n$ sont disjoints. Nous ne l'utiliserons qu'après le théorème 19, c'est pourquoi nous le montrerons en utilisant ce théorème.

On peut, pour N donné, trouver des ouverts $\mathcal{O}_n \supset \overline{A}_n$, $n = 0,1,\dots,N$, encore disjoints (corollaire 6 du théorème 11). Pour tout ouvert $\mathcal{O} \supset A$, on a $\mathcal{O} \supset \bigcup_{n=0} (\mathcal{O} \cap \mathcal{O}_n)$

Mais les $\mathcal{O} \cap \mathcal{O}_n$ sont ouverts donc mesurables, et disjoints, et le théorème 19 donnera $\mu(\mathcal{O}) \geqslant \sum_{n=0}^{N} \mu(\mathcal{O} \cap \mathcal{O}_n) \geqslant \sum_{n=0}^{N} \mu^*(A_n)$.

En prenant la borne inférieure pour tous les $\mathcal{O}$, on aura $\mu^*(A) \geqslant \sum_{n=0} \mu^*(A_n)$, d'où, en faisant tendre N vers $+\infty$, $\mu^*(A) \geqslant \sum_{n=0}^{\infty} \mu^*(A_n)$ et par suite

$$\mu^*(A) = \sum_{n=0}^{\infty} \mu^*(A_n)$$

; pour des A_n d'adhérences disjointes]. D'autre part on a des inégalités pour les mesures intérieures, toujours pour des A_n disjoints. Bornons-nous, comme plus haut, au cas où tous les $\mu_*(A_n)$ sont finis, la propriété étant évidente dans le cas contraire. Pour chaque n, on peut trouver un compact $K_n \subset A_n$, tel que $\mu_*(K_n) \geq \mu_*(A_n) - \frac{\varepsilon}{2^{n+2}}$ On peut ensuite, pour N donné, trouver des fonctions continues α_n, $n = 0, 1, 2, \ldots, N$, $0 \leq \alpha_n \leq 1$, α_n prenant la valeur 1 sur un voisinage de K_n, et à supports 2 à 2 disjoints (il suffit de prendre, en appliquant le corollaire 6 du théorème 11, des ouverts $\mathcal{U}_i$ 2 a 2 disjoints contenant les K_i, $i = 0, 1, 2, \ldots, N$, et d'appliquer le corollaire 1 de ce théorème à chaque couple $K_i \subset \mathcal{U}_i$). Soit alors $\varphi \in \mathcal{C}(X)$, $\varphi \geq 0$ et égale a 1 sur un voisinage de $K_0 \cup K_1 \cup \ldots \cup K_N$, telle que $\mu(\varphi) \leq \mu_*(K_0 \cup K_1 \ldots \cup K_N) + \frac{\varepsilon}{2}$.

Chaque $\alpha_n\varphi$ est ≥ 0, égale à 1 sur un voisinage de K_n, donc $\mu(\alpha_n\varphi) \geq \mu_*(K_n)$; mais comme les supports des α_n sont disjoints, $\varphi \geq \alpha_0\varphi + \cdots + \alpha_N\varphi$, donc

$$\mu(\varphi) = \sum_{n=0}^{N} \mu(\alpha_n\varphi) \geq \sum_{n=0}^{N} \mu_*(K_n),$$

ce qui donne

$$\mu_*\left(\bigcup_{n=0}^{N} K_n\right) \geq \sum_{n=0}^{N} \mu_*(K_n) - \frac{\varepsilon}{2}.$$

Finalement $\mu_*(A) \geq \mu_*\left(\bigcup_{n=0}^{N} A_n\right)$

$$\geq \mu_*\left(\bigcup_{n=0}^{N} K_n\right) \geq \sum_{n=0}^{N} \mu_*(K_n) - \frac{\varepsilon}{2} \geq \sum_{n=0}^{N} \mu_*(A_n) - \varepsilon.$$

Comme ceci vaut quel que soit $\varepsilon > 0$, on a

$\mu_*(A) \geq \sum_{n=0}^{N} \mu_*(A_n)$ et comme c'est vrai pour tout N, on a, <u>pour des A_n 2 à 2 disjoints; de réunion A</u> :

$$\mu_*(A) \geq \sum_{n=0}^{\infty} \mu_*(A_n). \tag{IV,3;7 bis}$$

On dit qu'une partie A de X est <u>pseudo-mesurable</u> pour μ, si ses mesures extérieures ~~et intérieures~~ sont égales (finies ou infinies). Leur valeur commune est appelée la mesure de A et notée $\mu(A)$ Nous avons dit pseudo-mesurable et non mesurable, car nous verrons dans quelques instants que cette définition n'est pas assez restrictive : on devra appeler mesurable une partie pseudo-mesurable satisfaisant a une propriété additionnelle.

Nous allons démontrer quelques propriétés de la mesure ainsi définie.

1°) Les ouverts sont pseudo-mesurables.

soit en effet $\mathcal{O}$ un ouvert et soit **M** fini< $\mu^*(\mathcal{O})$. Il existe $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq 1$, à support dans $\mathcal{O}$, telle que $\mu(\varphi) \geq M$.

soit K le support (compact) de φ. Pour toute Ψ de $\mathcal{C}(X)$, ≥ 0, et ≥ 1 sur un voisinage de K on a $\mu(\Psi) \geq \mu(\varphi) \geq M$; en prenant la borne Inférieure pour toutes ces Ψ, on a $\mu_*(K) \geq M$. Comme c'est vrai pour tout M, $\mu_*(\mathcal{O}) = \sup_{K \subset \mathcal{O}} \mu_*(K) = \mu^*(\mathcal{O})$, ce qui prouve que l'ouvert $\mathcal{O}$ est pseudo-mesurable.

2°) Les compacts sont pseudo-mesurables.

Soit K un compact quelconque.

soit $\varphi \in \mathcal{C}(X)$, ≥ 0, et 31 sur un voisinage de K, telle que $\mu(\varphi) \leq \mu_*(K) + \varepsilon$. Soit $\mathcal{O}$ un ouvert $\supset K$ sur lequel $\varphi(x) \geq 1$. Pour toute $\Psi \in \mathcal{C}(X)$, $0 \leq \Psi \leq 1$, Ψ à support dans $\mathcal{O}$, on a $\Psi \leq \varphi$, donc $\mu(\Psi) \leq \mu(\varphi)$; en prenant la borne supérieure, pour toutes les Ψ, on a $\mu^*(\mathcal{O}) \leq \mu(\varphi \leq \mu_*(K) + \varepsilon$. Donc, ε étant arbitraire, $\mu^*(K) = \operatorname{Inf}_{\mathcal{O} \supset K} \mu^*(\mathcal{O}) = \mu_*(K)$,

donc le compact **K** est pseudo-mesurable.

3°) Soit **A** une réunion finie ou dénombrable de parties A_n disjointes pseudo-mesurables. Les inégalités (IV,3;6), (IV,3;7bis), montrent immédiatement que A est pseudo-mesurable et que

(IV,3;7ter) $$\mu(A) = \sum_{n \geq 0} \mu(A_n)$$

(Voir le $\sum$ page 461).

4°) Soient **C** et **D** 2 parties disjointes de **X**

On sait déjà que

(IV,3;7quarto) $$\mu^*(C \cup D) \leq \mu^*(C) + \mu^*(D)$$

Soit $\mathcal{O}$ un ouvert $\supset (C \cup D)$, **K** un compact $\subset$ **C**. Alors $\mathcal{O} - K$ est un ouvert contenant D, donc $\mu(\mathcal{O} - K) \geq \mu^*(D)$. D'autre part $\mathcal{O}$ est réunion des 2 parties pseudo-mesurables disjointes **K** (compact) et $\mathcal{O} - K$ (ouvert); 3°) donne donc

$$\mu(\mathcal{O}) = \mu(K) + \mu(\mathcal{O} - K) \geq \mu(K) + \mu^*(D).$$

En prenant d'abord, pour $\mathcal{O}$ fixé, la borne supérieure pour tous les K, on obtient $\mu(\mathcal{O}) \geq \mu_*(C) + \mu^*(D)$; la borne inférieure pour tous les $\mathcal{O}$ donne l'inégalité remarquable

(IV,3;7 quinto) $$\mu^*(C \cup D) \geq \mu_*(C) + \mu^*(D)$$

(et on pourrait évidemment permuter C et D).

<u>Si en particulier C est pseudo-mesurable,</u> on déduit de (IV,3;7 quarto et quinto) <u>l'égalité remarquable</u> (remarquable parce qu'il est rare qu'on obtienne des égalités où figurent des ensembles non pseudo-mesurables) :

(IV,3;7 sexto) $$\mu^*(C \cup D) = \mu(C) + \mu^*(D) .$$

En ce qui concerne ia mesure intérieure, on a d'abord, par (IV,3;7bis) :

(IV,3;7 septimo) $$\mu_*(C \cup D) \geq \mu_*(C) + \mu_*(D) .$$

Soit K un compact $\subset (C \cup D)$, $\mathcal{O}$ un ouvert $\supset D$. Le compact $K \cap \complement\mathcal{O}$ est contenu dans C, donc sa mesure est $\leq \mu_*(C)$. Mais $K \cup \mathcal{O}$ est la réunion des 2 ensembles pseudo-mesurables disjoints $\mathcal{O}$ (ouvert) et $K \cap \complement\mathcal{O}$ (compact), donc il est pseudo-mesurable, et 3° donne :

$$\mu(K) \leq \mu(K \cup \mathcal{O}) = \mu(\mathcal{O}) + \mu(K \cap \complement\mathcal{O}) \leq \mu(\mathcal{O}) + \mu_*(C) .$$

En prenant successivement la borne supérieure pour tous les K et la borne inférieure pour tous les $\mathcal{O}$, on obtient:

(IV,3;7 octavo) $$\mu_*(C \cup D) \leq \mu_*(C) + \mu^*(D)$$

(et on pourrait évidemment permuter C et D).

Echangeons C et D , puis appliquons (IV,3;7 septimo et octavo); on obtient, si C est mesurable, l'égalité remarquable :

(IV,3;7 noveno) $$\mu_*(C \cup D) = \mu(C) + \mu_*(D) .$$

Alors (IV,3;7 sexto et noveno) montrent que, si C et $(C \cup D)$ sont pseudo-mesurables disjoints, C de mesure finie. alors D est pseudo-mesurable, puisque

$$\mu^*(D) = \mu_*(D) = \mu(C \cup D) - \mu(C) .$$

Ce résultat ne subsiste pas nécessairement si $\mu(C) = \mu(C \cup D) = +\infty$, car la différence est alors indéterminé. Il est d'ailleurs facile de donner un contre-exemple. Nous donnerons plus loin (page 464) un exemple d'un ensemble D non pseudo-mesurable, contenu dans un intervalle borné $[\![a, b]\!]$ de $\mathbb{R}$, pour $\mu = dx$.

Soit alors $C = \complement([a,b])$, qui est ouvert, de mesure $+\infty$; comme $(C \cup D) \supset C$, il est pseudo-mesurable de mesure $+\infty$; cependant D n'est pas pseudo-mesurable.

On peut mettre le résultat obtenu sous une autre forme équivalente, et posant $A = C \cup D$, $B = C$. Si A et B sont pseudo-mesurables, $A \supset B$, et si B est de mesure finie, alors $A - B = A \cap \complement B$ est pseudo-mesurable, et $\mu(A-B) = \mu(A) - \mu(B)$.

Le résultat ne subsiste pas si $\mu(B) = +\infty$. En prenant $A = X$, on voit que, si B est pseudo-mesurable, et de mesure finie, $\complement B$ est pseudo-mesurable (mais pas en général si $\mu(B) = +\infty$).

5") Soit A une réunion finie de parties pseudo-mesurables, $A = \bigcup_{n=0}^{N} A_n$; alors A est pseudo-mesurable (et l'on a évidemment $\mu(A) \leq \sum_{n=0}^{N} \mu(A_n)$ en vertu de (IV,3;6)). Bien noter que 5") diffère de 3°), car les A_i ne sont plus supposées disjointes. La propriété est évidente si l'un des A_n est de mesure infinie, car alors il en est de même de A ; supposons donc toutes les $\mu(A_i)$ finies. soit $\mathcal{O}_n$ un ouvert contenant A_i, tel que $\mu(\mathcal{O}_n) \leq \mu(A_n) + \frac{\varepsilon}{2(N+1)}$, et soit K_n un compact $\subset A_n$ tel que $\mu(K_n) \geq \mu(A_n) - \frac{\varepsilon}{2(N+1)}$.

Alors $\mathcal{O} = \bigcup_{n=0}^{N} \mathcal{O}_n$ est un ouvert contenant A et $K = \bigcup_{n=0}^{N} K_n$ est un compact $\subset A$; on a $\mu^*(A) - \mu_*(A) \leq \mu(\mathcal{O}) - \mu(K) = \mu(\mathcal{O} - K)$

$$\leq \sum_{n=0}^{N} \mu(\mathcal{O}_n - K_n) \left(\text{car } (\mathcal{O} - K) \subset \bigcup_{n=0}^{N} (\mathcal{O}_n - K_n)\right) \leq \sum_{n=0}^{N} \left(\mu(\mathcal{O}_n) - \mu(K_n)\right) \leq \varepsilon ;$$

donc, ε étant arbitraire, A est bien pseudo-mesurable.

6") Soit maintenant A une réunion dénombrable d'ensembles pseudo-mesurables A_i. Posons $B_n = A_0 \cup A_1 \cup \ldots \cup A_n$. Alors les B_n sont pseudo-mesurables d'après 5°), et A est la réunion de la suite croissante des B_n. Mais alors les $C_n = B_n - B_{n-1}$ sont pseudo-mesurables, pourvu que les mesures des B_n soient finies, d'après 4°).

Enfin A est la réunion des ensembles disjoints $B_0, C_1, C_2, \ldots, C_n, \ldots$ donc est pseudo-mesurable d'après 3°). Le résultat subsiste évidemment si l'un des A_n est de mesure infinie, car il en est de même de A. Nous avons donc démontré que : Toute réunion finie ou dénombrable de parties pseudo-mesurable est pseudo-mesurable; si les B_n sont une suite croissante d'ensembles pseudo-mesurables, de réunion B (ici $B = A$), alors (B est pseudo-mesurable d'après ce qui précède) on a $\mu(B) = \lim_{n\to\infty} \mu(B_n)$; en effet, d'après 3°), $\mu(B) = \mu(B_0) + \mu(C_1) + \cdots \mu(C_n) \ldots$, et d'après 4°), $\mu(C_n) = \mu(B_n) - \mu(B_{n-1})$ si toutes sont finies, ce qui donne bien $\mu(B) = \lim_{n\to\infty} \mu(B_n)$ si les $\mu(B_n)$ sont finies; et le même résultat est évident si l'un d'eux est Infini.

7°) Si A est une intersection finie de parties pseudo-mesurables de mesures finies, $A = \bigcap_{n=0}^{N} A_n$, alors A est pseudo-mesurable. Raisonnons en effet comme à 5°), avec les mêmes notations, mais posons cette fois $\mathcal{O} = \bigcap_{n=0}^{N} \mathcal{O}_n$, $K = \bigcap_{n=0}^{N} K_n$

Alors $\mu^*(A) - \mu_*(K) \leq \mu(\mathcal{O}) - \mu(K) = \mu(\mathcal{O} - K)$

$$\leq \sum_{n=0}^{N} \mu(\mathcal{O}_n - K_n) (\text{car } \mathcal{O} - K \subset \bigcup_{n=0}^{N} (\mathcal{O}_n - K_n)) = \sum_{n=0}^{N} (\mu(\mathcal{O}_n) - \mu(K_n)) \leq \varepsilon,$$

ce qui démontre notre affirmation, ε étant arbitraire. Mais ici, contrairement à 5°), le résultat ne subsiste pas si la mesure de l'un des A_n est infinie. Soit en effet D un ensemble non pseudo-mesurable, contenu dans un intervalle borné $[a,b]$ de $\mathbb{R}$ (voir page 464).

soit $A = [a,b]$, pseudo-mesurables; et soit $B = D \cup \complement[a,b]$ qui contient $\complement[a,b]$ et donc est pseudo-mesurable de mesure infinie. Ici $A \cap B = D$, non pseudo-mesurable.

8°) Toute intersection finie ou dénombrable d'ensembles pseudo-mesurables de mesures finies est pseudo-mesurable. Soit en effet $A = \bigcap_{n=0}^{\infty} A_n$. On posera $B_n = B_0 \cap B_1 \ldots \cap B_n$. Alors chaque B_n est pseudo-mesurable d'après 7°), la suite des B_n est décroissante et d'intersection A. Alors $B_n = B_0 - (B_0 - B_n)$; la suite des $B_0 - B_n$ est croissante, on peut donc lui appliquer 6°), d'où le résultat : $B_0 - (B_0 - A) = A$ est pseudo-mesurable. En outre $\mu(B_0 - A) = \lim_{n\to\infty} \mu(B_0 - B_n)$ donc! : si les B_n sont une suite décroissante de parties pseudo-mesurables de mesures finies, d'intersection23 (ici A), on a $\mu(B) = \lim_{n\to\infty} \mu(B_n)$

Les résultats obtenus sont très satisfaisants, sauf sur les points suivants :

Si **A** et B sont pseudo-mesurables, $A \supset B$, mais $\mu(B) = +\infty$, A-B n'est pas nécessairement pseudo-mesurable; l'intersection de deux ensembles pseudo-mesurables n'est pas nécessairement pseudo-mesurable, si l'un d'eux a une mesure infinie. On peut corriger ces défauts en introduisant une définition plus restrictive :

A est dit mesurable si, pour tout compact K, $A \cap K$ est pseudo-mesurable (nécessairement de mesure finie). Un ensemble mesurable est sûrement pseudo-mesurable, autrement dit la propriété est plus forte; en effet, X est réunion d'une suite de compacts K_n, si donc tous les $A \cap K_n$ sont pseudo-mesurables, **A** l'est aussi d'après 6°). D'autre part, tout ensemble pseudo-mesurable de mesure finie est mesurable; car alors $A \cap K$ est pseudo-mesurable d'après 7°). Mais un ensemble pseudo-mesurable de mesure infinie n'est pas nécessairement mesurable; par exemple, si D n'est pas pseudo-mesurable, et est contenu dans $[a,b] \subset \mathbb{R}$ l'ensemble $D \cup \complement [a,b]$ est pseudo-mesurable de mesure infinie, mais non mesurable, puisque son intersection avec le compact $[a,b]$ n'est pas pseudo-mesurable. Alors il est évident que, si **A et** B sont mesurables, $A \supset B$, A -B est toujours mesurable; car $(A-B)\cap K = (A \cap K) - (B \cap K)$. Mais la formule $\mu(A-B) = \mu(A) - \mu(B)$ reste dénuée de sens si $\mu(B) = +\infty$.

D'autre part une réunion ou intersection finie ou dénombrable A d'ensembles mesurables A_n est toujours mesurable; car $A \cap K$ est la réunion ou l'intersection des $A_n \cap K$. D'ailleurs le complémentaire $\complement A$ d'un mesurable A est mesurable; et on passe de la réunion à l'intersection par passage aux complémentaires. Par contre la formule $\mu(B) = \lim_{n\to\infty} \mu(B_n)$ pour une suite décroissante d'ensembles mesurables B_n d'intersection B, suppose toujours les B_n de mesures finies, comme le montre sa démonstration. Si d'ailleurs nous prenons la mesure dx sur $\mathbb{R}$, la suite des ensembles $B_n = \{x \in \mathbb{R} ; |x| \geq n\}$ est décroissante mais tous sant de mesure infinie, et leur intersection est vide, donc de mesure 0.

Ajoutons que tout ensemble fermé F est mesurable, car il est réunion de la suite des compacts $F \cap K_n$, si X est réunion des compacts K_n ; donc, par passage au complémentaire, tout ouvert est mesurable, même s'il est de mesure infinie.

Répétons les définitions données en cours de route :

Définition : Un ensemble $A \subset X$ est pseudo-mesurable, si $\mu^*(A) = \mu_*(A)$ et la valeur commune s'appelle sa mesure et se note $\mu(A)$. Un ensemble A est dit mesurable s'il est pseudo-mesurable, et si en outre, pour tout compact K, $A \cap K$ est pseudo-mesurable.

On a alors démontré le théorème suivant :

Théorème 19 1°/ Tous les ensembles ouverts et tous les ensembles fermés sont mesurables.

2°/ Si A et B sont deux parties mesurables de X, $A \supset B$, l'ensemble $A - B$ est mesurable; et $\mu(A-B) = \mu(A) - \mu(B)$, si $\mu(A)$ et $\mu(B)$ ne valent pas toutes deux $+\infty$. En particulier, si A est mesurable, $\complement A$ est aussi mesurable.

3°/ Si A est la réunion d'un nombre fini, ou d'une infinité dénombrable d'ensembles mesurables, $A = \bigcup_n A_n$, alors A est mesurable.

En outre on a l'inégalité

(IV,3;8)
$$\mu(A) \leq \sum_n \mu(A_n),$$

et cette inégalité devient une égalité, si les ensembles A_n sont deux a deux disjoints

3°/ bis - L'intersection d'un nombre fini ou d'une Infinité dénombrable d'ensembles mesurables est mesurable.

4°/ soit $A_1, A_2, \ldots A_n, \ldots$, une suite croissante d'ensembles mesurables; alors leur reunion A est mesurable, et en outre, on a la formule :

$$\mu(A) = \lim_{n \to \infty} \mu(A_n).$$

5°/ Si $A_0, A_1, \ldots, A_n, \ldots$ est une suite décroissante d'ensembles mesurables, alors leur intersection A est mesurable; en outre, si tous les A_n sont de mesure finie, on a la formule (IV,3;9), cette formule ne subsiste pas si tous les $\mu(A_n)$ sont infinis.

Remarque 1 - Nous démontrerons plus tard une formule nouvelle sur les mesures extérieures : si les A_i sont une suite croissante d'ensembles quelconques de X, de réunion A, on a toujours : $\mu^*(A) = \lim_{n \to \infty} \mu^*(A_i)$ (Voir corollaire 0 du théorème 36). Mais nous devons démontrer maintenant cette égalité dans un cas particulier qui nous sera utile prochainement : pour tout $A \subset X$, on a $\mu^*(A) = \operatorname{Sup} \mu^*(A \cap K)$, Sup étant pris pour tous les compacts K de X. C'est bien un cas particulier, puisque X est réunion d'une suite croissante de compacts K_n. Supposons d'abord $\mu^*(A) < +\infty$.
Il existe alors un ouvert $\mathcal{O} \supset A$ de mesure finie. Les $\mathcal{O} \cap \complement K_n$ forment une suite décroissante d'ensembles mesurables de mesures finies, d'intersection vide; d'après le théorème 19, partie 5°), $\mu(\mathcal{O} \cap \complement K_n)$ tend vers 0 pour n infini. Mais $(A \cap \complement K_n) \subset (\mathcal{O} \cap \complement K_n)$, donc $\mu^*(A \cap \complement K_n)$ tend aussi vers 0. Mais $A = (A \cap K_n) \cup (A \cap \complement K_n)$, donc $\mu^*(A) \leq \mu^*(A \cap K_n) + \mu^*(A \cap \complement K_n)$; donc nécessairement $\mu^*(A \cap K_n) \leq \mu^*(A)$ tend vers $\mu^*(A)$ pour n infini.
Supposons maintenant $\mu^*(A) = +\infty$. Supposons, pour simplifier, X métrisable et les boules fermées compactes (sinon, on utilise un raisonnement analogue à celui de la page 441).

Posons $A_n = A \cap \{x \in X \,;\, n \leq d(x,a) < n+1\}$.
Posons ensuite $B_0 = A_0 \cup A_2 \cup A_4 \ldots$, et $B_1 = A_1 \cup A_3 \cup A_5 \ldots$
Alors $A = B_0 \cup B_1$, donc $\mu^*(A) \leq \mu^*(B_0) + \mu^*(B_1)$, donc l'un des deux vaut $+\infty$, $\mu^*(B_0)$ par exemple. Mais B_0 est réunion d'ensembles d'adhérences **disjointes**, donc on a (voir page 461) : $+\infty = \mu^*(B_0) = \mu^*(A_0) + \mu^*(A_2) + \mu^*(A_4) + \cdots$
Donc la dernière série est divergente. Mais alors

$$\mu^*(A \cap \{x \in X \,;\, d(x,a) \leq 2N+1\}) \geq \mu^*(A_0 \cup A_2 \ldots \cup A_{2N})$$
$$= \mu^*(A_0) + \mu^*(A_2) \ldots + \mu^*(A_{2N}) \text{ tend vers } +\infty = \mu^*(A)$$

pour N infini.

Remarque 2 - Il n'est pas question d'étendre les résultats du théorème 19 à des réunions ou intersections de familles non dénombrables de parties. Par exemple, pour la mesure dx tout $A \subset \mathbb{R}$ (mesurable ou non, de mesure finie ou non) est réunion de ses points, qui sont des ensembles mesurables de mesure nulle,

<u>Exemples divers</u> 1°/ Soit μ la mesure atomique $\sum_\nu c_\nu \delta_{(a_\nu)}$, où tous les c_ν sont $\geqslant 0$; il résulte des remarque 1°/ page 459 et 460 que toutes les parties A de X sont mesurables, et que l'on a la formule

$$\mu(A) = \sum_{a_\nu \in A} c_\nu \leqslant +\infty .$$

2°/ Pour la mesure $\mu\, dx$ sur $\mathbb{R}$, il existe toujours des parties non mesurables, sauf si $\mu\, dx$ est la mesure 0 .

Nous allons donner un exemple d'un ensemble non mesurable dans un cas particulier. Soit X le cercle trigonométrique; il est muni d'une mesure de radon naturelle $\mu = d\,\theta$, bien connue, appelée la mesure angulaire. Voici par exemple comment on peut la définir.

Tout point de X étant repéré par un nombre réel défini à un multiple près de 2π , on peut identifier l'espace $\mathscr{C}(X)$ des fonctions continues sur X (nécessairement à support compact puisque X est compact) avec l'espace des fonctions continues périodiques sur la droite réelle $\mathbb{R}$, de période 2π . Appelons $\mathscr{G}_{2\pi}(\mathbb{R})$ l'espace de ces fonctions, muni de la norme $\|\varphi\| = \operatorname{Max}_{x \in \mathbb{R}} |\varphi(x)|$.

Il est donc équivalent de définir une mesure sur X ou de définir une forme linéaire continue sur $\mathscr{G}_{2\pi}(\mathbb{R})$. La mesure $\mu = d\theta$. est alors définie par la formule

$$\mu(\varphi) = \int_{[a,a+2\pi]} \varphi(x)\, dx , \quad \varphi \in \mathscr{G}_{2\pi}(\mathbb{R}) ,$$

où a peut-être choisi arbitrairement* le résultat n'en dépend pas puisque toutes les fonctions φ ont la période 2π * . Cette mesure $d\theta$ est invariante par les rotations (de centre origine) de X ; cela revient à dire que, quel que soit α , on a la formule

$$\mu(\varphi(x)) = \mu(\varphi(x-\alpha)) , \varphi \in \mathscr{G}_{2\pi}(\mathbb{R})$$

* Nous ramenons cette mesure à la notion connue d'intégrale de Riemann sur $\mathbb{R}$; mais il n'y aurait eu aucune difficulté à faire directement une théorie de l'intégrale de Riemann $\int_X \varphi(\theta)\, d\theta$ sur le cercle trigonométrique X .

Il en résulte aussi naturellement que la p-mesure des ensembles de X est invariante par rotation. Si A est une partie quelconque de X, et si A_α est son transformé par la rotation de l'angle α, on a la formule

$$\mu^*(A_\alpha) = \mu^*(A), \quad \mu_*(A_\alpha) = \mu_*(A)$$

; en outre A_α est mesurable si et seulement si A est mesurable, et on a alors $\mu(A_\alpha) = \mu(A)$.

Désignons alors par α Un nombre incommensurable avec π, choisi une fois pour toutes. Nous dirons que deux points x, y, de X sont congrus si $y - x$ est un multiple entier de α. On a là bien entendu une relation d'équivalence sur X.

Soit $\dot{X}$ l'ensemble quotient; il a bien entendu la puissance du continu comme X lui-même. En effet, d'une part il existe une surjection de X sur $\dot{X}$, donc $\operatorname{card} X \geq \operatorname{card} \dot{X}$. Mais d'autre part chaque classe d'équivalence est dénombrable, on a donc $\operatorname{card} X \leq \nu \operatorname{card} \dot{X}$ et comme $\operatorname{card} \dot{X} \geq$? (le contraire voudrait dire qu'il n'existerait qu'un nombre fini de classes; chacune étant dénombrable X serait dénombrable), on a, d'après le théorème 5 du chapitre 1, $\operatorname{card} X \leq \operatorname{card} \dot{X}$; finalement $\operatorname{card} X = \operatorname{card} \dot{X} = \gamma$ puissance du continu. Soit alors A une partie de X, contenant un point et un seul de chaque classe d'équivalence. (il est bien évident que dans Une classe d'équivalence aucun point ne se présente de façon plus naturelle que les autres, et qu'il n'y aura donc pas de règle permettant de faire simplement ce choix. Nous sommes donc obligés d'utiliser, pour cette infinité de choix ayant la puissance du continu, l'axiome de choix ou axiome de Zermelo (voir (*) page 24)).

Nous allons démontrer que A n'est pas μ- mesurable. En effet, les ensembles $A_{n\alpha}$, $n \in \mathbb{Z}$, sont deux a deux disjoints, puisque A contient au plus un point dans chaque classe d'équivalence ; d'autre part leur réunion est X entier, puisque A contient au moins un point dans chaque classe d'équivalence. Par ailleurs A et $A_{n\alpha}$ sont transformés l'un de l'autre par la rotation $n\alpha$, et par conséquent leurs mesures extérieures aussi bien qu'intérieures coïncident. Si alors A était mesurable, tous les $A_{n\alpha}$ seraient mesurables et de même mesure, et l'on aurait

$$2\pi = \mu(X) = \sum_{n=-\infty}^{+\infty} \mu(A_{n\alpha}) = \mu(A) \sum_{n=-\infty}^{+\infty} 1 .$$

Or cette égalité est impossible; si en effet A était de mesure nulle, on en déduirait que X serait de mesure nulle alors qu'il est de mesure 2π ; si A était de mesure > 0, on en déduirait que X serait de mesure infinie.

Donc A n'est pas mesurable. En outre l'inégalité (IV,3; 7bis) montre que l'on a nécessairement

$$2\pi = \mu_*(X) \geqslant \sum_{n=-\infty}^{+\infty} \mu_*(A_{n\alpha}) = \mu_*(A) \sum_{n=-\alpha}^{+\infty} 1 \ ,$$

donc que la mesure intérieure de **A** est nulle. Mais il n'est pas mesurable, et par conséquent sa mesure extérieure est >0. L'exemple cité ici a un certain caractère "monstrueux. C'est là un fait général pour toutes les mesures non nulles de la forme $p\,dx$ sur la droite $\mathbb{R}$, ou $p\,d\theta$ sur le cercle trigonométrique, p étant intégrable-Riemann : on ne sait pas former explicitement d'ensembles non mesurables, on peut seulement démontrer leur existence par l'axiome de choix. Cela signifie que tous les ensembles rencontrés dans les applications pratiques, et qui sont en général explicitement définis, sont nécessairement mesurables. Un intervalle ouvert, semi ouvert, ou fermé $|a,b|$, est mesurable, et sa mesure est $\int_{[a,b]} p(x)\,dx$.

3°/ Le théorème 19 permet de toute façon de donner une très large catégorie d'ensembles qui sont toujours mesurables. On appelle tribu de parties d'un ensemble X, un ensemble $\mathcal{T}$ de parties ayant les propriétés suivantes :

a) si $A \in \mathcal{T}$, alors $\complement A \in \mathcal{T}$;

b) toute réunion ou toute intersection finie ou dénombrable de parties appartenant à $\mathcal{T}$ appartient à $\mathcal{T}$,

Donnons des exemples de tribus : citons tout d'abord la tribu vide ne contenant aucun ensemble; ensuite la tribu constituée seulement par une partie et la partie complémentaire; ensuite la tribu de toutes les parties de X .

si $\mathcal{T}_1$ et $\mathcal{T}_2$ sont de deux tribus, les parties de X qui appartiennent aux deux à la fois forment encore une tribu; on l'appelle la tribu intersection $\mathcal{T}_1 \cap \mathcal{T}_2$.

On définira de même la tribu intersection $\bigcap_{i \in I} \mathcal{T}_i$ d'une famille quelconque de tribus finie ou infinie. Il en résulte en particulier que, si $\mathcal{S}$ est un ensemble quelconque de parties de X il existe une plus petite tribu qui contienne cet ensemble, à savoir l'intersection de toutes les tribus qui le contiennent. Si alors X est un espace topologique on appelle tribu borélienne la plus petite tribu qui contienne toutes les parties ouvertes de la topologie de X (et par conséquent aussi toutes les parties fermées). Une partie de X est dite borélienne si elle appartient à la tribu borélienne. La tribu borélienne de X ne dépend que de la topologie de X Par exemple, une intersection dénombrable d'ouverts (qui n'est plus ouverte), une réunion dénombrable de fermes (qui n'est plus fermée) sont boréliennes; mais il y en a bien d'autres ! Une partie borélienne est alors "universellement mesurable", en ce sens qu'elle est mesurable pour toutes les mesures de Hadon μ sur X .

En effet, d'après le théorème 19, l'ensemble de toutes les parties μ - mesurables de X est une tribu, qui naturellement dépend de μ ; or cette partie contient nécessairement les parties ouvertes et par conséquent contient la tribu borélienne. Des théorèmes très délicats permettent de montrer qu'il existe d'autres parties de X universellement mesurables que les parties boréliennes.

Ensembles de mesure nulle

Théorème 20 - Pour qu'une partie A de X soit de mesure nulle relativement à μ , il faut et il suffit, que quel que soit $\varepsilon > 0$, il existe un ouvert $\mathcal{O}$ de X contenant A , et tel que $\mu(\mathcal{O}) \leqslant \varepsilon$. Toute réunion d'un nombre fini, ou d'une infinité dénombrable d'ensembles de mesure nulle est encore de mesure nulle.

Démonstration - La première condition signifie en effet que la mesure extérieure $\mu^*(A)$ est nulle; comme il en est alors nécessairement de même de la mesure intérieure, qui est plus petite, l'ensemble est bien mesurable et de mesure nulle. Ce qui est relatif aux réunions finies ou dénombrables résulte alors du théorème 19, et du fait que la somme de toute série dont tous les termes sont nuls, est nulle.

Corollaire - Tout ensemble dénombrable de la droite réelle $\mathbb{R}$ est de mesure nulle, pour la mesure $\mu\,dx$.

Cela résulte en effet du théorème 20, et de ce qu'un ensemble réduit a un point est de mesure nulle.

Remarques 1°/ La mesure de la droite réelle $\mathbb{R}$ toute entière, ou d'un segment $[a,b]$ d'extrémités distinctes, n'étant pas nulle pour la mesure dx, il en résulte une nouvelle démonstration de ce que la droite ou un segment de droite ne sont pas dénombrables. Ces exemples montrent à nouveau — qu'une réunion d'une infinité non dénombrable d'ensembles de mesure nulle, et en particulier d'ensembles réduits à des points, n'a pas nécessairement une mesure nulle.

2°/ Le corollaire pourrait faire croire qu'il n'existe pas, sur la droite $\mathbb{R}$, d'autres ensembles de mesure nulle pour la mesure dx, que les ensembles dénombrables. Il n en est rien. Nous allons construire un ensemble E contenu dans l'intervalle $[0,1]$, qu'on appelle l'ensemble parfait discontinu de Cantor, qui a la puissance du continu et qui est de mesure nulle. Considérons l'ensemble E_1, réunion des intervalles $[0,1/3]$ et $[2/3,1]$. Partageons de nouveau chacun de ces intervalles en trois parties égales, ne retenons que la lère et la 3ème, et appelons E_2 l'ensemble ainsi formé; E_2 est donc la réunion des 4 intervalles $[0,\frac{1}{9}]$, $[\frac{2}{9},\frac{3}{9}]$, $[\frac{6}{9},\frac{7}{9}]$, $[\frac{8}{9},1]$. Partageons de nouveau chacun des 4 intervalles qui constituent E_2 en 3 intervalles égaux, ne retenons que le ler et le 3ème, et appelons E_3 l'ensemble ainsi formé, réunion de 8 intervalles; et ainsi de suite. Nous formons ainsi une suite décroissante d'ensembles fermés $E_1, E_2, \dots, E_n, \dots$; la mesure de E_n est évidemment égale à $\left(\frac{2}{3}\right)^n$; donc l'intersection E de tous les E_n est de mesure nulle. Montrons qu'elle a cependant la puissance du continu. On détermine en effet, d'une manière unique, le point le plus général de E, en choisissant une suite arbitraire $a_1, a_2, \dots, a_n, \dots$, où chacun des a_i est l'un des deux mots : "premier, deuxième".

Soit en effet $x \in E$. Le mot a_1 sera premier ou deuxième, selon que X sera dans le premier intervalle, $[0,\frac{1}{3}]$, ou le deuxième intervalles $[\frac{2}{3},1]$, dont la réunion constitue E_1. Soit par exemple a_1 = premier. Le mot a_2 sera Premier ou deuxième, selon que x est dans le premier sous-intervalle, $[0,\frac{1}{9}]$, ou le deuxième sous-intervalle, $[\frac{2}{9}.\frac{3}{9}]$ du partage de $[0,\frac{1}{3}]$ qui a servi à définir E_2. Et ainsi de suite. D'ailleurs, posons $b_n = 0$ si a_n = premier, $b_n = 2$ si a_n = deuxième; on voit immédiatement que la correspondance entre x et la suite des $b_n = 0$ ou 2 est entièrement déterminée par $x = \sum_{n=1}^{\infty} \frac{b_n}{3^n}$, et que cette formule définit une bijection entre E et l'ensemble $\{0,2\}^{\mathbb{N}}$ des applications de $\mathbb{N}$ dans $\{0,2\}$ ou des suites d'éléments 0 ou 2. Il résulte alors de ce qui a été vu au chapitre I que $card\ E = card\ \mathcal{P}(\mathbb{N}) = 2^{\aleph} = \aleph$ puissance du continu. On peut d'ailleurs très exactement définir l'ensemble E en utilisant la représentation des points de l'intervalle $[0,1]$, non pas décimale mais triadique, c'est-à-dire dans le système de base 3. L'ensemble E est alors exactement l'ensemble des points dont le développement triadique peut se faire en utilisant seulement les nombres 0 et 2, et sans utiliser le nombre 1 (un élément qui a deux développements triadiques appartiendra à E si et seulement si l'un d'eux peut se faire en utilisant seulement 0 et 2). L'ensemble E, décrit pour la première fois par Cantor au siècle dernier, a de plus la propriété suivante : il est <u>parfait</u>; on entend par là que non seulement il est fermé dans $\mathbb{R}$, mais qu'en outre aucun de ses points n'est isolé; tout voisinage de l'un de ses points contient même, comme l'on le démontre aisément, une infinité de points de E ayant la puissance du continu. Voici encore un autre exemple instructif. Soit $a_0, a_1, \ldots, a_n, \ldots$ une suite de points de $\mathbb{R}$ dense (par exemple l'ensemble des nombres rationnels rangés en une suite); et soit $c_0, c_1, \ldots, c_n, \ldots$, une suite de nombres >0 telle que $\sum_{n=0}^{\infty} c_n < +\infty$. Appelons alors E_ε l'ensemble $E_\varepsilon = \bigcup_{n=0}^{\infty} \{x \in \mathbb{R} ; |x - a_n| < \varepsilon c_n\}$. C'est un ensemble ouvert, comme réunion d'ouverts, et sa mesure est naturellement majorée par $2\varepsilon \sum_{n=0}^{\infty} c_n$. Appelons alors E l'intersection de tous les ensembles E_ε, correspondant à tous les $\varepsilon > 0$. C'est un ensemble de mesure nulle. On pourrait croire à priori qu'il est dénombrable et qu'il se réduit à la suite des a_n. Nous laissons aux lecteurs étonnés le soin de montrer que cet ensemble, en réalité, a toujours la puissance du continu.

Si, par ailleurs, nous prenons comme mesure la mesure $dx_1\, dx_2 \ldots dx_n$ sur $\mathbb{R}^n$, alors la mesure d'un ensemble correspond à son volume, et par conséquent une variété différentiable, de dimension $< n$, comme nous le verrons plus tard, est nécessairement de mesure nulle; on a là des exemples très simples d'ensembles de mesure nulle ayant la puissance du continu, 3°/ Pour une mesure de Dlrac $\delta_{(a)}$, l'ensemble $\complement\{a\}$ est de mesure nulle.

Pour la mesure 0 , tout ensemble est de mesure nulle.

<u>**Propriétés** vérifiées presque partout</u>.

Soit P une propriété relative aux points d'un espace localement **compact** X muni d'une mesure de Radon μ 3 0 . * Cette propriété est définie par l'ensemble A des points **de** X qui la vérifient.

* On peut définir une **propriété** P relative aux points **de** X comme une **application de** X dans l'ensemble **à** 2 éléments $\{$ **oui, non** $\}$. On dit que $x \in X$ vérifie P si $P(x) =$ oui.

On dit que cette **propriété** est **vérifiée μ-presque** partout, **ou μ-presque sûrement,** ou que μ - presque tous les points de X vérifient cette propriété, si le complémentaire $\complement A$ de A a une mesure nulle relativement à μ . On pourra se contenter de dire "presque partout", si la mesure μ a été spécifiée une fois pour toutes et si cela ne prête à **aucune** confusion. Remarquons bien que si, pour la mesure $\mu = dx$. la notion de "presque partout" correspond bien à l'intuition, il n'en est plus nécessairement de même pour d'autres mesures; par exemple, pour $\mu = \delta_{(a)}$, masse unité au point a , une propriété est vérifiée presque partout, simplement si elle est vérifiée au point a . Si la mesure μ est nulle, toute propridté est vérifiée presque partout (ainsi d'ailleurs que la négation de cette **propriété**). Voila ce que signifie, pour une fonction f définie sur X , supposé métrique, à valeursdans un espace **métrique** F, le fait d'etre <u>presque partout continue</u> :

$$\left(\exists A \in \mathfrak{P}(X),\ \ \mu\left(\complement A\right) = 0\right) \left(\forall a \in A\right) \left(\forall \varepsilon > 0\right) \cdot$$

$$\left(\exists \eta > 0\right) \left(\forall x \in X,\ d(a, x) \leq \eta\right) :$$

$$\left(d\left(f(a), f(x)\right) \leq \varepsilon\right)$$; ou encore : l'ensemble des points où f est discontinue est de mesure nulle pour μ . **Si** $\vec{F}$ est un-espace vectoriel, dire qu'une application $\vec{f}$ de X dans F est presque partout nulle, signifie que l'ensemble des points x de X où $\vec{f}(x) \neq \vec{0}$, est de mesure nulle.

Sur la droite **réelle,** pour la mesure $\mu = dx$, presque tous les points sont irrationnels et même transcendants.

<u>Théorème 21</u> - <u>**Soient**</u> P_0 , P_1 , ... , P_n ,... u<u>n nombre fini ou une infinité dénombrable de propriétés relatives **à** des points **de** X</u> ; <u>**si chacune d'elles est vérifiée presque partout,**</u> <u>alors</u> <u>**la propriété**</u> $P = P_1 \wedge P_2 \wedge \ldots \wedge P_n , \ldots$, <u>**qui consiste à** vérifier **à** la fois toutes les propriétés</u> P_n , <u>est elle même vérifiée presque part</u>out.

C'est surtout dans la théorie des probabilités que ce théorème est fondamental.

Démonstration. Soit A_n l'ensemble des points qui vérifient P_n; alors l'ensemble des points qui vérifient P est $A = \bigcap_n A_n$, son complémentaire est $\bigcup_n (\complement A_n)$, qui est de mesure nulle, comme réunion d'une infinité dénombrable d'ensembles de mesure nulle.

Naturellement, il n'y a pas d'énoncé analogue pour une infinité non dénombrable de propriétés.

Par exemple, pour $\mu = dx$ sur $\mathbb{R}$, la propriété P_a : x vérifie P_a si $x \neq a$, est vraiepresque partout; la propriété $P = \bigwedge_a P_a$ est la suivante : x est différent de tout point de $\mathbb{R}$; elle n'est vérifiée par aucun point.

Fonctions μ - mesurables à valeurs dans un espace métrisable separable

Définition provisoire.- Soit F un espace métrisable. On dit qu'il est ~~séparable~~, ou de type dénombrable, s'il existe dans F un ensemble dénombrable dense (La dénomination "séparable" est bien fâcheuse; cela n'a aucun rapport avec l'axiome de séparation de Hansdorff).

soit f une application de l'espace localement compact X, muni d'une mesure $\mu \geqslant 0$, dans l'espace métrisable séparable F. On dit que f est μ -mesurable, ou simplement mesurable si aucune confusion n'est à craindre, si l'image réciproque par f de tout ensemble ouvert de F est une partie μ - mesurable de X.

Dans cette définition, la raison pour laquelle nous supposons F séparable n'apparait pas immédiatement. On peut définir les fonctions mesurables à valeurs dans un espace topologique F arbitraire, mais alors on doit prendre une définition plus compliquée, qui n'est équivalente à la précédente que si F est métrisable séparable * . Il résulte de cette définition que toute application continue de X dans F est mesurable, car alors, pour tout ouvert de F l'image réciproque est un ouvert, donc est mesurable d'après le théorème 19, 1°/.

Naturellement, par passage au complémentaire, an pourrait dans la définition remplacer les parties ouvertes par les parties fermées. D'ailleurs l'ensemble des parties de F dont l'image réciproque par f est μ - mesurable est une tribu (voir page 464-2); si f est mesurable, cette tribu contient les ouverts donc les boréliens, et l'image réciproque de tout borélien de F sera μ - mesurable.

* Voir corollaire du théorème 33.

Nous ne répèterons pas, en général, que F est séparable, sauf si c'est essentiel pour la validité d'un théorème.

Soit A une partie de X, et soit φ sa fonction caractéristique; considérée comme application de X dans l'espace topologique séparable $\mathbb{R}$ * ou dans l'espace discret $\{0,1\}$ elle est mesurable, si et seulement si A est mesurable; car l'image réciproque de tout ouvert, par cette fonction caractéristique, est $\emptyset$, ou X, ou A, ou $\complement A$.

Théorème 21bis- Soient f et g deux applications de X dans F. Si ces deux applications sont presque partout égales, et si l'une est mesurable, l'autre est aussi mesurable.

Démonstration-Soit A l'ensemble des points x où l'on a $f(x) = g(x)$; $\complement A$ est un ensemble de mesure nulle par hypothèse, tous les ensembles qu'il contient sont donc de mesure nulle et en particulier mesurables. Supposons alors f mesurable et montrons que g est mesurable. Soit $\mathcal{O}$ une partie ouverte de F ; on passe de $f^{-1}(\mathcal{O})$ à $g^{-1}(\mathcal{O})$ par les deux opérations suivantes : On ajoute d'abord l'ensemble des éléments qui appartiennent à $g^{-1}(\mathcal{O})$ sans appartenir à $f^{-1}(\mathcal{O})$, c'est une partie de $\complement A$, donc une partie mesurable; ensuite on retranche l'ensemble des points qui appartiennent à $f^{-1}(\mathcal{O})$ sans appartenir à $g^{-1}(\mathcal{O})$, c'est encore une partie de $\complement A$, donc mesurable. Cela démontre bien que $g^{-1}(\mathcal{O})$ est mesurable, et que par conséquent la fonction g est mesurable.

Théorème 22 - Soient F et G des espaces métrisables, f une application mesurable de X dans F, g une application continue de F dans G ; composée application $g \circ f$ de X dans G est mesurable.

Démonstration - Si en effet $\mathcal{O}$ est un ouvert de G, l'image réciproque

$$(g\circ f)^{-1}(\mathcal{O}) = f^{-1}(g^{-1}(\mathcal{O}))$$

est l'image réciproque par f de l'ensemble $g^{-1}(\mathcal{O})$, qui est ouvert, et par conséquent elle est mesurable.

Théorème 23 - Soit X un espace localement compact muni d'une mesure $\mu \geq 0$, et soit F un espace métrisable; toute application f de X dans F, qui est limite μ-presque partout d'une suite $f_0, f_1, \dots, f_n, \dots$, d'applications μ-mesurables de X dans F, est aussi mesurable. En particulier, toute application qui est limite presque partout d'une suite d'applications continues est mesurable. Quand nous disons que f

* $\mathbb{R}$ est séparable, puisque l'ensemble dénombrable Q des rationnels est dense.

est limite presque partout de la suite des f_n, nous voulons dire que cette suite est, presque partout, convergente et de limite f ; autrement dit, qu'il existe un ensemble **A** dont le complémentaire est de mesure nulle pour μ, tel que, pour tout x de A, la suite des $f_n(x)$ converge vers $f(x)$ pour n tendant vers $+\infty$.

__Démonstration__ Supposons d'abord que la suite des f_n converge simplement __partout__ vers f Soit B une partie quelconque de F, appelons A_n l'image réciproque $f_n^{-1}(B)$, puis $\mathcal{A}_m$ l'intersection $\bigcap_{n \geq m} A_n$; c'est l'ensemble des x tels que tous les $f_n(x)$, $n \geq m$, soient dans B. Alors les $\mathcal{A}_m$ forment une suite croissante d'ensembles; appelons $\mathcal{A}$ leur réunion.

Si B est une partie ouverte ou fermée, chacun des A_n est mesurable, puisque chacune des f_n est mesurable, et par suite les ensembles $\mathcal{A}_m$ et $\mathcal{A}$ sont eux-mêmes mesurables.

Supposons B __ouvert__, et montrons que, dans ce cas, l'ensemble $\mathcal{A}$ est intercalé entre $f^{-1}(B)$ et $f^{-1}(\bar{B})$:

(IV,3;13) $$f^{-1}(B) \subset \mathcal{A} \subset f^{-1}(\bar{B}) .$$

Tout d'abord il est évident que l'on a $\mathcal{A} \subset f^{-1}(\bar{B})$.

En effet, si $x \in \mathcal{A}$, cela signifie que x appartient à un $\mathcal{A}_m$ pour un m convenable, alors $f_n(x) \in B$, pour $n \geq m$; donc $f(x) \in \bar{B}$ (théorème 15 du chapitre II). Ceci ne suppose pas nécessairement B ouvert.

Si maintenant B est ouvert, montrons que $f^{-1}(B) \subset \mathcal{A}$.

Si x appartient à $f^{-1}(B)$, cela prouve que $f(x)$, limite des $f_n(x)$, est un élément y de B. Comme alors B est ouvert donc est un voisinage de y, il résulte de la notion même de limite qu'il existe un entier m tel que, pour tout $n \geq m$, l'élément $f_n(x)$ appartienne à B ; alors x appartient à A_n pour tout $n \geq m$, par conséquent il appartient à $\mathcal{A}_m$ et a fortiori à $\mathcal{A}$. Nous ne savons donc pas que les images réciproques $f^{-1}(B)$ ou $f^{-1}(\bar{B})$ sont mesurables, mais nous savons que, si B est ouvert, il existe une partie $\mathcal{A}$ intercalée entre les deux, et qui est mesurable.

Soit alors B une partie __fermée__ quelconque de F. Appelons B_k la réunion des boules*ouvertes de ~~rayon~~ $\frac{1}{k}$, de centres

* On choisit sur F une métrique.

contenus dans B ; comme rdunion de boules ouvertes, c'est une partie ouverte. La suite des B_k, $k=1,2,\ldots$, est décroissante et l'intersection des B_k est B ; en effet, elle contient trivialement B , et d'autre part, sis appartient à cette intersection, la distance de x à l'ensemble fermé B est $\leq \frac{1}{k}$ quel que soit k , donc elle est nulle, et par suite x est dans B . Mais, d'autre part, l'intersection des adhérences $\bar{B}_k$ est aussi identique à B . En effet, d'une part, elle contient la précédente; mais, d'autre part, si un point appartient à l'adhérence $\bar{B}_k$ c'est qu'il est limite de points de B_k , alors, d'après la'continuité de la fonction distance (voir chapitre II, page 81) la distance de x à B est $\leqslant \frac{1}{k}$; donc, si x est dans 1 intersection des B_k, sa distance à B est encore nulle, etx est encore dans B. Il en résulte que nous avons pu former deux suites décroissantes d'ensembles : les B_k ouverts, et les $\bar{B}_k$, dont l'intersection est toujours B . Mais, comme B_k est ouvert, il résulte de ce qu'on a vu plus haut qu'il est possible d'intercaler entre $f^{-1}(B_k)$ et $f^{-1}(\bar{B}_k)$ une partie mesurable $\mathfrak{A}^{(k)}$; alors l'intersection des parties mesurables $\mathfrak{A}^{(k)}$, qui est nécessairement mesurable, doit contenir l'intersection des $f^{-1}(B_k)$ et être contenue dans l'intersection des $f^{-1}(\bar{B}_k)$. C'est donc nécessairement $f^{-1}(B)$, qui est par conséquent mesurable, et le théorème est démontré dans ce cas.

Plaçons nous maintenant dans le cas général, où la suite des f_n converge seulement presque partout vers f . Appelons A l'ensemble des x pour lesquels la suite des $f_n(x)$ converge vers $f(x)$. Modifions toutes les fonctions f_n en prenant $f_n(x) = c$, élément fixe quelconque de F , toutes les fois que x appartient à $\complement A$. Les f_n modifiées sont presque partout égales aux f_n Initiales, et par conséquent chacune d'elles est aussi mesurable (théorème 21). Mais maintenant la suite des f_n modifiées converge partout vers une fonction, qui est égale à f en tous les points de A , et à la constante c , en tous les points de $\complement A$. C'est une fonction qui est mesurable d'après ce que nous avons vu plus haut, et comme elle est presque partout égale à f, f est aussi mesurable. Ce théorème admet la très Importante réciproque suivante :

Fonctions μ - étagées

Les fonctions étagdes par rapport à une mesure de Hadon $\mu \geqslant 0$ sur un espace localement compact X , généralisent les fonctions en escalier, introduites pour la définition de l'intégrale de Riemann.

Si F est un ensemble quelconque, on dit qu'une application f de X dans F est μ-étagée, ou simplement étagée si aucune confusion n'est à craindre, s'il existe une partition de X en réunion d'un nombre fini de parties deux à deux disjointes $X_1, X_2, \ldots, X_n$, telles que chacune d'elles soit μ-mesurable, et que, sur chacune d'elles, l'application f soit constante. Une telle partition s'appelle partition admissible pour f. Naturellement il peut exister, pour une application f, une infinité de telles "décompositions"; de même que pour les fonctions en escalier, il en existe une qui est meilleure que les autres, c'est-à-dire pour laquelle le partage de X comprend le plus petit nombre de parties possible; mais ce n'est pas nécessairement à celle-là qu'on s'intéressera. Si $\vec{f}$ et $\vec{g}$ sont deux fonctions étagées à valeurs dans un espace vectoriel $\vec{F}$, alors la somme $\vec{f} + \vec{g}$ et le produit de $\vec{f}$ par une constante scalaire, sont encore étagées.

Cette dernière affirmation est évidente. La première résulte de ce que, si $(X_i)_{i \in I}$ est une partition admissible pour $\vec{f}$ et $(X_j)_{j \in J}$ une partition admissible pour $\vec{g}$, alors chacune des parties $X_i \cap X_j$ est mesurable, et dans chacune d'elles la fonction $\vec{f}$ et la fonction $\vec{g}$ sont constantes, donc aussi la fonction $\vec{f} + \vec{g}$. Autrement dit, les fonctions étagées sur X à valeurs dans un espace vectoriel $\vec{F}$ forment un sous-espace vectoriel de l'espace F^X de toutes les fonctions sur X à valeurs dans F.

Théorème 23 bis - pour qu'une application de X (espace localement compact dénombrable à l'infini) dans un espace métrisable F soit μ-mesurable, il faut et il suffit qu'elle soit limite μ-presque partout d'une suite de fonctions μ-étagées.

Démonstration - Toute fonction étagée est mesurable, parce que l'image réciproque, par cette fonction, de toute partie de F, est une des parties mesurables de X sur lesquelles la fonction est constante. Donc il en est de même d'une limite presque partout d'une suite de fonctions étagées, d'après le théorème 23.

Réciproquement soit f une application mesurable de X dans F métrique séparable. Soit $a_0, a_1, \ldots, a_n, \ldots$ une suite dense dans F, soit $\varepsilon > 0$ fixé. La réunion des boules $B(a_n, \varepsilon)$ est F ; mais elles ne sont pas disjointes. Formons donc la suite $C_{n,\varepsilon} = B(a_n, \varepsilon) \cap \complement B(a_0, \varepsilon) \cap \complement B(a_1, \varepsilon) \cdots \cap \complement B(a_{n-1}, \varepsilon)$

Les $C_{n,\varepsilon}$ sont maintenant disjointes et de réunion F. Si $D_{n,\varepsilon} = f^{-1}(C_{n,\varepsilon})$, les $D_{n,\varepsilon}$ sont disjointes et de réunion X. En outre $C_{n,\varepsilon}$ est borélienne, donc $D_{n,\varepsilon}$ mesurable. Soit K un compact de X. Les $D_{n,\varepsilon} \cap K$ ont pour réunion K, et sont disjoints et mesurables. La série $\sum_{n=0}^{\infty} \mu(D_{n,\varepsilon} \cap K)$ est donc convergente, de somme $\mu(K)$. On peut donc, pour tout $\delta > 0$, déterminer un entier $N(\varepsilon, K, \delta)$, dépendant de ε, K, δ, tel que la réunion des $D_{k,\varepsilon} \cap K$, $k > N$, soit de mesure $\leq \delta$; soit $E(\varepsilon, K, \delta)$ la réunion des $D_{k,\varepsilon}$, $k > N(\varepsilon, K, \delta)$; $\mu(E(\varepsilon, K, \delta) \cap K) \leq \delta$.

Nous déterminerons f_n comme suit. Soit K_n une suite croissante de compacts, tels que X soit réunion des intérieurs $\mathring{K}_n$ (voir lemme 3 du théorème 11, la suite des B_n ; on sait que, pour tout compact K de X, il existe un entier m tel que $K \subset K_n$ pour $n \geq m$. Posons

$$\begin{cases} f_n(x) = a_k, \text{ si } x \in D_{k,\frac{1}{n}}, & k \leq N\left(\frac{1}{n}, K_n, \frac{1}{2^{n+1}}\right) ; \\ = a \quad \text{ailleurs} & \end{cases}$$

f_n est, étagée, par construction, puisque les $D_{k,\frac{1}{n}}$ pour $k = 0, 1, 2, \ldots$, sont disjoints mesurables de réunion X. Si $x \in D_{k,\frac{1}{n}}$, $k \leq N\left(\frac{1}{n}, K_n, \frac{1}{2^{n+1}}\right)$, on a $f_n(x) = a_k$, et $f(x) \in B(a_k, \frac{1}{n})$, puisque $f(D_{k,\frac{1}{n}}) \subset C_{k,\frac{1}{n}} \subset B(a_k, \frac{1}{n})$; donc $d(f(x), f_n(x)) \leq \frac{1}{n}$. Donc $f_n(x)$ ne convergera pas vers $f(x)$ pour n infini, seulement si, pour une infinité de n, x est dans la réunion des $D_{k,\frac{1}{n}}$, $k > N\left(\frac{1}{n}, K_n, \frac{1}{2^{n+1}}\right)$, c'est-à-dire dans $E\left(\frac{1}{n}, K_n, \frac{1}{2^{n+1}}\right)$. L'ensemble E des points de X, où $f_n(x)$ ne converge pas vers $f(x)$ pour n infini, vérifie donc $E \subset E\left(\frac{1}{n}, K_n, \frac{1}{2^{n+1}}\right)$ pour une infinité de n ; donc, quel que soit m, $E \subset \bigcup_{n \geq m} E\left(\frac{1}{n}, K_n, \frac{1}{2^{n+1}}\right)$.

Soit K un compact quelconque de X, et soit m un entier tel que $K_m \supset K$. Alors $E \cap K \subset \bigcup_{n \geq m} \left(E\left(\frac{1}{n}, K_n, \frac{1}{2^{n+1}}\right) \cap K_n\right)$, donc $\mu(E \cap K) \leq \sum_{n \geq m} \frac{1}{2^{n+1}} = \frac{1}{2^m}$.

Mais on peut remplacer m par n'importe quel entier plus grand, donc $\mu(E \cap K) = 0$; ceci étant vrai pour tout compact K de X, $\mu(E) = 0$, et f_n converge presque partout vers f pour n infini.

Corollaire 1 - Soient $F_1, F_2, \dots, F_m$, des espaces métriques, X Pour qu'une application f produit $F_1 \times F_2 \times \dots \times F_m$, définie par des applications f_i de X dans les F_i, soit mesurable, il faut et il suffit que chacune des applications composantes f_i soit mesurable. En particulier, pour qu'une application de X dans un espace affine de dimension finie, muni d'un référentiel, soit mesurable, il faut et il suffit que chacune de ses composantes le soit.

Démonstration - Il est bien évident que, si f est mesurable, chacune des applications composantes f_i est mesurable; en effet, elle s'obtient en composant f avec la projection de F sur F_i qui est une application continue, et il suffit alors d'appliquer le théorème 22.

La réciproque est d'un caractère plus profond; elle s'appuie sur le théorème 23 bis Par hypothèse, f_i est une application mesurable; donc il existe une suite $(f_i)_n$ d applications étagées de X dans F_i, qui convergent vers f_i, pour n tendant vers $+\infty$, sauf sur un ensemble B_i de mesure nulle de X Si alors nous appelons B la réunion des B_i, il est de mesure nulle; en tout point x de $\complement B$, la suite des $f_n(x) = ((f_1)_n(x), (f_2)_n(x), \dots, (f_m)_n(x))$ converge vers $f(x)$, d'après la définition de la topologie produit; par conséquent f est limite presque partout d'une suite de fonctions étagées, donc elle est mesurable.

Corollaire 2 - Si $\vec{F}$ est un espace vectoriel normé, si $\vec{f}$ et $\vec{g}$ sont des applications mesurables de X dans $\vec{F}$, et si k est une fonction scalaire mesurable, alors les applications $\vec{f}+\vec{g}$ et $k\vec{f}$ sont mesurables. Les fonctions μ - mesurables "sur X, a valeurs dans $\vec{F}$, forment un espace vectoriel.

Démonstration - L'application $x \to \vec{f}(x) + \vec{g}(x)$ est composée de l'application $x \longrightarrow (\vec{f}(x), \vec{g}(x))$ de X dans $\vec{F} \times \vec{F}$, qui est mesurable d'après le corollaire?, et de l'application $(\vec{f}, \vec{g}) \to \vec{f} + \vec{g}$, de $\vec{F} \times \vec{F}$ dans $\vec{F}$, qui est continue; donc elle est mesurable d'après le théorème 22. De même, $x \to k(x)\vec{f}(x)$ est composée de l'application mesurable $x \to (k(x), \vec{f}(x))$ de X dans $\mathbb{K} \times \vec{F}$, et de l'application continue $(k, \vec{f}) \to k\vec{f}$ de $\mathbb{K} \times \vec{F}$ dans $\vec{F}$.

Corollaire 3 - Soient $\vec{F}, \vec{G}, \vec{H}$, des espaces vectoriels normés, $\vec{f}$ et $\vec{g}$ des applications mesurables de X dans $\vec{F}$ et $\vec{G}$ respectivement, B une application bilinéaire continue de $\vec{F} \times \vec{G}$ dans $\vec{H}$. Alors l'application $B(\vec{f}, \vec{g}) : x \to B(\vec{f}(x), \vec{g}(x))$ de X dans $\vec{G}$ est mesurable.

Démonstration - Elle est composée de l'application mesurable $x \to (\vec{f}(x), \vec{g}(x))$ de X dans $\vec{F} \times \vec{G}$ (corollaire 1), et de l'application continue B de $\vec{F} \times \vec{G}$ dans $\vec{H}$, donc mesurable d'après le théorème 22 * .

Les différents théorèmes précédents permettent de comprendre que toutes les fonctions rencontrées dans la pratique soient mesurables. Si X est la droite, et $\mu = dx$, personne n'a jamais pu donner explicitement une fonction réelle non mesurable. Mais, bien entendu, en utilisant l'axiome de choix. on peut montrer l'existence de fonction réelles non mesurables, par exemple la fonction caractéristique de l'ensemble A construit page - 2

Fonctions boréliennes

On dira qu'une application f d'un espace topologique X dans un espace topologique F est borélienne si l'image réciproque par f de toute partie ouverte de F est borélienne. Comme les parties de F dont l'image réciproque par f est boré.lienne forment évidemment une tribu de parties de F , celle-ci ne peut contenir les ouverts (ou les fermés) sans contenir tous les boréliens : on peut donc; dans la définition, remplacer les ouverts de F par les fermés ou les ou les boréliens de F . Cette notion ne dépend que des topologies de X et de F , par exemple, une partie de X est borélienne si et seulement si sa fonction

* L'hypothèse "B est bilinéaire" est superflue. Nous donnons ce corollaire parce qu'il est utile dans la pratique.

caractéristique est borélienne. On voit donc que, si X est localement compact dénombrable à l'infini et F métrisable séparable, une fonction borélienne sur X à valeurs dans F est universellement mesurable, c'est-à-dire μ-mesurable sur X pour toute mesure de radon $\mu \geq 0$.

Les fonctions boreliennes ont les propriétés suivantes :

1°/ La composée de deux applications boréliennes est borélienne : si f est borelienne de X dans Y, g borélienne de Y dans Z, $g \circ f$ est borélienne de X dans Z; car l'image réciproque d'un ensemble borélien est borelienne.

2°/ Toute limite d'une suite d'applications boréliennes de X dans F métrisable, est borelienne. La démonstration est la même que celle du théorème 23, en remplaçant partout mesurable par borélien.

3°/ Il ne semble pas que toute fonction borélienne soit limite simple d'une suite de fonctions étagées-boréliennes (on entend par là une fonction qui ne prend qu'un nombre fini de valeurs, chacune sur un ensemble borélien; c'est sans rapport avec les mesures de Radon). Reprenons les notations de la démonstration du théorème 23 bis. Soit u_n la fonction égale à a_n dans $D_{k,\frac{1}{n}}$; elle est borélienne, parce que les $B(a_k,\frac{1}{n})$, donc les $C_{k,\frac{1}{n}}$, donc les $D_{k,\frac{1}{n}}$ sont boréliens et $d(f,u_n) \leq \frac{1}{n}$; mais elle a une infinité dénombrable d'étages, les $D_{k,\frac{1}{n}}$ donc n'est pas étagée-borélienne. Mais u_n est limite simple pour m infini d'une suite de fonctions $v_{n,m}$, étagées-boréliennes, où $v_{n,m} = a_k$ dans $D_{k,\frac{1}{n}}$, pour $k \leq m$, et $v_{n\,m} = a \in F$ ailleurs; donc toute fonction f borélienne à valeurs dans F métrisable séparable est limite uniforme d'une suite de fonctions boreliennes u_n, qui sont, chacune, limite simple d'une suite de fonctions étagées-boréliennes $v_{n\,m}$.

4°/ Pour qu'une application f de X dans un produit $F_1 \times \dots \times F_m$ soit borelienne, il faut, et il suffit si les F_i sont métrisables séparables, que chacune des composantes f_i de f soit borélienne de X dans F_i. En effet chaque f_i est composée de f et de la projection de F sur F_i, continue donc f_i est bien borélienne si f est borélienne. Inversement supposons les f_i boréliennes et les F_i métrisables séparables, On pourra reproduire la démonstration du corollaire 1 du théorème 23 bis, mais avec deux passages à la limite successifs en utilisant ce que nous venons de voir à 3°/, et cela montrera que f est borélienne.

Remarque - On peut alors notablement améliorer le théorème 23 bis

1°/ soit d'abord A un ensemble mesurable, de mesure finie. Il existe une suite décroissante d'ouverts $\mathcal{O}_n \supset A$, tels que $\mu(\mathcal{O}_n) \leq \mu(A) + \frac{1}{n}$; l'intersection A^* des $\mathcal{O}_n$ est un ensemble borélien contenant A, et $\mu(A^* - A) = 0$. En remplaçant les $\mathcal{O}_n$ ouverts $\supset A$ par des compacts $K_n \subset A$, on construit un borélien $A_* \subset A$, avec $\mu(A - A_*) = 0$. Si A n'est pas de mesure finie, on le représentera comme la réunion des $A \cap K_n$, où les K_n sont une suite croissante de compact: de réunion X, rt, dans tous les cas, on pourra trouver 2 ensembles boréliens A^*, A_*, tels que $A_* \subset A \subset A^*$, et $\mu(A^* - A_*) = 0$. On peut dire qu'à un ensemble de μ - mesure nulle près, tout ensemble μ - mesurable est borélien. Ceci est intéressant parce que les ensembles boréliens sont universellement mesurable et ne dépendent que de la topologie de X. On peut encore dire : la tribu des parties μ - mesurables de X (qui dépend de μ) est la plus petite tribu contenant tous les ensembles boréliens (qui ne dépendent que de la topologie de X) et les parties de μ - mesure nulle.

2°/ Soit maintenant f une fonction μ - mesurable, à valeurs dans l'espace métrique séparable F. Dans la construction des fonctions étagées f_n, au théorème 23 bis, on peut, au lieu de prendre $f_n = a_k$ sur $D_{k,\frac{1}{n}}$, prendre $f_n = b_k$, où b_k est n'importe quel point de $B(a_k, \frac{1}{n})$, On peut alors choisir b_k dans l'image $f(X) \subset F$ (car, si $D_{k,\frac{1}{n}}$ n'est vide $f(X) \cap B(a_k, \frac{1}{n})$ n'est pas vide) : on peut donc supposer que toutes les f_n prennent leurs valeurs dans

3°/ Soient A les étages de f_n. Chacun est réunion d'un borélien $B_{i,n}$ et d'un ensemble de mesure nulle. La fonction g_n, égale à f_n sur les $A_{i,n}$ et à n'importe quel point fixe de $f(X)$ ailleurs, est alors borélienne et presque partout égale à f_n ; les g_n peuvent donc entièrement remplacer les f_n. Les g_n convergent seulement presque partout vers f. Soit N l'ensemble des points où elles ne convergent pas. Soit N^* un ensemble borélien, encore de mesure nulle, contenant N. Soit h_n la fonction égale à g_n sur $\complement N^*$, à un élément fixe de $f(X)$ sur N^*. Alors les h_n sont encore boréliennes à valeurs dans $f(X)$, mais maintenant elles convergent partout vers une limite h, presque partout égale à f et à valeurs dans $f(X)$. La limite d'une suite de fonctions boréliennes étant borélienne h est borélienne. Ainsi on a finalement :

<u>Toute fonction f sur X , a valeurs dans F , mesurable, est presque partout égale à une fonction borelienne Prenant ses valeurs dans $f(X)$, qui elle-même est limite Partout d'une suite de fonctions étagées-boréliennes prenant leurs valeurs dans $f(X)$</u> .

$4°/$ voici un autre genre d'amélioration. Supposons que, F soit compact, ou simplement que $f(X)$ soit contenu dans un compact de F . Alors, pour tout n , un nombre fini des boules $B(a_k, \frac{1}{n})$ le recouvre. On peut prendre toutes ces boules en nombre fini, pour construire f_n , au lieu d'en choisir un nombre fini $N(\frac{1}{n}, K_n, \frac{1}{2^{[illegible]}})$ comme dans le cas gknéral. Alors on a $d(f_n, f) \leq \frac{1}{n}$: <u>les</u> f_n <u>étagées convergent partout et uniformément vers</u> f . On pourra, même dans ce cas, les prendre telles que $f_n(X) \subset f(X)$, mais on ne pourra pas les prendre boréliennes (sauf si f elle-même est borélienne, auquel cas, d'après leur construction même, elles sont boréliennes).

Supposons par exemple que $F = \mathbb{C}$ et que f soit borélienne bornée à support compact K . Elle Prend ses valeurs dans un compact de $\mathbb{C}$, donc elle est limite uniforme d'une suite de fonctions étagées boréliennes, bornées dans leur ensemble, et à support dans un compact fixe (car on peut toujours les multiplier par la fonction caractéristique de K).

Intégrale d'une fonction vectorielle étagée

Si $\vec{f}$ est μ-étagée, a support compact, on appelle <u>intégrale de 'la fonction étagée $\vec{f}$</u> par rapport à la mesure μ, la quantité suivante :

$$\int \vec{f}\, d\mu = \sum_{i \in I} \vec{f}_i \, \mu(X_i) , \tag{IV,3;14}$$

relative à une partition admissible pour la fonction $\vec{f}$; $\vec{f}_i$ est la valeur constante de $\vec{f}$ sur X_i. On voit immédiatement, comme nous l'avons déjà fait pour l'intégrale de Riemann, qu'une telle quantité est inddpendante de la partition choisie pour représenter $\vec{f}$; l'intégrale de la fonction caractéristique d'un ensemble mesurable d'adhérence compacte est sa mesure; on volt de même aussitôt que l'intégrale, ainsi définie, est une application linéaire de l'espace des fonctions étagées à support compact à valeurs dans $\vec{F}$, dans l'espace vectoriel $\vec{F}$. D'autre part, si $\vec{f}$ est une fonction étagée, nulle dans le complémentaire d'un ensemble mesurable **A** , on a immédiatement la majoration

$$\left\| \int \vec{f}\, d\mu \right\| \leq \int \| \vec{f} \| \, d\mu \leq \||| \vec{f} \||| \, \mu(A) . \tag{IV,3;15}$$

Si $\vec{f}$ et $\vec{g}$ sont étagées, et presque partout égales, elles ont même intégrale.

Intégrale supérieure d'une fonction réelle ≥ 0

Définition - Soit X un espace localement compact, dénombrable à l'infini, muni d'une mesure de Radon $\mu \geq 0$; soit f une fonction réelle ≥ 0 définie sur X, à valeurs finies ou infinies * . Si f est bornée et à support compact, on appelle intégrale supérieure de f par rapport à μ, la borne inférieure des intégrales, par rapport à μ, des fonctions g étagées à support compact $\geq f$ Si maintenant f est quelconque, on appelle intégrale supérieure de f par rapport à μ, la borne supérieure, finie ou égale à $+\infty$, des intégrales supdrieures des fonctions ≥ 0 bornées, à support compact, qui sont $\leq f$.

L'intégrale supérieure de f par rapport à μ se note $\int^* f\,d\mu$, ** ou $\int^* f(x)\,d\mu(x)$, ou $\int^* f$ si aucune confusion n'est à craindre.

Remarque - Soient M un nombre ≥ 0 et K un compact de X ; appelons $f_{M,K}$, la fonction égale à $f(x)$ en tous les points x appartenant au compact K et pour lesquels $f(x) \leq M$, à M aux points x de K où $f(x) > M$, et à 0 en dehors de K. Alors $f_{M,K}$ est bornée à support compact, et son intégrale supérieure se définit d'après la première partie de la définition. Alors la deuxième partie est manifestement équivalente à la suivante: l'intégrale supérieure de f est la borne supérieure des intégrales supdrieures des $f_{M,K}$.

* le mot "fonction réelle' est alors abusif, elle est à valeurs dans $\overline{\mathbb{R}}$ et non dans $\mathbb{R}$.

** La différence essentielle entre l'intégrale supérieure de Riemann (relativement à dx) et l'intégrale supérieure de Lebesgue (relativement à $\mu \geq 0$ quelconque), est que la lère s'obtient par un seul passage à la limite : une borne inférieure; alors que la 2ème nécessite 2 passages à la limite: une borne supérieure et une borne inférieure. Ici nous avons fait 4 passages à la limite : 2 pour définir les ensembles mesurables, et 2 maintenant. Mais ce n'est là qu'une apparence, due à l'exposé, qui, comme nous l'avons dit (note * page 457), n'est pas le meilleur; on peut, en fait. se ramener à 2 passages à la limite; mals pas a 1.

On verra plus loin la relation précise entre l'intégrale supérieure de Riemann et celle de Lebesgue : pour une fonction $f \geq 0$, bornée à support compact, sur $\mathbb{R}$, et pour $\mu = dx$, l'intégrale supérieure de Riemann est $\geq$ à l'intégrale supérieure de Lebesgue.

On peut aussi dire que, si $M_0, M_1, \dots, M_n, \dots$, est une suite de nombres tendant vers $+\infty$ lorsque n tend vers $+\infty$, et $K_0, K_1, \dots, K_n, \dots$, une suite croissante de compacts dont la réunion est X (une telle suite existe toujours puisque X est supposé dénombrable à l'infini), alors l'intégrale supérieure de f est la limite, pour n tendant vers $+\infty$, des intégrales supérieures des f_{M_n, K_n}.

<u>Théorème 24</u> - <u>Si et g sont presque partout égales, alors leurs intégrales supérieures sont égales.</u>

<u>Démonstration</u> - Supposons d'abord que f et g, presque partout égales, soient toutes les deux bornées par un nombre $M \geq 0$ et de supports contenus dans un compact K. Soit B l'ensemble, de mesure nulle, des points x où $f(x) \neq g(x)$.

Soit alors f_1 une fonction étagée $\geq f$ · si on appelle g_1 l'enveloppe supérieure de f_1 et du produit de M par la fonction caractéristique de B, alors g_1 est étagée, $g_1 \geq g$, et, comme f_1 et g_1 sont presque partout égales, on a $\int^* g \leq \int g_1 = \int f_1$. En prenant alors la borne inférieure des intégrales des f_1 ; on voit qu'on a nécessairement $\int^* g \leq \int^* f$. Le même raisonnement en sens inverse, montre alors que l'on a $\int^* f = \int^* g$.
Supposons maintenant! et g quelconques. Soit alors f_2 une fonction bornée à support compact, $\leq f$; si nous appelons g_2 l'enveloppe inférieure de g et de f_2, on a $g_2 \leq g$; en outre f_2 et g_2 sont presque partout égales, bornées, à support compact, et par suite $\int^* g \geq \int^* g_2 = \int^* f_2$.

En passant à la borne supérieure des intégrales supérieures des f_2, on en déduit $\int^* g \geq \int^* f$. Le même raisonnement en sens inverse montre alors que $\int^* g = \int^* f$.

<u>Corollaire</u> - <u>Si une fonction</u> $f \geq 0$ <u>est presque partout nulle, son intégrale supérieure est nulle.</u>

<u>Théorème 25</u> - <u>Soient f et g deux fonctions</u> ≥ 0 <u>sur</u> X, et k <u>une constante scalaire. Alors, si</u> $f \leq g$ <u>on a</u> $\int^* f \leq \int^* g$. <u>On a toujours</u>

$$(\text{IV},3;16) \qquad \int^{*}(f+g) \leqslant \int^{*} f + \int^{*} g \,, \quad \text{et} \quad \int^{*} k f = k \int^{*} f$$

La première formule est l'inégalité de convexité. Contrairement à ce qui se passe pour l'intégrale de Riemann elle s'étendra en une inégalité de convexité dénombrable * et c'est ce qui fait la supériorité essentielle de l'intégrale de Lebesgue sur l'intégrale de Riemann.

Démonstration - Ces propriétés sont évidentes; démontrons par exemple la première formule (IV,3;16).

Supposons d'abord! et g bornées à support compact. Alors, à d'évidentes modifications près, la démonstration est exactement la même que celle qui a été donnée pour l'intégrale de Riemann au théorème 2. Prenons maintenant le cas général où f et g sont quelconques. Si alors on Introduit les fonctions $f_{M,K}$ définies plus haut, on a l'inégalité

$$(\text{IV},3;17) \qquad (f+g)_{M,K} \leqslant f_{M,K} + g_{M,K} \,, \qquad \text{d'où}$$

$$\int^{*}(f+g)_{M,K} \leqslant \int^{*} f_{M,K} + \int^{*} g_{M,K} \leqslant \int^{*} f + \int^{*} g \,,$$

et ceci entraîne Immédiatement le résultat, en prenant la borne supérieure de $\int^{*}(f+g)_{M,K}$ pour toutes les valeurs de M, et tous les compacts K de X.

On voit que, si A est une partie quelconque de X, et φ_A sa fonction caractéristique, on a

$$(\text{IV},3;18) \qquad \int^{Y} \varphi_A = \mu^{*}(A) \quad {}^{**}$$

* Corollaire 3 du thdorème 36.

** C'est évident, si $\overline{A}$ est compact. Si A est quelconque, on le voit aussitôt par la formule $\mu^{*}(A) = \sup_{K \subset A} \mu^{*}(A \cap K)$ (K compacts) de la remarque 1 après le théorème 19.

En particulier, $\int^* 1$ est la mesure de X, c'est-&-dire la norme de y- .

Théorème **26.** Si l'intégrale supérieure de $f \geq 0$ est finie, la fonction f est presque partout finie; si Pintegrale supérieure de f est nulle, la fonction f est presque partout nulle.

Démonstration - Soit d'abord $\int^* f < +\infty$.

Appelons **A**, l'ensemble des points x où $f(x) \geq$ n . SI alors φ_n est sa fonction caractéristique, on a $f \geq n\varphi_n$, on en déduit l'inégalité

$$\int^* f \geq n\, \mu^*(A_n)\ . \tag{IV,3;19}$$

Si, a fortiori, on appelle **A** l'ensemble des points x où $f(x) = +\infty$, on a l'inégalité

$$\int^* f \geq n\, \mu^*(A)\ . \tag{IV,3;20}$$

Cette inégalité étant vraie quel que soit n, on voit bien que $\mu^*(A) = 0$, ce qui montre bien que f a presque partout une valeur finie.

Supposons maintenant $\int^* f = 0$. Appelons cette fois B_k l'ensemble des points x où $f(x) \geq \frac{1}{k}$, φ_k sa fonction caractéristique. Alors $f \geq \frac{1}{k}\varphi_k$, par conséquent on a l'inégalité

$$0 = \int^* f \geq \frac{1}{k}\mu^*(B_k)\ , \tag{IV,3;21}$$

qui montre que $\mu^*(B_k) = 0$. Alors l'ensemble B des points x où $f(x) > 0$ est la réunion de l'infinité dénombrable des ensembles B_k de mesure nulle, et par suite il est lui-même de mesure nulle, et f est bien presque partout nulle.

Intégrabilité des fonctions à valeurs vectorielles

Définition - Soient X un espace localement compact muni d'une mesure de Radon $\mu \geq 0$, et $\vec{F}$ un espace vectoriel normé sur le corps des réels ou des complexes. On dit qu'une fonction $\vec{f}$, définie sur X, à valeursdans $\vec{F}$, est

μ-intégrable, ou intégrable relativement à μ, ou simplement intégrable si aucune confusion n'est à craindre, si, quel que soit $\varepsilon > 0$ il existe une fonction $\vec{g}$, définie sur X, à valeurs dans $\vec{F}$, μ-étagée, à support compact, telle que l'on ait l'inégalité

$$\int^* \| \vec{f} - \vec{g} \| \leq \varepsilon . \tag{IV,3;22}$$

Il est encore équivalent de dire qu'il existe une suite $\vec{f}_1, \vec{f}_2, \dots, \vec{f}_n, \dots,$ de fonctions étagées à support compact, définies sur X, à valeurs dans $\vec{F}$, telles que les quantités $\int^* \| \vec{f} - \vec{f}_n \|$ tendent vers 0 pour n tendant vers l'infini. Une telle suite s'appelle suite d'approximation de $\vec{f}$ pour l'intégrale de Lebesgue par rapport à μ. Au lieu de intégrable, on dit aussi sommable.

Si deux fonctions, définies sur X, à valeurs dans $\vec{F}$ sont presque partout égales, et si l'une est intégrable, il en est de même de l'autre, avec les mêmes suites d'approximation (théorème 24).

Remarques. Les remarques 1°, 2°, 3°, des pages 407, 408 peuvent se répéter ici, en remplaçant "intégrable-Riemann, relativement à dx sur $\mathbb{R}$", par "à support compact, et intégrable-Lebesgue relativement à μ sur X" (la restriction "à support compact" étant inutile pour la remarque 1°) et "en escalier" par "μ-étagée". Si nous voulons y référer, nous dirons : remarque 1°, 2°, ou 3°, page 478.

Intégrale de Lebesgue d'une fonction à valeurs vectorielles

Théorème 27 - Soit $\vec{f}$ une application de X dans l'espace de Banach $\vec{F}$. Si $\vec{f}$ est intégrable, et si les $\vec{f}_n$ forment une suite d'approximation de $\vec{f}$ par des fonctions étagées à support compact, alors les intégrales $\int \vec{f}_n$ ont une limite, qui est indépendante du choix de la suite d'approximation.

Démonstration analogue à celle du théorème 3.

Définition - Dans les conditions du précédent théorème, la limite commune des intégrales $\int \vec{f}_n$, pour toutes les suites d'approximation, s'appelle intégrale de la fonction $\vec{f}$ par rapport à μ.

On la note par les diverses notations

$$\int \vec{f}\,d\mu,\ \int \vec{f}\mu,\ \int \vec{f}(x)\,d\mu(x)\ ,\ \int \vec{f}\ .^{*} \tag{IV,3;23}$$

Si deux fonctions intégrables sont presque partout égales, elles ont même intégrale, puisqu'elles ont les mêmes suites d'approximation.

Si $\vec{f}$ est une application d'une partie Y de X dans $\vec{F}$, on dit que $\vec{f}$ est intégrable sur Y, si la fonction $\tilde{\vec{f}}$ sur X, égale à $\vec{f}$ sur Y et à $\vec{0}$ ailleurs, est intégrable, et dans ces conditions, on pose

$$\int_Y \vec{f} = \int \tilde{\vec{f}}\ ^{**}\ . \tag{IV,3;24}$$

Remarque - Naturellement toute fonction étagée à support compact est intégrable, et son intégrale n'est autre que celle qui a été définie à la formule (IV,3;14).

Théorème 28 - Si $\vec{f}$ et $\vec{g}$ sont deux fonctions intégrables sur X à valeurs dans $\vec{F}$, k une constante scalaire, alors $\vec{f}+\vec{g}$ et $k\vec{f}$ sont aussi intégrables; en outre on a l'égalité

$$\int k\vec{f} = k\int \vec{f}\ ,\qquad \int (\vec{f}+\vec{g}) = \int \vec{f} + \int \vec{g}\ . \tag{IV,3;25}$$

Démonstration analogue à celle du théorème 4.

Ce théorème prouve que les fonctions μ-intégrables sur X, à valeurs dans $\vec{F}$, forment elles-mêmes un espace vectoriel par rapport au corps des scalaires $\mathbb{K}$, et que l'intégrale est une application linéaire de cet espace vectoriel dans $\vec{F}$.

Corollaire 1 - Si Y et Z sont 2 parties de X, $\vec{f}$ est intégrable sur leur réunion, si et seulement si elle est intégrable sur chacun d'eux, et alors

$$\int_{Y\cup Z} \vec{f} = \int_Y \vec{f} + \int_Z \vec{f}\ . \tag{IV,3;26}$$

Démonstration analogue à celle du corollaire du théorème 4.

* Il y a tellement de notations différentes de l'intégrale de $\vec{f}$ par rapport à μ, qu'on pourrait presque dire : toute formule où figurent une fonction vectorielle $\vec{f}$ et une mesure $\mu \geq 0$ veut dire l'intégrale de $\vec{f}$ par rapport à μ.

** Si $\vec{f}$ est déjà définie partout sur X, $\tilde{\vec{f}}$ n'est autre que le produit $\vec{f}\varphi_Y$ de $\vec{f}$ par la fonction caractéristique φ_Y de Y.

Corollaire 2 - Si $\vec{f}$ est une fonction étagée, $\vec{f} = \sum_i \vec{f_i}\varphi_{X_i}$ où les $\vec{f_i}$ sont des constantes, et les X_i des ensembles mesurables de mesure finie (mais n'ayant pas nécessairement une adhérence compacte) de fonctions caractéristiques φ_{X_i}, alors $\vec{f}$ est intégrable et on a encore (IV,3;14). En particulier, si A est un ensemble mesurable et de mesure finie, on a toujours $\int \varphi_A = \mu(A)$.

Démonstration - A cause du théorème 28, il suffit de montrer la dernière affirmation. Soit K_n une suite croissante de compacts, de réunion X.

On a $\mu(A) = \lim_{n\to\infty} \mu(A \cap K_n)$ (théorème 19) et $\lim_{n\to\infty} \mu(A - A\cap K_n) = 0$. Par ailleurs $\varphi_{A\cap K_n}$ est étagée à support compact, donc $\int \varphi_{A\cap K_n} = \mu(A\cap K_n)$, et (d'après (IV,3;18)) $\int^*(\varphi_A - \varphi_{A\cap K_n}) = \mu^*(A - A\cap K_n) = \mu(A - A\cap K_n)$ tend vers 0 ; donc les $\varphi_{A\cap K_n}$ forment une suite d'approximation pour φ_A, donc φ_A est intégrable, et $\int \varphi_A = \lim_{n\to\infty}\int \varphi_{A\cap K_n} = \lim_{n\to\infty} \mu(A\cap K_n) = \mu(A)$.

Théorème 29 - Si la fonction $\vec{f}$, définie sur X, à valeurs dans $\vec{F}$, est intégrable, alors la fonction $\|\vec{f}\|$, définie par $x \to \|\vec{f}(x)\|$, est aussi intégrable, et en outre on a l'inégalité

$$\left\| \int \vec{f} \right\| \leq \int \|\vec{f}\| . \tag{IV,3;27}$$

Démonstration analogue à celle du théorème 5.

Corollaires - On peut répéter ici les corollaires du théorème 5 ; dans les corollaires 3 et 3 bis, on peut remplacer $[a,b]$ par une partie mesurable quelconque Y de X, et $(b-a)$ par $\mu(Y)$. Nous y référerons en disant : corollaires 1, 2, 3, 3bis, 4, 5, du théorème 29.

Théorème 30 - Si φ est une fonction, à valeurs complexes, appartenant à $\mathcal{C}(X)$, alors elle est integrable par rapport à μ, et son integrale $\int\varphi$ coïncide avec $\mu(\varphi)$.

Ce théorème est évidemment essentiel, il indique que la construction assez compliquée que nous avons faite de l'intégration, donne bien un prolongement de la forme linéaire initiale $\varphi \longrightarrow \mu(\varphi)$ définie sur L&(X) .

Démonstration - Démontrons d'abord que toute $\varphi \in \mathcal{C}(X)$ est intégrable, et plus généralement que toute fonction $\vec{f}$ sur X, à valeurs dans $\vec{F}$, continue à support compact, est intégrable. Soit $\mathcal{V}$ un voisinage compact de K et soit $\varepsilon > 0$. Pour tout a de K, soit $\mathcal{V}_a \subset \mathcal{V}$ un voisinage de a tel que, pour $x \in \mathcal{V}_a$, on ait $\|\vec{f}(x) - \vec{f}(a)\| \leq \frac{\varepsilon}{\mu(\mathcal{V})}$. Un nombre fini d'entre eux recouvre K, soit $\mathcal{V}_{a_1}, \mathcal{V}_{a_2}, \ldots, \mathcal{V}_{a_n}$. Posons $C_1 = \mathcal{V}_{a_1}$, $C_2 = \mathcal{V}_{a_2} \cap \complement\, \mathcal{V}_{a_1}, \ldots, C_n = \mathcal{V}_{a_n} \cap \complement(\mathcal{V}_{a_1} \cup \mathcal{V}_{a_2} \cup \ldots \cup \mathcal{V}_{a_{n-1}})$. Chacun des C_i est borélien, donc mesurable. Appelons $\vec{f}_\varepsilon$ la fonction égale à $\overrightarrow{f(a_i)}$ dans C_i, à $\vec{0}$ dans le complémentaire de la réunion des $\mathcal{V}_{a_i}$. Alors $\vec{f}_\varepsilon$ est étagée, à support compact, et $\|\vec{f} - \vec{f}_\varepsilon\| \leq \frac{\varepsilon}{\mu(\mathcal{V})}$, donc $\int^* \|\vec{f} - \vec{f}_\varepsilon\| \leq \varepsilon$; ε étant arbitraire, $\vec{f}$ est bien intégrable.

Il faut maintenant montrer que, pour toute $\varphi \in \mathcal{C}(X)$, $\mu(\varphi) = \int \varphi\, d\mu$. Nous aurons besoin de 2 lemmes.

Lemme 1 - Soit A une partie mesurable de X , φ une fonction continue à support compact, majorée par la fonction caractéristique χ_A de A . Alors $\mu(\varphi) \leq \mu(A)$.

Démonstration du lemme 1 - On a $\int \varphi \leq \int \chi_A = \mu(A)$; mais ce qui nous intéresse est $\mu(\varphi)$ et non $\int\varphi$, et nous ne connaissons pas encore leur égalité.

soit $\mathcal{O}$ un ouvert $\supset A$, si nous montrons que $\mu(\varphi) \leq \mu(\mathcal{O})$, nous aurons le résultat et prenant la borne inférieure pour tous les $\mathcal{O}$. Tout revient donc à oublier A, et à le remplacer par $\mathcal{O}$ ouvert. On peut ensuite supposer $\mathcal{O}$ compact; car, si $\mathcal{V}$ est un voisinage compact du support (supposé compact) de φ, on peut remplacer $\mathcal{O}$ par $\mathcal{O} \cap \mathcal{V}$, d'adhérence compacte. On peut enfin supposer $\varphi \geq 0$, sans quoi on la remplace par φ^+. Soit alors K_n le compact $\left\{x \in X \,;\, \varphi(x) \geq \frac{1}{n}\right\}$, $n \geq 1$. Soit α_n une fonction continue, $0 \leq \alpha_n \leq 1$ à support dans $\mathcal{O}$, égale à 1 sur K_n (corollaire 1 du théorème 11).

Posons $\varphi_n = \alpha_n \varphi$. Alors φ_n est continue; à support dans $\mathcal{O}$, $0 \leqslant \varphi_n \leqslant 1$, donc $\mu(\varphi_n) \leqslant \mu(\mathcal{O})$, (définition page 458). Mais $|\varphi - \varphi_n| = (1-\alpha_n)\varphi \leqslant \frac{1}{n}$,
et cette fonction a son support dans $\overline{\mathcal{O}}$; donc $\mu(\varphi-\varphi_n) \leqslant \frac{1}{n}\|\mu\|_{\overline{\mathcal{O}}}$ et par suite $\mu(\varphi) \leqslant \mu(\varphi_n) + \frac{1}{n}\|\mu\|_{\overline{\mathcal{O}}}$
$\leqslant \mu(\mathcal{O}) + \frac{1}{n}\|\mu\|_{\overline{\mathcal{O}}}$; n étant arbitraire cela démontre le lemme.

<u>Lemme 2</u> - soit $\vec{g}$ une fonction vectorielle étagée à support compact, à N étages, majorée en norme par M. Quel que soit $\delta > 0$, il existe une fonction $\vec{\gamma}$ continue, à support compact, décomposable * telle que :

1°) $\|\vec{g} - \vec{\gamma}\| \leqslant NM$

2°) $\vec{\gamma} = \vec{g}$, sauf sur un ensemble de mesure $\leqslant \delta$;

3°) si g est scalaire, donc $\gamma \in \mathcal{C}(X)$, alors $\left|\int \gamma - \mu(\gamma)\right| \leqslant \delta$;

4") si $g \geqslant 0$, alors $\gamma \geqslant 0$.

<u>Démonstration du lemme 2</u> -

soit $\vec{g} = \sum_{i=1}^{N} \vec{g}_i \chi_{A_i}$, χ_{A_i} fonction caractéristique de l'étage A_i de $\vec{g}$, sur lequel $\vec{g}$ est la constante $\vec{g}_i \neq \vec{0}$. Soit K_i un compact, $\mathcal{O}_i$ un ouvert, tels que $K_i \subset A_i \subset \mathcal{O}_i$, $\mu(\mathcal{O}_i) - \mu(K_i) \leqslant \frac{\delta}{NM'}$, $M' = Max(M,1)$.
soit γ_i une fonction continue, $0 \leqslant \gamma_i \leqslant 1$, à support dans $\mathcal{O}_i$, égale à 1 sur un voisinage de K_i . Posons

$$\vec{\gamma} = \sum_{i=1}^{N} \vec{g}_i \gamma_i \, .$$

1°) $0 \leqslant \chi_{A_i} \leqslant 1$, $0 \leqslant \gamma_i \leqslant 1$, donc $|\gamma_i - \chi_{A_i}| \leqslant 1$,
et $\|\vec{g} - \vec{\gamma}\| \leqslant \sum_{i=1}^{N} \|\vec{g}_i\| \, |\chi_{A_i} - \gamma_i| \leqslant NM$.

2°) On a $\gamma_i = \chi_{A_1} = 1$ sur K_i , $= 0$ sur $\complement \mathcal{O}_i$, donc partout sauf sur $\mathcal{O}_i - K_i$, de mesure $\leqslant_N \frac{\delta}{MM'} \leqslant \frac{\delta}{N}$;
donc $\vec{g} = \vec{\gamma}$ sauf sur un ensemble $\subset \bigcup_{i=1}^{N} (\mathcal{O}_i - K_i)$, de mesure $\leqslant N \frac{\delta}{N} = \delta$ ..

* Voir cette notion au corollaire 8 du théorème 11.

<u>Remarque</u> - De 1° et 2°) on déduit

$$\int \| \vec{g} - \vec{\gamma} \| \leq \delta N M .$$

3°) Soit $\vec{F} = \complement \int \gamma_i$ et $\mu(\gamma_i)$ sont toutes deux comprises entre $\mu(K_i)$ et $\mu(\mathcal{O}_i)$, donc

$\left|\int \gamma_i - \mu(\gamma_i)\right| \leq \mu(\mathcal{O}_i) - \mu(K_i) \leq \frac{\delta}{NM'} \leq \frac{\delta}{NM}$. Alors

$$\left|\int \gamma - \mu(\gamma)\right| \leq \sum_{i=1}^{N} |g_i| \left|\int \gamma_i - \mu(\gamma_i)\right| \leq M \frac{\delta}{NM} N = \delta .$$

4°) Si les g_i sont ≥ 0, γ est ≥ 0.

<u>Démonstration du théorème</u> - soit $\varphi \in \mathcal{L}(X)$. soit $\varepsilon > 0$. Il existe des fonctions g et h étagées à support compact, $h \geq 0$, telles que $|\varphi - g| \leq h$, et $\int h \leq \frac{\varepsilon}{2}$. (remarque 2° page 478). Soit M le maximum des modules de 1, φ, g, h et N le nombre maximum d'étages de g, h.

Appliquons le lemme 2, successivement à g, et à $h \geq 0$, pour $\delta = \frac{\varepsilon}{2(1+5NM)}$; on trouve 2 fonctions continues à support compact, α, et $\beta \geq 0$.

On a

$$\left|\int \varphi - \mu(\varphi)\right| \leq \left|\int \varphi - \int \alpha\right| + \left|\int \alpha - \mu(\alpha)\right| + \left|\mu(\alpha) - \mu(\varphi)\right| . \tag{IV,3;29}$$

1°) $|\varphi - \alpha| \leq |\varphi - g| + |g - \alpha| \leq h + |g - \alpha|$.

On a $\int h \leq \frac{\varepsilon}{2}$. Ensuite $g - \alpha$ est nulle, sauf sur un ensemble de mesure $\leq \delta$ où elle est majorée en module par NM ; donc $\int |g - \alpha| \leq \delta NM$. Finalement

$$\left|\int \varphi - \int \alpha\right| \leq \frac{\varepsilon}{2} + \delta NM. \tag{IV,3;30}$$

$$\text{2°)} \quad \left|\int \alpha - \mu(\alpha)\right| \leq \delta . \tag{IV,3;31}$$

3°) $|\mu(\alpha) - \mu(\varphi)| = |\mu(\alpha - \varphi)| \leq \mu(|\alpha - \varphi|)$.

Mais $|\alpha - \varphi| \leq |\varphi - g| + |g - \alpha| \leq h + |g - \alpha|$ $\leq \beta + |h - \beta| + |g - \alpha|$. Ensuite $|g - \alpha|$ est nulle sauf sur un ensemble de mesure δ où elle est majorée par NM ; et $|h - \beta|$ est nulle, aauf sur un ensemble de mesure δ' où elle est majorée par NM . Donc $|\alpha - \varphi| - \beta$ est une fonction continue, à support compact, partout, sauf sur un ensemble mesurable de mesure $\leqslant 2\delta$ sur lequel elle est majorée par 2 NM . En appliquant le lemme 1 à $\frac{|\varphi-\alpha|-\beta}{2NM}$, on obtient

(IV,3;32) $$\left|\mu(\alpha)-\mu(\varphi)\right| \leqslant 2\delta \,.\, 2NM = 4NM\delta \,.$$

On déduit de (IV,3;30, 31 et 32) :

(IV,3;33) $$\left|\int\varphi-\mu(\varphi)\right| \leqslant \frac{\varepsilon}{2} + \delta NM + \delta + 4\delta NM \leqslant \varepsilon$$

d'aprés le choix de δ .

ε étant arbitraire, on a bien $\int\varphi = \mu(\varphi)$.

Remarque - Il est assez ridicule d'avoir à démontrer ce theoreme; c'est le châtiment de la méthode choisie pour faire le prolongement de Lebesgue (voir * page 457).

Corollaire 1 - Une mesure de Radon $\mu \geqslant 0$ sur X est connue si l'on connait les mesures des ouverts (ou des compacts), *

Tout d'abord les formules $\mu(K) = \operatorname*{Inf}_{\sigma \supset K} \mu(\sigma)$, $\mu(\sigma) = \operatorname*{Sup}_{K \subset \sigma} \mu(K)$, montrent que, si l'on connait les mesures des ouverts, on connaît celles des compacts, et vice versa. On connaît alors les mesures de tous les ensembles, donc les intégrales de toutes les fonctions. Alors $\mu(\varphi) = \int\varphi\, d\mu$ montre qu'on connait $\mu(\varphi)$ pour Q $\in \mathcal{C}(X)$, donc μ est connue.

Corollaire 2 - Une mesure de Radon $\mu \geqslant 0$ sur la droite réelle $\mathbb{R}$ est connue, si l'on connait les mesures des intervalles ouverts (ou des intervalles fermés).

Tout d'abord les formules

$$\mu([a,b]) = \lim_{\varepsilon\to 0} \mu(]a-\varepsilon, b+\varepsilon[) \,,$$

$$\mu(]a,b[) = \lim_{\varepsilon\to 0} \mu([a+\varepsilon, b-\varepsilon])$$ montre que l'un des cas se ramène à l'autre.

Ensuite on peut montrer aisément que tout ouvert de $\mathbb{R}$ est réunion d'une infinité dénombrable d'intervalles ouverts disjoints; alors on connait les mesures de tous les ouverts, et on applique le corollaire 1.

* Nous ne voulons pas dire qu'on peut choisir arbitrairement les mesures des ouverts; mais que 2 mesures de Radon $\geqslant 0$, donnant aux ouvertsoles mêmes mesures, coïncident.

Par exemple la théorie de l'intégrale de Riemann (§ 1) a permis de construire la mesure de Radon dx sur $\mathbb{R}$, qui donne à tout intervalde $|a, b|$ la mesure $b - a$; le présent corollaire dit que c'est la seule à avoir cette propriété.

Remarque Soit maintenant φ une fonction scalaire continue à support non compact, mais telle que le support de μ et le support de φ aient une intersection K compacte. Nous avons défini $\mu(\varphi)$ (théorème 16), et cela même si μ est une mesure vectorielle. Il est facile de voir que, si μ est $\geqslant 0$, φ est μ -intégrable, et que ce $\mu(\varphi)$ coïncide encore avec $\int \varphi \, d\mu$.

Si, en effet, $\alpha \in \mathcal{C}(X)$ est égale à 1 sur un voisinage de **K**, nous avons défini $\mu(\varphi)$ comme $\mu(\alpha\varphi)$; mais $\alpha\varphi$ est dans $\mathcal{C}(X)$, donc intégrable, et alors $\mu(\alpha\varphi) = \int(\alpha\varphi)\, d\mu$ d'après le théorème **30**; mais $\alpha\varphi$ et φ sont μ- presque partout égales, puisqu'elles sont égales sur le support de μ *, donc φ est bien intégrable et $\int \varphi \, d\mu = \int(\alpha\varphi)\, d\mu = \mu(\alpha\varphi) = \mu(\varphi)$.

Théorème **31** - 1°/ Si f est une fonction réelle $\geqslant 0$ bornée, à support compact, sur $\mathbb{R}$, son intégrale supérieure de Riemann pour dx est au moins égale à son intégrale superieur de Lebesgue.

2°/ Si $\vec{f}$ est une fonction définie sur la droite réelle $\mathbb{R}$, à valeurs dans $\vec{F}$ et intégrable-Riemann pour la mesure dx, alors elle est a fortiori intégrable-Lebesgue, et son intégrale au sens de Lebesgue est égale à son intégrale au sens de Riemann.

* En tout point du support de μ, on a ou bien $\alpha = 1$ (si ce point est dans **K**) ou bien $\varphi = 0$, donc toujours $\alpha\varphi = \varphi$. D'autre part, nous avons bien dit page 459 remarque 3°/, que le complémentaire du support est de mesure nulle.

<u>Démonstration</u> - Démontrons d'abord 1°/. Appelons $\int^{*R} f$ (resp. $\int^{*L} f$) l'intégrale supérieure au sens de Riemann (resp. Lebesgue). Comme toute fonction en escalier est a fortiori étagée, en se rappelant les définitions de ces 2 intdgrales supérieures, on voit aussitôt que

$$\int^{*L} f \leqslant \int^{*R} f .$$

<u>L'inégalité n'est pas nécessairement une égalité, on peut avoir $<$.</u> Prenons par exemple la fonction f , égale à 1 en tous les points rationnels de $[0,1]$, et à 0 ailleurs. Nous avons vu pages 405 et 406 que $\int^{*R} f = 1$; or l'ensemble des rationnels est de mesure nulle, donc $\int^{*L} f = 0$.

Soit maintenant $\vec{f}$ une fonction intégrable-Riemann, $\vec{f}_n$ une suite d'approximation formée de fonctions en escalier à support compact, donc a fortiori étagées. Alors

$$\int^{*L} \|\vec{f} - \vec{f}_n\| \leqslant \int^{*R} \|\vec{f} - \vec{f}_n\|$$ tend vers 0 , donc les $\vec{f}_n$ forment a fortiori une suite d'approximation pour l'intégrale de Lebesgue, donc $\vec{f}$ est intégrable-Lebesgue, et son intdgrale $\int^{L} \vec{f}$, limite des $\int^{L} \vec{f}_n = \int^{R} \vec{f}_n$, est $\int^{R} \vec{f}$.

Au contraire <u>la fonction $f \geq 0$ vue ci-dessus est intégrable Lebesgue,</u> et d'intégrale nulle (elle est presque partout nulle), <u>mais n'est pas intégrable Riemann</u>.

<u>Théorème 32</u> - <u>Toute fonction continue sur X , à valeurs dans $\vec{F}$, a support compact, est integrable.</u> <u>Soit $\vec{f}$ une fonction définie sur X à valeurs dans $\vec{F}$. Pour [redacted] inté[redacted]rable, [redacted]l faut et il suffit que, quel que soit $\varepsilon > 0$, il existe une fonction $\vec{g}$, définie sur X , à valeurs dans $\vec{F}$, continue, à support compact, telle que</u> $\int^{*} \|\vec{f} - \vec{g}\| \leqslant \varepsilon$; <u>ou encore, il faut et il suffit qu'il existe une suite d'approximation de $\vec{f}$ formée par des fonctions continues à support compact.</u>

<u>Ces fonctions peuvent même être choisies décomposables * .</u>

<u>Démonstration</u> - L'intégrablIité des fonctions continues- à support compact a été démontrée au début de la démonstration du théorème 30. Donc toute fonction ayant une suite d'approximation formée de fonctions continues à support compact est intégrable.

* Voir définition au corollaire 8 du théorème 11.

Soit maintenant $\vec{f}$ intégrable quelconque. Il existe $\vec{g}$ étagée à support compact, telle que $\int \| \vec{f} - \vec{g} \| \leq \frac{\varepsilon}{2}$.
Ensuite, d'après le lemme 2 de la démonstration du théorème 30, il existe $\vec{\gamma}$ continue, à support compact décomposable, telle que $\int \| \vec{g} - \vec{\gamma} \| \leq \delta N M$; et choisissant $\delta = \frac{\varepsilon}{2NM}$, on a $\int \| \vec{g} - \vec{\gamma} \| \leq \frac{\varepsilon}{2}$,
d'où $\int \| \vec{f} - \vec{\gamma} \| \leq \varepsilon$, c.q.f.d.

<u>Remarque-</u>Ce théorème est particulièrement avantageux si $\vec{F}$ est le corps des complexes, car alors il donne $\int f$ comme une limite de quantités $\mu(g)$, lesquelles résultent immédiatement de la définition de μ comme forme linéaire sur $\mathscr{C}(X)$.

Intégrabilité et intégrales des fonctions définies presque partout

Soit d'abord f une fonction réelle ≥ 0, définie seulement <u>presque partout</u> sur X. Prolongeons là de n'importe quelle manière, en une fonction $\tilde{f} \geq 0$ définie partout sur X. Alors $\int^* \tilde{f}$ est indépendant du prolongement choisi, parce que deux de ces prolongements sont des fonctions presque partout égales. La valeur commune s'appelle l'intégrale supérieure de f et se note $\int^* f$. De la même manière, si $\vec{f}$ est une fonction définie seulement presque partout sur X, à valeurs dans $\vec{F}$, et qu'on appelle $\tilde{f}$ un prolongement quelconque de cette fonction, en une fonction définie sur X tout entier, à valeurs dans $\vec{F}$, l'intégrabilité de $\tilde{f}$ et la valeur de son intégrale ne dépendent pas du prolongement choisi.

Si $\vec{\tilde{f}}$ est intégrable, on dit que $\vec{f}$ est intégrable, et la valeur commune des intégrales $\int \vec{\tilde{f}}$ s'appelle intégrale de $\vec{f}$ et se note $\int \vec{f}$.

§ 4 THÉORÈME DE CONVERGENCE DE LEBESGUE. L'ESPACE L^1

Le théorème de convergence que nous allons donner est <u>beaucoup plus puissant</u> que tous ceux qui ont été donnés à propos de l'intégrale de Riemann, et c'est lui qui fait la supériorité essentielle de l'intégrale de Lebesgue sur l'intégrale de Riemann; des conditions, tout juste un peu plus restrictives que la convergence simple des fonctions, suffiront à entraîner la convergence des intégrales. Nous allons d'abord donner un théorème qui montre comment on peut se ramener de la convergence simple à une espèce de convergence uniforme.

Théorème 33 (Egoroff). Soit f_n une suite d'applications mesurables de X dans un espace métrique F, et supposons que cette suite converge simplement μ-presque partout vers. Alors, quel que soit le compact K de X, et le nombre $\delta > 0$, il existe une partie compacte K_δ de K, telle que $\mu(K - K_\delta) \leq \delta$, et que, sur K_δ, les f_n convergent uniformément vers f.

Démonstration Une fois pour toutes, nous nous restreindrons au compact K. Etant donné un nombre ε' fixé, appelons $A_{n,\varepsilon'}$ l'ensemble des points x de K pour lesquels $d(f_n(x), f(x)) > \varepsilon'$.

Montrons d'abord que cet ensemble $A_{n,\varepsilon'}$ est μ-mesurable. Comme chacune des fonctions f_n, f, est mesurable (théorème 23) l'application $x \longrightarrow (f_n(x), f(x))$, est une application mesurable de K dans le produit $F \times F$ (corollaire 1 du théorème 23 bis); comme alors $(x,y) \rightarrow d(x,y)$ est une application continue de $F \times F$ dans $\mathbb{R}$ (théorème 17 bis du chapitre II), il résulte du théorème 22 que la fonction réelle $x \longrightarrow d(f_n(x), f(x))$ est mesurable. Alors

l'ensemble $A_{n,\varepsilon'}$ n'est autre que l'intersection avec K de l'image réciproque, par cette application, de l'ensemble ouvert $]\varepsilon', +\infty[$ de $\mathbb{R}$ · c'est donc bien un ensemble mesurable, d'après la définition. Appelons alors $\mathfrak{A}_{m,\varepsilon'}$ la réunion $\bigcup_{n \geq m} A_{n,\varepsilon'}$; elle est mesurable. Un point x de K appartient à cette réunion, si et seulement si il existe un entier $n \geq m$, tel que l'on ait $d(f_n(x), f(x)) > \varepsilon'$. La suite des $\mathfrak{A}_{m,\varepsilon'}$ est une suite décroissante d'ensembles mesurables. Si un point x appartient à l'intersection $\bigcap_{m=0}^{\infty} \mathfrak{A}_{m,\varepsilon'}$ des $\mathfrak{A}_{m,\varepsilon'}$, sûrement la suite des $f_n(x)$ ne converge pas vers $f(x)$; et, comme les f_n convergent presque partout vers f, cette Intersection est nécessairement de mesure nulle. Alors, la suite décroissante des $\mathfrak{A}_{m,\varepsilon'}$ ayant pour intersection un ensemble de mesure nulle, il résulte du théorème 19 5°), que, si $\delta' > 0$ est donné, il existe un entier m tel que l'on ait $\mu(\mathfrak{A}_{m,\varepsilon'}) \leq \delta'$; appelons $\mathfrak{A}_{\delta',\varepsilon'}$ cet $\mathfrak{A}_{m,\varepsilon'}$.

* C'est ici que nous faisons intervenir le tait que nous nous sommes restreints au compact K : les $\mathfrak{A}_{m,\varepsilon'}$ sont de mesure finie, et le théorème est applicable.

Nous avons donc montré le résultat suivant :

Quels que soient les nombres δ' et $\varepsilon' > 0$, il existe **un entier** $m' = m'(\delta', \varepsilon')$, et un ensemble $\mathcal{O}_{\delta', \varepsilon'} \subset K$ de mesure $\leqslant \delta'$, tels que l'on ait $d(f_n(x), f(x)) \leqslant \varepsilon'$ dès que $n \geqslant m(\delta', \varepsilon')$ et que $x \in K - \mathcal{O}_{\delta', \varepsilon'}$.

Prenons alors successivement $\delta' = \frac{\delta}{2^{\nu+1}}$, $\varepsilon' = \frac{1}{2^\nu}$, $\nu = 1,2,\ldots$. Considérons la suite des entiers $m_\nu = m\left(\frac{\delta}{2^{\nu+1}}, \frac{1}{2^\nu}\right)$ et des ensembles $\mathcal{O}_\nu = \mathcal{O}_{\frac{\delta}{2^{\nu+1}}, \frac{1}{2^\nu}}$. Appelons $\mathcal{O}_\delta$ la réunion des $\mathcal{O}_\nu$; c'est un ensemble mesurable, contenu dans K, de mesure $\leqslant \frac{\delta}{2}$.

Alors, quel que soit 3 , on a, pour $n \geqslant m_\nu$, l'inégalité $d(f_n(x), f(x)) \leqslant \frac{1}{2^\nu}$, dès que x n'appartient pas à l'ensemble $\mathcal{O}_\delta$. Cela prouve bien que, sur le complémentaire K'_δ de cette partie $\mathcal{O}_\delta$ par rapport à K, la suite des f_n converge uniformément vers f. On a $\mu(K'_\delta) = \mu(K - \mathcal{O}_\delta)$ $\geqslant \mu(K) - \frac{\delta}{2}$; alors, d'après la définition de la mesure des ensembles, il existe un compact K_δ contenu dans K'_δ, tel que $\mu(K_\delta) \geqslant \mu(K) - \delta$, ou $\mu(K - K_\delta) \leqslant \delta$; et, sur K_δ, les f_n convergent uniformément vers f, ce qui démontre le théorème .

Remarques 1°/ On pourrait espérer qu'il soit possible de trouver une partie K_δ de K telle que $\mu(K - K_\delta) = 0$,et que, sur K_δ, les f_n convergent uniformément vers f, mais il n'en est rien.

Si, par exemple, on considère la suite des fonctions réelles définies sur $\mathbb{R}$ par la formule

$$(\mathrm{IV},4;1) \qquad f_n(x) = \begin{cases} 1 \text{ pour } 0 < x < \frac{1}{n} \\ 0 \text{ ailleurs} \end{cases}$$

on voit que, lorsque n tend vers $+\infty$, f_n converge simplement vers 0 ; et, sur le compact $[0,1]$, quel que soit ε, $0 < \varepsilon < 1$, et quel que soit n, l'ensemble des points x pour lesquels $|f_n(x)| > \varepsilon$ est nécessairement de mesure $\frac{1}{n} > 0$, ce qui rendrait inexacte la conjecture précédente.

2°/ Si X n'est pas compact, on ne peut pas remplacer K par X lui-même. En effet, si nous considérons des fonctions réelles de variable réelle, et si nous prenons pour f_n la fonction définie comme suit

$$(\text{IV},4;2) \qquad f_n(x) = \begin{cases} 0 \text{ pour } x < n \\ 1 \text{ pour } x \geq n, \end{cases}$$

on voit que la suite des f_n converge simplement vers la fonction 0 ; cependant, quel que soit ε, $0 < \varepsilon < 1$, et quel que soit n, l'ensemble des points x où l'on a l'inégalité $|f_n(x)| > \varepsilon$, est un ensemble de mesure infinie.

Corollaire (propriété de Lusin) - Soit f une application de X dans un espace métrisable séparable F. Si f est μ-mesurable, alors, quel que soit le compact K de X, et quel que soit $\delta > 0$, il existe un compact $K_\delta \subset K$, tel que $\mu(K - K_\delta) \leq \delta$, et que la restriction de f à K_δ soit une application continue de K_δ dans F ; et réciproquement.

Démonstration - 1°/ Soit d'abord f étagée; sur un nombre fini d'ensembles A_i $i = 1, 2, \ldots, N$ disjoints et de réunion X, elle prend des valeurs constantes f_i. Soit K un compact; pour tout i, soit K_i un compact $\subset A_i \cap K$, tel que $\mu(K_i) \geq \mu(A_i \cap K) - \frac{\delta}{N}$, il suffit de prendre pour K_δ la réunion des K_i. K_δ est compact, $\mu(K - K_\delta) \leq \sum_{i=1}^{N} \frac{\delta}{N} = \delta$; en outre K_δ est réunion des ensembles K_i disjoints et fermés; alors la restriction f_δ de f à K_δ est continue sur K_δ puisque l'image réciproque par f_δ de tout ensemble fermé (et même de tout ensemble) de F est une réunion de K_i, donc fermée, Soit maintenant f mesurable quelconque. Il existe, d'après le théorème 23 bis, une suite de fonctions étagées f_n qui, pour n infini, converge presque partout vers f. D'après le théorème d'Egoroff, si nous choisissons une métrique sur F, il existe un compact K'_δ, $\mu(K-K'_\delta) \leq \frac{\delta}{2}$, sur lequel les f_n convergent uniformément vers f. Alors, f_n étant étagée, il résulte de ce qu'on vient de voir qu'on peut trouver un compact $K_n \subset K'_\delta$, $\mu(K_n - K'_\delta) \leq \frac{\delta}{2^{n+2}}$, tel que f_n soit continue sur K_n. Prenons maintenant $K_\delta = \bigcap_{n=0}^{\infty} K_n$; c'est un compact de K et $K - K_\delta \subset (K-K'_\delta) \cup \left(\bigcup_{n=0}^{\infty}(K'_\delta - K_n)\right)$, donc $\mu(K-K_\delta) \leq \mu(K-K'_\delta) + \sum_{n=0}^{\infty} \mu(K'_\delta - K_n) \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta$ En outre, sur K_δ toutes les f_n sont continues, et convergent uniformément vers f, donc f est continue sur K_δ. Nous avons utilisé Egoroff, qui suppose f métrique (pour pouvoir parler de convergence uniforme); mais le résultat final (continuité de f sur K_δ) est évidemment indépendant de cette métrique.

2°/ Réciproquement, supposons que la propriété de Lusin soit vérifiée par f. Soit $\mathcal{O}$ un ouvert de F, et soit $B = f^{-1}(\mathcal{O})$. Comme f est continue sur K_δ, $B \cap K_\delta$ est un ouvert du compact K_δ, donc l'intersection d'un ouvert de X et de K_δ, donc μ-mesurable. Alors

$$\mu(B \cap K_\delta) \leq \mu_*(B\cap K) \leq \mu^*(B \cap K) \leq \mu(B \cap K_\delta)$$

$$+ \mu^*(B \cap (K-K_\delta))$$

$$\leq \mu(B\cap K_\delta) + \mu(K-K_\delta) \leq \mu(B \cap K_\delta) + \delta \text{ , donc}$$

$\mu^*(B \cap K) - \mu_*(B \cap K) \leq \delta$. Comme δ est arbitraire, on a $\mu^*(B \cap K) = \mu_*(B \cap K)$, donc $B \cap K$ est mesurable ; comme c'est vrai pour tout compact K, et que X est réunion dénombrable de compacts, $B = f^{-1}(\mathcal{O})$ est mesurable, et f est une application mesurable.

Remarque - La réciproque 2°/ est valable pour F topologique arbitraire (pourvu que X soit dénombrable à l'infini, comme toujours). Par contre le 1°/ utilise le théorème d'Egoroff, donc suppose F métrisable, et le théorème 23 bis, qui exige, comme le montrait sa démonstration, F métrisable séparable. Donc la propriété de Lusin est plus forte en général que la mesurabilité, mais lui est équivalente pour F métrisable séparable. Mais rappelons que nous n'avons donné page 467 qu'une définition provisoire de la mesurabilité, supposant F métrisable séparable. Nous donnons maintenant la définition définitive: $\vec{f}$ est mesurable, pour F topologique quelconque, si elle vérifie la propriété de Lusin. On peut alors montrer que les théorèmes 21, 22, les corollaires 1,2,3, du théorème 23 bis, sont toujours vrais, sans condition sur F ; les théorèmes 23, 23bis, sont vrais seulement pour F métrisable; le théorème 33 d'Egoroff suppose F métrique (pour parler de convergence uniforme); enfin, le dernier corollaire (Lusin) pourra s'exprimer en disant que, si $\vec{f}$ est mesurable, l'image réciproque de tout borélien de F est mesurable, tandis que la réciproque n'est vraie que si F est métrisable séparable. Nous avons pris la précaution, conformément à ce qui a été annoncé page 467, note *, tout en ne considérant que la définition provisoire, supposant F métrisable séparable, d'exprimer les énoncés avec les hypothèses qui correspondent à la définition définitive.

Théorème 34 - (Théorème de Lebesgue particulier). Soient X un espace localement compact muni d'une mesure de Radon $\mu \geq 0$, et $\vec{F}$ un espace de Banach. Soit $\vec{f}_n$ une suite de fonctions, définies sur X, à valeurs dans $\vec{F}$, ayant toutes leur support dans un même compact K, et toutes majorées en norme par un même nombre $M \geq 0$.

Si les $\vec{f}_n$ sont intégrables et mesurables (*) et convergent simplement presque partout vers $\vec{f}$ pour n tendant vers $+\infty$, alors $\vec{f}$ est intégrable et mesurable, les $\int \vec{f}_n$ convergent vers $\int \vec{f}$, et même $\int \|\vec{f} - \vec{f}_n\|$ converge vers 0.

* Les hypothèses sont excessives. Le théorème 39 1°, dira qu'une fonction bornée à support compact est intégrable, si et seulement si elle est mesurable. Mais nous ne le savons pas encore.

Démonstration - Soit donné $\varepsilon > 0$.

Il résulte du théorème précédent que, étant donné $\delta = \frac{\varepsilon}{4M}$, on peut trouver une partie mesurable K_δ de K, telle que $\mu(K - K_\delta) \leqslant \delta$, et que, sur K_δ, la suite des $\vec{f}_n$ converge uniformément vers $\vec{f}$. Alors il existe un entier p tel que, pour $n \geq p$, on ait, pour tout x de K_δ, l'inégalité

$$\|\vec{f}(x) - \vec{f}_n(x)\| \leqslant \frac{\varepsilon}{2\mu(K)} . \tag{IV,4;3}$$

Il en résulte que la fonction $\|\vec{f} - \vec{f}_n\|$ est majorée, pour $n \geqslant p$, par la fonction, prenant la valeur $\frac{\varepsilon}{2\mu(K)}$ sur l'ensemble K_δ, la valeur $2M$ sur l'ensemble $K - K_\delta$, et la valeur 0 sur le complémentaire de K. On a donc la majoration, pour n 3 p,

$$\int^* \|\vec{f} - \vec{f}_n\| \leqslant \frac{\varepsilon}{2\mu(K)} \mu(K) + 2M\delta = \varepsilon . \tag{IV,4;4}$$

Cette majoration montre que la suite des fonctions intégrables $\vec{f}_n$ est une suite d'approximation pour $\vec{f}$, que par conséquent $\vec{f}$ est intégrable, que $\int \vec{f}$ est la limite des $\int \vec{f}_n$, et que $\int \|\vec{f} - \vec{f}_n\|$ converge vers 0.

<u>Théorème 35 - (Théorème de Lebesgue général) - Soit X un espace localement compact, muni d'une mesure de Radon $\mu \geqslant 0$, et soit $\vec{F}$ un espace de Banach. Supposons d'une part que la suite des $\vec{f}_n$, intégrables et mesurables * converge simplement presque partout vers $\vec{f}$, d'autre part qu'il existe une fonction g réelle $\geqslant 0$, définie sur X, intégrable, telle que l'on ait l'inégalité $\|\vec{f}_n\| \leq g$. Alors la fonction $\vec{f}$ est intégrable et mesurable, et $\int \vec{f}_n$ converge vers $\int \vec{f}$, et même $\int \|\vec{f} - \vec{f}_n\|$ converge vers 0.</u>

Naturellement le théorème de Lebesgue général contient le théorème de Lebesgue particulier;, il suffit en effet de prendre pour g le produit de M par la fonction caractéristique du compact K. En outre le théorème 35 va être démontré indépendamment de 34, de sorte qu'on aurait pu donner 34 comme un simple corollaire de 35; nous avons préféré le démontrer avant, à cause de sa simplicité.

* Ici de même, la mesurabilité est superflue : le théorème 39, 2°, dira qu'une fonction intégrable est mesurable.

Démonstration - Donnons-nous un nombre $\varepsilon > 0$. Comme g est intégrable, il existe une fonction $\gamma \geq 0$, bornée en module par un nombre $M > 0$, à support "compact K, telle que l'on ait l'inégalité

$$\int^* |g - \gamma| \leq \frac{\varepsilon}{8}. \tag{IV,4;5}$$

D'après le théorème 33, si $\delta = \frac{\varepsilon}{8M}$, il existe une partie K_δ de K, telle que $\mu(K - K_\delta) \leq \delta$, et que, sur K_δ, $\vec{f}_n$ converge uniformément vers $\vec{f}$. Alors il existe un entier p tel que, pour $n \geq p$, on ait, sur K_δ, l'inégalité

$$\| \vec{f} - \vec{f}_n \| \leq \frac{\varepsilon}{2\mu(K)}. \tag{IV,4;6}$$

On a alors la majoration

$$\int^* \| \vec{f} - \vec{f}_n \| \leq \int_{K_\delta} \| \vec{f} - \vec{f}_n \| + \int^*_{\complement K_\delta} \| \vec{f} - \vec{f}_n \|. \tag{IV,4;7}$$

La première intégrale est celle d'une fonction majorée par $\frac{\varepsilon}{2\mu(K)}$, et par suite cette intégrale est $\leq \frac{\varepsilon}{2}$ pour $n \geq p$. Comme les $\| \vec{f}_n \|$ sont majorées par g, et que pour presque tous les x, $\vec{f}_n(x)$ tend vers $\vec{f}(x)$, $\| \vec{f} \|$ est presque partout majorée par g, donc la deuxième intégrale est majorée par

$$\int_{\complement K_\delta} 2g \leq \int_{\complement K_\delta} 2\gamma + \int^* 2|g - \gamma| \tag{IV,4;8}$$

$$= \int_{K - K_\delta} 2\gamma + \int^* 2|g - \gamma| \leq 2M\delta + \frac{2\varepsilon}{8} = \frac{\varepsilon}{2}.$$

Au total, nous avons l'inégalité

$$\int^* \| \vec{f} - \vec{f}_n \| \leq \varepsilon, \quad \text{pour} \quad n \geq p. \tag{IV,4;9}$$

Cette inégalité montre que la suite des $\vec{f}_n$ est une suite d'approximation de $\vec{f}$ par des fonctions intégrables; il en résulte bien que $\vec{f}$ est intégrable, que $\int \vec{f}_n$ converge vers $\int \vec{f}$, et que $\int \| \vec{f} - \vec{f}_n \|$ converge vers 0 .

Exemples d'applications du théorème de Lebesgue

1°) Considérons l'intégrale

$$\int_{[0,1]} x^n \, dx = \frac{1}{n+1} \; . \tag{IV,4;10}$$

Elle tend vers 0 lorsque n tend vers l'infini.

Le théorème de Lebesgue permettait de la prévoir; car x^n converge vers 0 pour n tendant vers l'infini, en tous les points du segment semi-ouvert $]0,1]$, donc presque partout sur l'intervalle d'intégration $[0,1]$; et par ailleurs les fonctions $|x^n|$ sont majorées par la constante 1. On est ici dans le cas d'application du théorème de Lebesgue particulier (Mais on voit bien que les x^n ne convergent pas uniformément vers leur limite).

2°) Considérons la suite des fonctions réelles définies, pour $n \geq 1$, par

$$f_n(x) = \begin{cases} n^\alpha & \text{pour} \quad 0 < x \leq \frac{1}{n} \\ 0 & \text{ailleurs} \end{cases} \tag{IV,4;11}$$

Elles convergent encore simplement vers 0 . L'intégrale de f_n est manifestement $n^{\alpha-1}$, et par conséquent elle tend vers 0 pour $\alpha < 1$.

Or si nous cherchons la plus petite majorante commune de toutes les fonctions f_n, nous trouvons qu elle est donnée par

$$g(x) = \begin{cases} \left[\frac{1}{x}\right]^\alpha & \text{pour } 0 < x \leq 1 \; {}^{*} \\ 0 & \text{ailleurs} \end{cases} \tag{IV,4;12}$$

Il en résulte que le théorème de Lebesgue est effectivement applicable pour $\alpha < 1$, parce qu'alors la fonction g est intégrable, comme majorée, sur l'intervalle $[0,1]$, par

* $[u]$ = partie entière de u .

la fonction intégrable $\frac{1}{x^\alpha}$ * . Au contraire, pour $\alpha \geq 1$,

le théorème de Lebesgue ne peut certainement pas s'appliquer, puisque nous avons une suite de fonctions convergeant simplement vers 0 et dont l'intégrale ne converge pas vers 0 ; et, en effet, lés fonctions ne sont pas majorées par une même fonction ≥ 0 intégrable, car $\left[\frac{1}{x}\right]^\alpha \geq \frac{1}{x} - 1$ n'est pas intégrable sur $[0,1]$. Cet exemple montre bien que les restrictions qui figurent dans l'énoncé du théorème de Lebesgue sont inévitables. Il y a cependant un cas où de telles restrictions n'existent pas, c'est celui où les f_n forment une suite croissante de fonctions réelles. (Voir théorème 36).

Remarque - Le théorème de Lebesgue est énoncé seulement pour une suite $\vec{f}_n$. Mais soit T un espace topologique, A une partie de T , a un point de T adhérent à A. Supposons que $\vec{f}_t$ soit une fonction intégrable sur X à valeurs dans $\vec{F}$, dépendant du paramètre $t \in A$. Supposons que, lorsque $t \in A$ tend vers a , $\vec{f}_t$ converge simplement presque partout vers $\vec{f}$, et qu'on ait une majoration $\|\vec{f}_t\| \leq g$, où g est fixe, ≥ 0 , intégrable. Peut-on en déduire que $\vec{f}$ soit intégrable, et que $\int \vec{f}_t \in \vec{F}$ converge vers $\int \vec{f}$, quand t tend vers a en restant dans A ? . Il en est bien ainsi, si T est métrisable. Choisissons en effet une métrique sur T . Tout d'abord, a, étant adhérent à A , est limite d'au moins une suite $a_1, a_2, \ldots, a_n, \ldots$ de points de A (théorème 15 du chapitre II); alors les $\vec{f}_{a_n}$ convergent simplement presque partout vers $\vec{f}$ pour n infini, et vérifient la majoration $\|\vec{f}_{a_n}\| \leq g$; le théorème de Lebesgue dit donc que $\vec{f}$ est intégrable. En outre, pour toute suite de ce type, $\int \vec{f}_{a_n}$ converge vers $\int \vec{f}$. Supposons que $\int \vec{f}_t$ ne converge pas vers $\int \vec{f}$ pour $t \in A$ convergeant vers a. Alors il existera un $\varepsilon > 0$ tel que, quel que soit $\eta > 0$, on puisse trouver un point t de A , vérifiant $d(t, a) \leq \eta$, et $\left\|\int \vec{f}_t - \int \vec{f}\right\| > \varepsilon$.

* Nous n'avons pas encore montré que $\frac{1}{x^\alpha}$ est intégrable sur $[0,1]$, si et seulement si $\alpha < 1$; mais il ne s'agit ici que d'un exemple. On Sa déjà vu plus ou moins en Mathématiques Spéciales !

Pour $\eta = 1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots$, on déterminera une suite a_1 , $a_2, \dots, a_n, \dots$, d'éléments de A , tels que $d(a_n, a) \leq \frac{1}{n}$, $\left\| \int \vec{f}_{a_n} - \int \vec{f} \right\| > \varepsilon$; c'est une contradiction, puisque $\int \vec{f}_{a_n}$ doit converger vers $\int \vec{f}$, et notre affirmation est démontrée.

Théorème 36 (Théorème de Fatou) - Soit $f_0, f_1, \dots, f_n, \dots$ une suite croissante de fonctions ≥ 0 (à valeurs finies ou infinies) sur X ; autrement dit, on a $f_0 \leq f_1 \leq f_2 \dots \leq f_n \leq \dots$.
Dans ces conditions, si f est la limite des f_n , on a toujours la relation :

(IV,4;13) $$\lim_{n \to \infty} \int^* f_n = \int^* f \leq +\infty \; .$$

Démonstration - Il est évident que le premier membre est majore par le deuxième, il nous faut démontrer une inégalité en sens inverse.

Nous démontrerons d'abord un lemme :

Lemme - Soit $f_0, f_1, f_2, \dots, f_n, \dots$ une suite croissante de fonctions ≥ 0 sur X ; soit $g_0, g_1, g_2, \dots, g_n, \dots$, une suite (non nécessairement croissante) de fonctions intégrables, telle que

(IV,4;14) $$\int g_n \leq \int^* f_n + \frac{\varepsilon}{2^{n+1}}$$

Si on appelle γ_n l'enveloppe supérieure de $g_0, g_1, \dots, g_n$, on a

(IV,4;15) $$\int \gamma_n \leq \int^* f_n + \varepsilon\left(1 - \frac{1}{2^{n+1}}\right) \; .$$

Cette Inégalité est évidemment vraie pour $n = 0$. Démontrons la par récurrence, en la supposant vraie pour $n - 1$, et en la démontrant pour n .

Si u et v sont deux nombres réels quelconques, on a l'égalité

(IV,4;16) $$u + v = \operatorname{Sup}(u, v) + \operatorname{Inf.}(u, v) \; .$$

Il en résulte que l'on peut écrire

(IV,4;17) $$\int \gamma_n = \int \operatorname{Sup.}(\gamma_{n-1}, g_n) = \int \gamma_{n-1} + \int g_n - \int \operatorname{Inf.}(\gamma_{n-1}, g_n) \; .$$

Examinons les termes du deuxième membre.

La première intégrale est majorée par $\int^* f_{n-1} + \varepsilon\left(1 - \frac{1}{2^n}\right)$, en vertu de l'hypothèse de récurrence.

La deuxième est majorée par $\int^* f_n + \frac{\varepsilon}{2^{n+1}}$, à cause de l'inégalité (IV,4;14).

D'autre part, comme g_{n-1} majore f_{n-1}, γ_{n-1} majore aussi f_{n-1} ; il en est de même de g_n, qui majore f_n et a fortiori f_{n-1} ; on a donc certainement l'inégalité $\mathrm{Inf}(\gamma_{n-1}, g_n) \geq f_{n-1}$, d'où

$$\int \mathrm{Inf.}(\gamma_{n-1}, g_n) \geq \int^* f_{n-1} \quad . \tag{IV,4;18}$$

Par conséquent finalement l'expression $\int \gamma_n$ admet la majoration

$$\int \gamma_n \leq \int^* f_{n-1} + \varepsilon\left(1 - \frac{1}{2^n}\right) + \int^* f_n + \frac{\varepsilon}{2^{n+1}} - \int^* f_{n-1} \leq \int^* f_n + \varepsilon\left(1 - \frac{1}{2^{n+1}}\right) \tag{IV,4;19}$$

et ainsi l'inégalité (IV,4;15) est démontrée par récurrence.

1°/ Supposons d'abord f bornée par $M \geq 0$, à support compact K. Donnons nous $\varepsilon > 0$. Pour chaque fonction f_n , on peut trouver une fonction $g_n \geq f_n$, étagée, bornée, à support compact, telle que l'on ait l'inégalité (IV,4;14).

On peut bien entendu supposer que g_n est elle-même majorée par M et de support dans le compact K ; sans quoi, en la remplaçant par 0 en tous les points x qui n'appartiennent pas à K , par M en tous les points x pour lesquels $g_n(x) > M$, on formerait une nouvelle fonction étagée qui vérifierait encore a fortiori (IV,4;14).

Appelons γ_n l'enveloppe supérieure des fonctions $g_0, g_1, \dots g_n$. Alors la suite des γ_n est une suite croissante de fonctions étagées ≥ 0 , elles ont donc une limite γ , qui est elle-même majorée par M et de support dans K. D'après le théorème de Lebesgue particulier (théorème 34), γ est intégrable, et $\int \gamma_n$ converge vers $\int \gamma$: D'autre part. d'après le lemme, on a (IV,4;15).

Comme alors nous avons vu que $\int \gamma_n$ converge vers $\int \gamma$, on a nécessairement l'inégalité

$$\int \gamma \leq \lim_{n\to\infty} \int^* f_n + \varepsilon \ , \tag{IV,4;20}$$

la limite écrite existant sûrement, puisqu'il s'agit d'une suite croissante de nombres ≥ 0 . Mais on a $\gamma \geq \gamma_n \geq f_n$, et comme ceci est vrai pour tout n , on a $\gamma \geq f$, et par suite on en déduit

(IV,4;21) $$\int^* f \leq \lim_{n \to \infty} \int^* f_n + \varepsilon .$$

Comme ceci est vrai quel que soit $\varepsilon > 0$, on a l'inégalité

(IV,4;22) $$\int^* f \leq \lim_{n \to \infty} \int^* f_n ;$$

c'est ce que nous voulions démontrer.

2°/ Supposons maintenant $f \geqslant 0$ quelconque. Appelons, comme toujours, $f_{M,K}$ la fonction égale à f en tous les points x appartenant à K , pour lesquels $f(x) \leq M$ égale à 0 en dehors de K, et à M aux points x de K pour'lesquels $8^{(x)} > M$.

Faisons de même pour les fonctions f_n . Alors la suite des fonctions $(f_n)_{M,K}$ est une suite croissante de fonctions $\geqslant 0$, et' elle converge vers $f_{M,K}$. D'après ce que nous venons de voir, on a donc

(IV,4;23) $$\int^* f_{M,K} = \lim_{n \to \infty} \int^* (f_n)_{M,K} \leq \lim_{n \to \infty} \int^* f_n .$$

En prenant alors la borne supérieure pour toutes les valeurs de M et de K , on en déduit bien de nouveau (IV,4;22), et le théorème de Fatou est démontre.

Corollaire o - Soit $A_0, A_1, A_2, \ldots, A_n, \ldots$ une suite croissante de parties de X , de réunion A . Alors

(IV,4;23 bis) $$\mu^*(A) = \lim_{n \to \infty} \mu^*(A_n) .$$

Cela résulte immédiatement du théorème, et de (IV,3;18).

Corollaire 1 - Soit f_n une suite de fonctions de signe quelconque sur X , à valeurs dans $\overline{\mathbb{R}}$, intégrables; supposons que cette suite soit monotone, c'est-à-dire ue l'on ait toujours $f_n \geq f_{n-1}$, ou toujours $_{n\;-\;1}$ y que cette suite converge vers f . Alors :

1°) Si $\int^* |f| < +\infty$, f est intégrable, et $\int f_n$ converge vers $\int f$.

2°) si $\int^* |f| = +\infty$, f n'est pas intégrable, et $\int f_n$ tend vers $\pm \infty$.

<u>Démonstration</u> - Démontrons 1°). Supposons, pour fixer les idées, que la suite des f_n soit croissante. Alors on a l'inégalité : $f_0 \leq f_n \leq f$, donc $|f_n| \leqslant |f_0| + |f|$, donc

(IV, 4;24) $$\left|\int f_n\right| \leqslant \int^* |f| + \int |f_0| ,$$

donc la suite des nombres réels $\int f_n$ est croissante et bornée; elle admet donc une limite finie. Alors, quel que soit $\varepsilon > 0$, il existe un entier p, tel que l'on ait, pour $m \geqslant n \geqslant p$, l'inégalité

(IV, 4;24 bis) $$\int f_m - \int f_n \leqslant \varepsilon \quad \text{ou} \quad \int (f_m - f_n) \leqslant \varepsilon .^*$$

Mais la suite des fonctions $f_m - f_n$, pour n fixé et m tendant vers l'infini, est une suite croissante de fonctions $\geqslant 0$. Nous pouvons donc appliquer le théorème de Fatou, et écrire

(IV,4;25) $$\int^* (f - f_n) \leqslant \varepsilon , \text{ pour } n \geqslant p .$$

Par ailleurs, $\int^* (f - f_n) = \int^* |f - f_n|$, puisqu'il s'agit d'une fonction $\geqslant 0$; on en déduit que la suite des f_n est une suite d'approximation de f par des fonctions intégrables, et ceci démontre notre affirmation.

Pour 2°), supposons au contraire $\int^* |f| = +\infty$. Si f était intégrable, il en serait de même de $|f|$, ce qui serait contradictoire avec l'égalité $\int^* |f| = +\infty$; donc f n'est pas intégrable.

D'autre part, en supposant, comme précédemment, la suite des f_n croissante, les $f_n - f_0$ forment une suite croissante de fonctions $\geqslant 0$, elles convergent vers $f - f_0$, le théorème de Fatou permet alors d'écrire

(IV, 4;26) $$\int^* (f - f_0) = \left(\lim_{n\to\infty} \int f_n\right) - \int f_0 ;$$

mais nous avons l'inégalité

(IV,4;27) $$|f| \leqslant |f_0| + |f - f_0|, \text{ donc } +\infty = \int^* |f| \leqslant \int |f_0| + \int^* |f - f_0| ,$$

donc $\int^* |f - f_0| = +\infty$, de sorte que (IV,4;26) démontre notre affirmation.

* $f_m - f_n$ n'est pas défini aux points où f_m et f_n sont toutes deux infinies; mais ces points forment un ensemble de mesure nulle, c'est sans importance.

Les résultats précédents faisaient des hypothèses sur $\int |f|$. On peut au contraire faire des hypothèses sur les quantités $\int f_n$; et alors on a les résultats équivalents suivants :

Corollaire 2 - Dans les conditions du corollaire 1 :

1°/ Si les quantités $\left|\int f_n\right|$ sont bornées, alors f est intégrable, et $\int f_n$ converge vers $\int f$;

2°/ Si les quantités $\left|\int f_n\right|$ ne sont pas bornées, alors f n'est pas intégrable, et $\int f_n$ converge vers $\pm\infty$.

Une conséquence du théorème de Fatou est le théorème de convexité dénombrable, dont nous avons déjà signalé l'importance.

Corollaire 3 -(Inégalité de convexité dénombrable).

soit $u_0, u_1, \dots u_n, \dots$ une suite de fonctions réelles ≥ 0 (a valeurs finies ou infinies) sur X ; on a toujours l'inégalité de convexité dénombrable

$$\int^*\left(\sum_{n=0}^{\infty} u_n\right) \leq \sum_{n=0}^{\infty}\int^* u_n \leq +\infty . \tag{IV,4;28}$$

Démonstration - On a en effet, pour une somme finie $\sum_{n=0}^{m}$, l'inégalité de convexité (théorème 25). Le second membre converge, pour m infini, vers le second membre de (IV,4;28), que celui-ci soit fini ou infini.

D'autre part, étant donné que la somme $u_0 + u_1 + \cdots + u_m$ est croissante avec m et converge, pour m tendant vers l'infini, vers la somme $\sum_{n=0}^{\infty} u_n$, le théorème de Fatou indique que le premier membre converge vers le premier membre de (IV,4;28), d'où le résultat.

Remarque - Il en résulte que, si $\sum_{n=0}^{\infty}\int^* u_n < +\infty$, on a $\int^*\left(\sum_{n=0}^{\infty} u_n\right) < +\infty$, donc, d'après le théorème 26, la série $\sum_{n=0}^{\infty} u_n$ est presque partout convergente.

Corollaire 4 - Interversion du signez et du signe $\int$ pour des séries à termes positifs.

Si les u_n forment une suite de fonctions ≥ 0 (à valeurs finies ou Infinies) sur X , intégrables, on a toujours l' alité

$$\int^{*}\left(\sum_{n=0}^{\infty} u_n\right) = \sum_{n=0}^{\infty}\int u_n \leqslant +\infty \; ; \qquad \text{(IV,4;29)}$$

si l'un auelconoue des deux membres a une valeur finie, alors la fonction $S = \sum_{n=0} u_n$ est lntégrable, et l'on a

$$\int \sum_{n=0}^{\infty} u_n = \sum_{n=0}^{\infty}\int u_n \; . \qquad \text{(IV,4;30)}$$

L'égalité $\int S_m = \sum_{n=0}^{m} \int u_n$ (où $S_m = \sum_{n=0}^{m} u_n$) est en effet vraie pour une somme finie; il suffit alors de passer à la limite pourmtendant vers l'infini,en tenant compte du théorème et des corollaires 1 et 2.

Nous voyons ainsi combien le théorème de Fatou est utile pour l'intégration des séries, lorsqu'il s'agit de séries de fonctions positives. Au contraire, le théorème 35 de Lebesgue, sous sa forme directe, est assez peu maniable pour des séries; car, si on suppose qu'il s'agisse de fonctions à valeursvectorielles $\vec{u}_n$ pour pouvoir écrire (IV,4;30), il faudrait savoir que le8 différences $\vec{S} - \vec{S}_n$, c'est-à-dire les restes $\vec{R}_n$, sont majorés en norme,indépendamment den , par une fonction $g \geq 0$ intégrable; or ceci, en général, n'est pas facile à voir.

On a cependant le très important théorème suivant :

Théorème 37 - Soit $\vec{u}_n$ une suite de fonctions sur X à valeurs dans l'espace de Banach $\vec{F}$, chacune définie μ - presque partout; alors, si la série $\sum_{n=0}^{\infty} \int^{*} \|\vec{u}_n\|$ est convergente, la série $\sum_{n=0}^{\infty} \vec{u}_n(x)$ converge absolument, pour /&-presque toutes les valeurs de x ; elle définit donc p-presque partout un vecteur $\vec{S}(x)$. Si les $\vec{u}_n$ sont intégrables, la fonction $\vec{S}$, ainsi définie p-presque partout sur X , à valeurs dans $\vec{F}$, est μ - intégrable, et l'on a l'égalité :

$$(IV,4;31) \qquad \int \vec{S} = \sum_{n=0}^{\infty} \int \vec{u}_n \ , \quad \text{ou} \quad \int \sum_{n=0}^{\infty} \vec{u}_n = \sum_{n=0}^{\infty} \int \vec{u}_n \ ,$$

et en outre :

$$(IV,4;32) \qquad \left\| \int \vec{S} \right\| \leq \int \| \vec{S} \| \leq \sum_{n=0}^{\infty} \int \| \vec{u}_n \| \ ,$$

, et

$$(IV,4;33) \qquad \lim_{m \to \infty} \int \| \vec{R}_m \| = 0 \ ,$$

où $\vec{R}_m$ est le reste $\sum_{n=m+1}^{\infty} \vec{u}_n$.

Démonstration - On peut d'abord prolonger les $\vec{u}_n$ là où elles ne sont pas définies, de manière à pouvoir les supposer partout définies.

Si nous supposons la série $\sum_{n=0}^{\infty} \int^* \| \vec{u}_n \|$ convergente, et si nous appelons f la fonction $\geqslant 0$ définie par

$$(IV,4;33bis) \qquad f(x) = \sum_{n=0}^{\infty} \| \vec{u}_n (x) \| \leq +\infty \ ,$$

on a, d'après (IV,4;29), $\int^* f < +\infty$.

On déduit alors du théorème 26 que la fonction f est presque partout finie; mais, en un pointx où $f(x)$ est finie, la série $\sum_{n=0}^{\infty} \vec{u}_n(x)$ est absolument convergente dans $\vec{F}$, et par conséquent a fortiori convergente d'après le théorème 55 du chapitre II. Si nous appelons $\vec{S}(x)$ la somme, on voit donc que? est bien un. fonction définie presque partout sur X à valeurs dans F . Si alors nous posons $\vec{S}_m = \sum_{n=0}^{m} \vec{u}_n$ et $\vec{R}_m = \vec{S} - \vec{S}_m$, on a pour $\vec{R}_m$ la majoration

$$(IV,4;34) \qquad \int^* \| \vec{R}_m \| \leqslant \int^* \Big(\sum_{n=m+1}^{\infty} \| \vec{u}_n \| \Big) = \sum_{n=m+1}^{\infty} \int^* \| \vec{u}_n \| \ ;$$

il en résulte que $\int^* \| \vec{R}_m \|$ converge vers 0 pour m tendant vers l'infini, comme reste de la série $\sum \int^* \| \vec{u}_n \|$, à termes $\geqslant 0$ convergente; donc, si les $\vec{u}_n$ sont intégrables, la suite des $\vec{S}_m$ est une suite d'approximation de $\vec{S}$ par des fonctions intégrables; par suite,? est intégrable, et $\int \vec{S}_m$ converge vers $\int \vec{S}$, d'où (IV,4;31). Nous avons même démontré plus, puisque nous savons que $\int \|\vec{S}-\vec{S}_m\| = \int \|\vec{R}_m\|$ tend vers 0 pour m infini, c'est-a-dire (IV,4;33); (IV,4;32) est (IV,4;34) appliquée à $\vec{S}$ au lieu de $\vec{R}_m$.

Théorème 38 - Soit $\vec{f}$ une fonction définie presque partout sur X, à valeurs dans $\vec{F}$, et soit $\vec{f}_n$ une suite d'approximation de $\vec{f}$, autrement dit une suite de fonctions telles que $\int^* \|\vec{f} - \vec{f}_n\|$ converge vers 0, pour n tendant vers l'infini. Alors, de la suite des $\vec{f}_n$, on peut extraire une suite partielle qui converge presque partout vers $\vec{f}$ *.

On peut en effet, de procne en proche, déterminer une suite strictement croissante d'entiers $p_0, p_1, \cdots, p_n, \cdots$ telle que

$$\int^* \|\vec{f} - \vec{f}_{p_n}\| \leqslant \frac{1}{2^n} . \tag{IV,4;35}$$

Alors on a la majoration

$$\int^* \left\|\vec{f}_{p_{n+1}} - \vec{f}_{p_n}\right\| \leqslant \frac{1}{2^{n-1}} , \tag{IV,4;36}$$

et par conséquent, si nous posons $\vec{u}_n = \vec{f}_{p_{n+1}} - \vec{f}_{p_n}$, on a la ma joratlon $\sum\limits_{n=0}^{\infty} \int^* \|\vec{u}_n\| < +\infty$.

On déduit alors du théorème 37 que la série $\sum\limits_{n=0}^{\infty} \vec{u}_n$ est presque partout convergente, autrement dit que la suite partielle des $\vec{f}_{p_n}$ converge presque partout vers une fonction limite $\vec{g}$; en outre, il résulte de ce que nous avons vu au théorème 37, que $\int^* \|\vec{f}_{p_n} - \vec{g}\|$ converge vers 0 , et comme aussi $\int^* \|\vec{f}_{p_n} - \vec{f}\|$ converge vers 0 , on en déduit que $\int \|\vec{f} - \vec{g}\|$ est nulle, par conséquent $\vec{f}$ et $\vec{g}$ sont presque partout égales,ce qui montre bien que la suite des $\vec{f}_{p_n}$ converge presque partout vers $\vec{f}$.

* Les $\vec{f}_n$ et $\vec{f}$ ne sont pas supposées intégrables, mals ce sera le cas le plus intéressant. Ce théorème (qu'on aurait, en fait, pu démontrer bien avant),donne une-idée intéressante du comportement des suite d'approximation.

Remarque Ce **résultat** est très remarquable, il montre que l'on peut extraire de la suite ou de chacune de ses suites partielles une sous suite partielle convergeant presque partout vers $\vec{f}$. Mais on aurait tort de croire que la suite
Z **elle-même** des f_n converge nécessairement presque partout vers f

Considérons en effet l'exemple suivant.

Prenons pour X le cercle trigonométrique, et pour mesure μ la mesure $d\theta$. Considérons la suite des points a_n définie par $a_n = 1 + \frac{1}{2} + \cdots + \frac{}{n}$, $n \geqslant 1$, et appelons f_n la fonction égale à 1 dans l'intervalle $[a_n , a_{n+1}]$, et à 0 ailleurs. On a $\int f_n = \frac{1}{n+1}$, et par conséquent $\int |f_n|$ converge vers 0 pour n tendant vers l'infini; les f_n forment une suite d'approximation de la fo ction 0 . Cependant il est évident que la suite des f_n ne converge pour aucune valeur de x vers 0 ; en effet, comme la série $\sum_{n=1}^{\infty} \frac{1}{n}$ est divergente, les points a_n font une infinité de fois le tour de la circonférence, et par suite tout point x est, pour une infinité de valeurs de n , contenu dans l'intervalle $[a_n , a_{n+1}]$, et par conséquent on a, pour une infinité de valeurs de n, $f_n(x) = 1$.

Ce que nous dit par contre le théorème, c'est qu'on peut extraire de la suite des f_n une suite partielle convergeant presque partout vers 0 ; si en effet nous déterminons la suite des entiers p_n de façon que l'intervalle $[a_{p_n}, a_{p_n+1}]$ contienne un point fixe c , comme la longueur de cet intervalle tend vers 0 , on voit que la suite des f_{p_n} converge vers 0 en tout point $\neq c$, c'est-à-dire effectivement partout sauf en un point, et par conséquent presque partout.

Caractérisation des fonctions intégrables. Intégrabilité et mesurabilité

Théorème 39 - Soit $\vec{f}$ une application de l'espace localement compact X muni d'une mesure de Radon $\mu \geq 0$, dans l'espace de Banach $\vec{F}$.

1°/ Si $\vec{f}$ est bornée à support compact, pour qu'elle soit intégrable, il faut et il suffit qu'elle soit mesurable;

2°/ Si $\vec{f}$ est quelconque, pour qu elle soit intégrable, il faut et il suffit qu'elle soit mesurable, et que l'intégrale supérieure $\int^* \|\vec{f}\|$ soit finie.

Démonstration - Remarquons d'abord que la première partie du théorème est un cas particulier de la deuxième.

Supposons d'abord $\vec{f}$ intégrable. Alors, d'après le théorème 29, $\|\vec{f}\|$ est aussi intégrable, et par conséquent, d'après le corollaire 4 du théorème 29, $\int^* \|\vec{f}\| = \int \|\vec{f}\|$ est nécessairement finie. D'autre part, si $\vec{f}_n$ est une suite d'approximation de $\vec{f}$ par des fonctions étagées, il résulte du théorème 38 qu'on peut extraire de la suite des $\vec{f}_n$ une suite partielle qui converge presque partout vers $\vec{f}$; $\vec{f}$ est donc presque partout la limite d'une suite de fonctions mesurables, et par suite, d'après le théorème 23, elle est bien mesurable.

Démontrons maintenant la réciproque.

Supposons que $\vec{f}$ soit mesurable et que $\int^* \|\vec{f}\|$ soit finie.

1°/ Supposons d'abord $\vec{f}$ bornée en norme par $M \geq 0$ et de support compact K.

On sait alors que $\vec{f}$ est limite presque partout d'une suite de fonctions étages $\vec{f}_n$ (théorème 23 bis); en remplaçant éventuellement $\vec{f}_n$ par $\vec{f}_n \dfrac{M}{\|\vec{f}_n\|}$ sur les étages où la norme est $> M$, ce qui ne change pas la convergence presque partout vers $\vec{f}$, on voit qu'on peut supposer que les $\|\vec{f}_n\|$ sont bornées dans leur ensemble par M ; en outre, on peut les supposer à support dans K, sans quoi on les multiplie par la fonction caractéristique de K. Il résulte alors du théorème de Lebesgue particulier 34 que $\vec{f}$ est intégrable; en outre les $\vec{f}_n$ en forment une suite d'approximation.

2°/ Passons maintenant au cas général où f est quelconque.

Soit $K_1, K_2, \ldots, K_n, \ldots$, une suite croissante de compacts, de réunion X.

Appelons alors $\vec{f}_n$ la fonction égale à $\vec{f}$ sur l'ensemble H_n des points $x \in K_n$ où $\|\vec{f}(x)\| \leq n$ et à $\vec{0}$ ailleurs.

H_n est l'intersection de K_n et de l'image réciproque, par la fonction mesurable $\|\vec{f}\|$ (composée de $\vec{f}$ et de la norme, théorème 22), de l'intervalle $[0, +n]$; donc c'est une Intersection d'ensembles mesurables et par suite H_n est mesurable. Alors $\vec{f}_n$ n'est autre que le produit $\vec{f}\varphi_{H_n}$, et par conséquent elle est aussi mesurable (corollaire 3 du théorème 23 bis)*; étant bornée et à support compact, elle est aussi intégrable. Alors nous savons aussi que la fonction $\|\vec{f}_n\|$ est intégrable; mais ces fonctions forment une suite croissante de fonctions ≥ 0, et leur limite est précisément $\|\vec{f}\|$; alors, en vertu de l'hypothèse $\int^* \|\vec{f}\| < +\infty$, il résulte du corollaire 1 du théorème de Fatou, que $\|\vec{f}\|$ est intégrable. Mais alors les $\vec{f}_n$ convergent vers $\vec{f}$ pour n tendant vers l'infini, et sont majorées en norme par la fonction $\|\vec{f}\| \geq 0$, intégrable; d'après le théorème 35 de Lebesgue, $\vec{f}$ est intégrable.

Corollaire 1 - Toute fonction $\vec{f}$ mesurable, bornée, nulle en dehors d'un ensemble Y de mesure extérieure finie, est intégrable.

Car
$$\int^* \|\vec{f}\| \leq \|\|\vec{f}\|\| \mu^*(Y) < +\infty .$$

Corollaire 2 - Si $\|\mu\| < +\infty$, toute fonction $\vec{f}$, μ -mesurable et bornée, et en particulier toute fonction $\vec{f}$, continue et bornée, est intégrable.

En effet,
$$\int^* \|\vec{f}\| \leq \|\|\vec{f}\|\| \|\mu\|$$

Corollaire 3 - Pour que $\vec{f}$ soit intégrable, il faut et il suffit qu'elle soit mesurable, et que les intégrales supérieures des fonctions $\|\vec{f}\|_{M,K}$, définies page 474 , soient bornées indépendamment de M et de K ; il suffit même qu'il en soit ainsi pour une suite particulière de nombres M_n tendant vers l'infini, et une suite croissante particulière de compacts K_n , de réunion X .

Cela résulte en effet de ce que $\int^Y \|\vec{f}\|$ est alors la limite des quantités $\int^* \|\vec{f}\|_{M_n, K_n}$.

Corollaire 4 - (Convexité et valeur moyenne) - Soit $\vec{C}$ un ensemble convexe fermé de $\vec{F}$. Si $\vec{f}$ est intégrable sur un ensemble mesurable A de mesure finie $\neq 0$, et si $\vec{f}(A) \subset \vec{C}$ sa moyenne sur A : $\frac{1}{\mu(A)} \int_A \vec{f}\, d\mu$ appartient à $\vec{C}$.

Démonstration - Supposons d'abord une fois pour toutes que $\vec{0} \in \vec{C}$. On prolongera $\vec{f}$ par $\vec{0}$ dans $\complement A$, de sorte qu'on aura maintenant $\overrightarrow{f(X)} \subset \vec{C}$.

Soit d'abord $\vec{f}$ étagée ; on l'écrira $\vec{f} = \sum_{i=1}^{N} \vec{g}_i \chi_{A_i}$, $\vec{g}_i \in \vec{C}$, χ_{A_i} fonction caractéristique de A_i ; contrairement à ce qu'on fait dans d'autre cas, nous écrirons aussi le terme $\vec{0}\chi_A$ si $\vec{f}$ prend la valeur $\vec{0}$, de façon à pouvoir supposer les A_i disjoints et $\bigcup_{i=1}^{N} A_i = X$. Alors $\frac{1}{\mu(A)} \int_A \vec{f}\, d\mu = \frac{1}{\mu(A)} \sum_{i=1}^{N} \vec{g}_i \mu(A_i \cap A)$, qui est dans $\vec{C}$ d'après la propriété barycentrique des convexes, puisque $\sum_{i=1} \mu(A_i \cap A) = \mu(A)$. Supposons maintenant $\vec{f}$ bornée à support compact. La partie 1°/ de la démonstration du théorème 39 nous a fourni, en utilisant le théorème 23 bis et sa remarque 2°/, une suite d'approximation de $\vec{f}$ formée de fonctions étagées $\vec{f}_n$, à valeurs dans $\overrightarrow{f(X)} \subset \vec{C}$. Alors

$\frac{1}{\mu(A)}\int_A f_n \, d\mu$ est dans $\vec{C}$ d'après ce qui précède, donc aussi $\frac{1}{\mu(A)}\int_A \vec{f} \, d\mu$ puisque $\vec{C}$ est fermé.

Soit enfin $\vec{f}$ quelconque. La partie 2°/ de la démonstration du théorème 39 nous a donné une suite d'approximation de $\vec{f}$ formée de fonctions $\vec{f}_n$ bornées à support compact, prenant leurs valeurs dans $\vec{C} \cup \{\vec{0}\} = \vec{C}$ Alors un nouveau passage à la limite donne le résultat. La démonstration est donc terminée si $\vec{0} \in \vec{C}$. Mais, dans le cas général, soit $\vec{c} \in \vec{C}$; alors le translaté $\vec{C} - \vec{c}$ est convexe fermé et contient $\vec{0}$; la fonction $\vec{f} - \vec{c}$ prend ses valeurs dans ce convexe; on a donc

$$\frac{1}{\mu(A)}\int_A \vec{f} \, d\mu = \vec{c} + \frac{1}{\mu(A)}\int_A (\vec{f} - \vec{c}) \, d\mu$$

$$\in \vec{c} + (\vec{C} - \vec{c}) = \vec{C} \text{ , c.q.f.d}$$

<u>Remarque</u> - On peut supprimer la restriction $\mu(A) \neq 0$ en écrivant

$$\int_A \vec{f} \, d\mu \in \mu(A)\, \vec{C} \text{ ,}$$

<u>Remarques</u> 1°) Comme nous l'avons dit page 492 les fonctions usuelles sont généralement mesurables. Alors l'intégrabilité de $\vec{f}$ est équivalente l'inégalité $\int^* \|\vec{f}\| < \infty$.

2°) Soit f une fonction <u>mesurable ≥ 0 sur X</u>. si $\int^* f$ est fini, elle est intégrable, et alors (corollaire 4 du théorème 29), on a $\int^* f = \int f$. Si $\int^* f = +\infty$, $\int f$ n'a pas de sens. Mais il est maintenant raisonnable, <u>lorsque f est mesurable</u> ≥ 0 et que $\int^* f = +\infty$, d'écrire encore $\int f = \int^* f = +\infty$ (mais naturellement on ne dira pas que f est intégrable !) Alors, <u>si f est mesurable ≥ 0,</u> <u>on a dans tous les cas</u>, $\int^* f = \int f$. D'ailleurs, pour une partie A <u>mesurable</u> de X, nous avions déjà écrit $\mu(A) = +\infty$ si $\mu^*(A) = +\infty$.

Théorie de l'intégration à partir des fonctions continues et semi-continues inférieurement

Dans ce qui précède, nous avons fait jouer un rôle essentiel, pour la théorie du prolongement de Lebesgue, à la mesure des ensembles et aux fonctions étagées. On aurait pu les éviter complètement, en utilisant les fonctions semi-continues inférieurement à valeurs dans $\overline{\mathbb{R}}$ (définition page 84).

Tout d'abord on voit immédiatement que l'enveloppe supérieure ou inférieure d'un nombre fini de fonctions semi-continues inférieurement est encore semi-continue inférieurement; et que l'enveloppe supérieure (mais non l'enveloppe inférieure) d'une famille quelconque, finie ou non, de fonctions semi-continues inférieurement est semi-continue inférieurement. En particulier l'enveloppe supérieure d'une famille de fonctions continues est semi-continue inférieurement (les fonctions semi-continues supérieurement ont des propriétés analogues; mais cette fois,- c'est l'enveloppe inférieure d'une famille quelconque de fonctions semi-continues supérieurement (en particulier de fonctions continues) qui est semi-continue supérieurement; d'ailleurs f est semi-continue supérieurement, si et seulement si $-f$ est semi-continue inférieurement). La fonction caractéristique d'une partie **A** de **X** est semi-continue inférieurement (resp. supérieurement) si et seulement si cette partie est ouverte (resp. fermée); les propriétés des ouverts et des fermés relativement aux réunions ou intersections finies ou quelconques sont l'expression des propriétés des fonctions semi-continues inférieurement ou supérieurement relativement aux: enveloppes supérieures ou inférieures de familles finies ou quelconques. La somme d'un nombre fini de fonctions semi-continues supérieurement ou inférieurement $\geqslant 0$, ou à valeurs dans $\mathbb{R}$ (mais pas à valeurs dans $\overline{\mathbb{R}}$, car elle pouvait être indéterminée, $+\infty-\infty$) a la même propriété; la somme d'une famille infinie de fonctions semi-continues inférieurement (en particulier continues) $\geqslant 0$ est semi-continue inférieurement, car c'est l'enveloppe supérieure des sommes partielles finies.

Une fonction semi-continue inférieurement (ou supérieurement) f est borélienne, donc mesurable pour toute mesure de Radon $\geqslant 0$. En effet, $f^{-1}(]a,+\infty])$ est ouvert, quel que soit a ; donc, par différence, $f^{-1}(]a,b])$ est borélien quels que soient a et b, donc aussi $f^{-1}(]a,b[) = \bigcup_{n=1,2,\dots} f^{-1}\left(\left]a, b-\frac{1}{n}\right]\right)$; mais tout ouvert $\mathcal{O}$ de $\mathbb{R}$ est réunion d'une infinité dénombrable d'intervalles ouverts disjoints, donc $f^{-1}(\mathcal{O})$ est bien borélien, et f borélienne,

Théorème 39 bis - L'intégrale supérieure d'une fonction semi-continue inférieurement $f \geqslant 0$ (à valeurs finies ou non) est la borne supérieure des intégrales des fonctions $\varphi \geqslant 0$ de $\mathcal{C}(X)$ qui la minorent.

Démonstration - Soit f à valeurs dans $\overline{\mathbb{R}}_+$, semi-continue inférieurement. Pour toute $\varphi \in \mathcal{C}(X)$, $0 \leqslant \varphi \leqslant f$, on a $\mu(\varphi) = \int \varphi \leqslant \int^* f$. Nous devons montrer que, quel que soit $A < \int^* f$, il existe $\varphi \in \mathcal{C}(X)$, telle que $0 \leqslant \varphi \leqslant f$, et $\mu(\varphi) \geqslant A$.

Tout d'abord, d'après la définition de l'intégrale supérieure, si $A < B < \int^* f$, il existe un nombre $M \geqslant 0$ et un compact K de X tels que * $\int^* f_{M,K} \geqslant B$; au lieu de celà introduisons un voisinage ouvert $\mathcal{V}$ de K, sa fonction caractéristique $\chi_{\mathcal{V}}$, et la fonction $g = \operatorname{Inf}(f, M\chi_{\mathcal{V}})$; on a $f_{M,K} \leqslant g \leqslant f$, donc $\int^* g \geqslant B$, et g est encore semi continue inférieurement, comme enveloppe inférieure de deux telles fonctions. Mais maintenant est bornée et à support compact. Soit $\chi_{k,n}$ la fonction caractéristique de l'ensemble $\mathcal{O}_{k,n} = g^{-1}(]\frac{kM}{n}, M])$; cet ensemble est ouvert puisque g est semi-continue inférieurement. Posons alors $g_n = \sum_{k=1}^{n-1} \frac{M}{n} \chi_{k,n}$; c'est la fonction étagée prenant les valeurs $\frac{kM}{n}$ dans les ensembles $g^{-1}(]\frac{kM}{n}, \frac{(k+1)M}{n}])$, et 0 ailleurs; on a donc $g_n \leqslant g$, et $g - g_n \leqslant \frac{M}{n}$; donc, pour n infini, les g_n convergent uniformément vers g, tout en ayant leur support dans le compact fixe $\mathcal{V}$; donc $\int g_n$ converge vers $\int g$. Donc il existe un entier n tel que $\int g_n > A$. L'entier n étant ainsi fixé, posons $\delta = \int g_n - A > 0$. Soit, pour tout k, une fonction continue $\varphi_k \geqslant 0$ à support compact dans $\mathcal{O}_{k,n}$, telle que $\mu(\varphi_k) \geqslant \mu(\mathcal{O}_{k,n}) - \frac{\delta}{M}$. La fonction $\varphi = \sum_{k=1}^{n-1} \frac{M}{n} \varphi_k$ est $\geqslant 0$, à support compact, majorée par g_n donc par g donc par f, et

$$\mu(\varphi) \geqslant \sum_{k=1}^{n-1} \frac{M}{n} \mu(\mathcal{O}_{k,n}) - \sum_{k=1}^{n-1} \frac{M}{n} \frac{\delta}{M}$$
$$= \int g_n - \frac{n-1}{n}\delta > A$$

* Voir la définition de $f_{M,K}$ page 474.

Remarque - Ce résultat est trés particulier aux fonctions semi-continues inférieurement ≥ 0. Si l'on prend par exemple la fonction caractéristique f de l'ensemble des irrationnels sur $\mathbb{R}$ son intégrale supérieure est $+\infty$, mais, pour toute $\varphi \in \mathcal{C}(X)$ telle que $0 \leq \varphi \leq f$, on a $\mu(\varphi) = 0$.

Pour une fonction ≥ 0 arbitraire, on a ce qui suit :

Théorème 39 ter - L'intégrale supérieure d'une fonction $f \geq 0$ est la borne inférieure des intégrales supérieures des fonctions semi-continues inférieurement qui la majorent.

Démonstration - Soit $\varepsilon > 0$. Soit d'abord f bornée à support compact. Soit g étagée à support compact, $g \geq f$, telle que $\int g \leq \int^* f + \frac{\varepsilon}{2}$. La fonction g est de la forme $\sum_{i=1}^{N} g_i \chi_{A_i}$, g_i constantes > 0, A_i ensembles mesurables de mesures finies; soit $M = \operatorname{Max}_{i=1}^{N} |g_i|$. Pour tout i, soit $\mathcal{O}_i$ un ouvert $\supset A_i$ tel que $\mu(\mathcal{O}_i) \leq \mu(A_i) + \frac{\varepsilon}{2NM}$; posons $h = \sum_{i=1}^{N} g_i \chi_{\mathcal{O}_i}$. Alors chaque $\chi_{\mathcal{O}_i}$ est semi-continue inférieurement puisque $\mathcal{O}_i$ est ouvert, donc h est encore semi-continue inférieurement; on a $f \leq g \leq h$, et $\int h \leq \sum_{i=1}^{N} g_i (\mu(A_i) + \frac{\varepsilon}{2NM}) \leq \int g + \frac{\varepsilon}{2} \leq \int^* f + \varepsilon$, ce qui démontre le théoréme dans ce cas.

Soit maintenant f quelconque. Considérons la fonction f_{n,K_n} (définition des $f_{M,K}$ page 474; les K_n sont une suite croissante de compacts de réunion X). Elle est bornée à support compact; donc il existe g_n semi-continue inférieurement, telle que $f_{n,K_n} \leq g_n$, $\int g_n \leq \int^* f_{n,K_n} + \frac{\varepsilon}{2^{n+1}}$. Si alors nous posons $\gamma_n = \operatorname{Sup}(g_0, g_1, \ldots, g_n)$, il résulte du lemme utilisé dans la démonstration du théorème 36 de Fatou que l'on a $\int \gamma_n \leq \int^* f_{n,K_n} + \varepsilon\left(1 - \frac{1}{2^{n+1}}\right)$. La suite des γ_n est croissante; toutes sont semi-continues inférieurement; leur limite γ pour n infini, qui est encore leur enveloppe supérieure, est encore semi-continue inférieurement; on a $f_{n,K_n} \leq \gamma_n \leq \gamma$ pour tout n, donc $f \leq \gamma$. D'autre part, d'après le théorème de Fatou

$$\int^* \gamma = \lim_{n\to\infty} \int^* \gamma_n \leq \lim_{n\to\infty} \int^* f_{n,K_n} + \varepsilon = \int^* f + \varepsilon ,$$

ce qui démontre le théorème dans le cas général.

Remarque- 1°/ On ne pourrait évidemment pas remplacer les fonctions semi-continues inférieurement par des fonctions continues. Par exemple si f est la fonction caractéristique des rationnels sur $\mathbb{R}$, son intégrale supérieure est nulle, mais toute fonction continue qui la majore est ≥ 1 et a une intégrale infinie.

2°/ Comme nous l'avions annoncé page 474, note (**), nous obtenons ici l'intégrale supérieure d'une fonction $\geq$ CI par 2 passages à la limite, à partir des $\mu(\varphi)$:

(IV,4;40 bis) $$\int^* f = \underset{(g \geq f,\, g \text{ semi-cont.inf})}{\mathrm{Inf}} \left(\underset{\substack{0\leq\varphi\leq g \\ \varphi\in\mathcal{C}(X)}}{\mathrm{Sup}} \mu(\varphi) \right)$$

Il semble bien que l'introduction de l'intégrale supérieure par cette formule prise comme définition soit la meilleure et la plus courte. Mais cette nouvelle méthode est plus abstraite. Ensuite, on pourra directement définir les fonctions scalaires intégrables par le théorème 32, c'est-a-dire en utilisant une suite d'approximation par des fonctions de $\mathcal{C}(X)$ $\int\varphi$ étant alors, par définition, $\mu(\varphi)$ (ce qui supprime le théorème idiot 30). On définira alors la mesure des ensembles après l'intégrale des fonctions, comme l'intégrale de leur fonction caractéristique. Bien entendu, les théorèmes fondamentaux de la théorie seront les mêmes; et certains auront une démonstration plus difficile que dans l'exposition adoptée antérieurement,

Pour définir l'intégrale des fonctions vectorielles, on introduira l'espace $\mathcal{C}(X)\otimes\vec{F}$ des fonctions continues à support compact décomposables, c'est-a-dire de la forme $\vec{g} = \sum_{i=1}^{N} \vec{g}_i\, \varphi_i$, $\vec{g}_i$ constante de $\vec{F}$, $\varphi_i \in \mathcal{C}(X)$

Par définition, on posera $\int \vec{g} = \sum_{i=1}^{N} \vec{g}_i\, \mu(\varphi_i)$ (ce qui coïncide bien avec son intégrale dans la théorie adoptée antérieurement) on vérifie que le résultat est bien indépendant de l'expression particulière d'une fonction $\vec{g}$ continue décomposable sous la forme précédente (comme nous avons déja procédé en utilisant l'intégrale des fonctions étagées). On retrouve bien les mêmes fonctions intégrables qu'antérieurement et la même intégrale, En effet, toute fonction intégrable dans cette nouvelle théorie admet une suite d'approximation formée de fonctions intégrables dans l'ancienne (les fonctions de $\mathscr{C}(X)\otimes\vec{F}$), avec la même intégrale; et réciproquement, d'après le théorème 32.

Pour les fonctions semi-continues inférieurement ≥ 0, il existe une généralisation du théorème de Fatou, valable pour des familles non dénombrables :

Théorème 39 quarto - Soit $(f_i)_{i\in I}$ une famille quelconque de fonctions semi-continues inférieurement ≥ 0, et soit f leur enveloppe supérieure. On suppose en outre que la famille est "croissante", au sens suivant : quels que soient $i, j \in I$, il existe $k \in I$ tel que $f_k \geq \operatorname{Sup}(f_i, f_j)$. Alors on a, $\int^* f = \operatorname{Sup}_{i\in I} \int^* f_i$.

Démonstration - Il est bien évident que $\int^* f \geq \operatorname{Sup}_{i\in I} \int^* f_i$. Soit alors A un nombre quelconque $< \int^* f$. Il existe une fonction $\varphi \in \mathscr{C}(X)$, $0 \leq \varphi \leq f$, telle que $\int \varphi > A$, d'après le théorème 39 bis. Soit K le support de φ. Soit $\delta = \frac{\int \varphi - A}{\mu(K)}$. Pour tout a de K, il existe un indice $i_a \in I$ tel que $f_{i_a}(a) \geq j(a) - \frac{\delta}{2} \geq \varphi(a) - \frac{\delta}{2}$. Il existe alors un voisinage ouvert $\mathcal{V}_a$ de a dans X tel que, pour tout $x \in \mathcal{V}_a$, $f_{i_a}(x) \geq \varphi(x) - \delta$, puisque φ est continue et f_{i_a} semi-continue inférieurement. K est recouvert par un nombre fini des $\mathcal{V}_a$, soit $\mathcal{V}_{a_1}, \mathcal{V}_{a_2}, \dots, \mathcal{V}_{a_n}$. Si alors nous appelons k un indice tel que $f_k \geq \operatorname{Sup}(f_{i_{a_1}}, f_{i_{a_2}}, \dots, f_{i_{a_n}})$ (condition de croissance) on aura $f_k \geq \varphi - \delta$ sur K tout entier, et par suite on aura partout $f_k \geq (\varphi - \delta)\chi_K$, puisque, en dehors de K, la 2ème fonction est nulle et la 1ère ≥ 0. On aura alors

$$\int^* f_k \geq \int(\varphi) - \delta\,\mu(K) \geq A$$

, ce qui démontre notre affirmation. C'est à peu près la seule propriété de la théorie de l'intégration qui soit valable pour une limite de familles non dénombrables. Le même résultat serait faux pour $f_i \geq 0$ arbitraires. Par exemple, soit $f = 1$ sur $\mathbb{R}$, $\mu = dx$. Prenons pour f_i les fonctions caractéristiques des ensembles finis. On a $\int^* f_i = 0$, $\int^* f = +\infty$.

Remarque - Nous avons toujours essentiellement supposé X dénombrable à l'infini. Voici un exemple de ce qui peut se passer si l'on ne fait pas cette hypothèse. Soit X l'espace somme d'une infinité non dénombrable de fois la droite $\mathbb{R}$;

plus spécialement, I sera un ensemble d'indices non dénombrable, et, pour tout $i \in I$, $\mathbb{R}_i$ sera isomorphe à $\mathbb{R}$, les $\mathbb{R}_i$ seront supposés disjoints dans X et de réunion X. La topologie de X sera la suivante : une partie est ouverte si son intersection avec tout $\mathbb{R}_i$, $i \subset I$, est ouverte. X n'est pas dénombrable à l'infini; car, si K est un compact de X, les $\mathbb{R}_i$ en forment un recouvrement ouvert, dont il existe un sous-recouvrement fini, autrement dit K est contenu dans la réunion d'un nombre fini de $\mathbb{R}_i$; donc une réunion dénombrable de compacts de X est contenue dans une réunion dénombrable de $\mathbb{R}_i$, et ne peut pas être X tout entier. Mettons sur chaque $\mathbb{R}_i$ la mesure dx. On définit ainsi une mesure de Radon $\mu \geq 0$ sur X ; si en effet $\varphi \in \mathscr{C}(X)$, son support, compact, est contenu dans une réunion finie de $\mathbb{R}_i$, soit $\bigcup_{i \in J} \mathbb{R}_i$, et on pourra poser $\mu(\varphi) = \sum_{i \in J} \int \varphi_i\, dx$, φ_i restriction de φ à $\mathbb{R}_i$. Considérons alors l'ensemble A formé des points 0 de tous les $\mathbb{R}_i$. A est fermé; cependant il n'est pas mesurable, au sens de la page 462. En effet, tout ouvert le contenant doit contenir des ouverts non vides sur tous les $\mathbb{R}_i$, donc sa mesure est sûrement $+\infty$; tout compact contenu dans A ne contient que des points 0 dans un nombre fini de $\mathbb{R}_i$, donc sa mesure est nulle; $\mu^*(A) = +\infty$, $\mu_*(A) = 0$. On n'a donc pas, par exemple $\mu^*(A) = \operatorname{Sup.} \mu^*(A \cap K)$ pour tous les compacts K de X (page). Soit χ_A la fonction caractéristique de A. Elle n'est donc pas mesurable au sens de la page 467; cependant elle vérifie la propriété de Lusin, qui entraîne la mesurabilité dans le cas dénombrable à l'infini. Son intégrale supérieure, définie comme page 474, est la borne supérieure des intégrale supérieures des fonctions bornées a support compact qui la minorent; or toute fonction ≥ 0 a support compact qui la minore a une intégrale nulle, donc $\int^* \chi_A = 0 \neq \mu^*(A)$. Par contre, toute fonction semi-continue inférieurement $\geq \chi_A$ a une intégrale infinie; dans la théorie de l'intégration où l'on définit l'intégrale supérieure par les fonctions semi-continues inférieurement, on a $\int^* \chi_A = \mu^*(A) = +\infty$.

Donc les diverses théories de l'intégration ne sont plus équivalentes. Si l'on tient absolument à englober dans la théorie les X même non dénombrables à l'infini, on doit prendre la définition de l'intégrale par les fonctions semi-continues inférieurement. La mesure des ensembles se définit par l'intégrale de leur fonction caractéristique; les fonctions mesurables se définissent par Lusin; un ensemble sera dit mesurable si sa fonction caractéristique est mesurable, c'est-à-dire vérifie Lusin. Dans ce cas, une grande partie des théorèmes antérieurs subsistent, mais pas tous. Ayant une fois pour toutes supposé X dénombrable à l'infini, nous ne l'avons répété dans les énoncés que lorsque c'était

Indispensable, même dans la théorie de l'intégration qui part des fonctions semi-continues inférieurement. Mais tout individu doué d'une prudence élémentaire se gardera même de songer à des espaces X localement compacts non dénombrables à l'infini, où les pires horreurs sont possibles.

Espaces $\mathcal{L}^p(X, \mu; F)$ Soit f une fonction réelle ≥ 0, à valeurs finies ou infinies, définie μ-presque partout sur X. Appelons norme d'ordre p de f où p est un nombre réel fini ≥ 1, la quantité définie comme suit :

$$N_p(f) = N_p(X, \mu; f) = \left(\int^* f^p\right)^{\frac{1}{p}} . \tag{IV,4;41}$$

Théorème 40 - Si $f \leq g$, on a $N_p(f) \leq N_p(g)$.

On a les relations :

$$N_p(f+g) \leq N_p(f) + N_p(g), \quad N_p(kf) = |k| N_p(f) ; \tag{IV,4;42}$$

(inégalité de convexité ou de Minkowski); plus généralement, on a l'inégalité de convexité dénombrable :

$$N_p\left(\sum_{n=0}^{\infty} u_n\right) \leq \sum_{n=0}^{\infty} N_p(u_n) , \tag{IV,4;43}$$

si les u_n sont des fonctions ≥ 0 sur X.

Démonstration - La relation relative à kf est évidente. Considérons deux fonctions f et g. Supposons d'abord qu'elles soient étagées à support compact, et que l'on puisse ecrire, en prenant une décomposition commune :

$$f = \sum_{i=1}^{n} f_i \varphi_{X_i} , \quad g = \sum_{i=1}^{n} g_i \varphi_{X_i} , \; f_i \text{ et } g_i \text{ constantes} \tag{IV,4;44}$$

les ensembles X_i étant des parties mesurables de X.

Alors l'intégrale $N_p(f)$ n'est autre que la norme d'ordre p : $\|\vec{Z}\|_p$, où $\vec{Z}$ est le vecteur de $\mathbb{C}^n$ de composantes $Z_1, Z_2, \ldots, Z_n$, avec $Z_i = f_i$, relativement aux poids $c_1, c_2, \ldots, c_n$, où $c_i = \mu(X_i)$.

On opère de même pour g, et alors l'inégalité de convexité n'est autre que l'inégalité de Minkowski, démontrée au théorème 35 du chapitre III.

Supposons maintenant f et g quelconques, mais bornées à support compact. Désignons alors par f_1 et g_1 deux fonctions étagées à support compact qui les dépassent. Alors on a l'inégalité :

$$(IV,4;45)\qquad N_p(f+g) \leqslant N_p(f_1+g_1) \leqslant N_p(f_1) + N_p(g_1).$$

En prenant la borne Inférieure du second membre, on en déduit bien l'inégalité de **convexité** cherchée pour f et g.
Passons enfin au cas où f et g sont $\geqslant 0$ quelconques.
SI alors nous construisons les fonctions $f_{M,K}$, $g_{M,K}$, $(f+g)_{M,K}$, déjà définies antérieurement, on a **l'inégalité**

$$(IV,4;46)\qquad (f+g)_{M,K} \leqslant f_{M,K} + g_{M,K} \quad ,$$

d'où l'on déduit **l'inégalité**

$$(IV,4;47)\qquad N_p\Big((f+g)_{M,K}\Big) \leqslant N_p(f_{M,K}) + N_p(g_{M,K}) \leqslant N_p(f) + N_p(g).$$

Il suffit alors de prendre la borne supérieure du ler membre pour toutes les valeurs de M et tous les K , pour obtenir le résultat.

Démontrons enfin l'inégalité de convexité dénombrable. Pour la somme finie desmpremiers termes, on a l'inégalité

$$(IV,4;48)\qquad N_p(u_0+u_1+\cdots+u_m) \leqslant N_p(u_0) + N_p(u_1) + \cdots + N_p(u_m).$$

Il suffit alors de passer **à** la limite pourmtendant vers l'infini, en utilisant le théorème 36 de **Fatou,** pour obtenir le résultat.

<u>Définition</u> Soit X un espace localement compact **dénombrable** <u>à l'infini</u>, muni d'une mesure de Radon $\mu \geqslant 0$, et soit $\vec{F}$ un espace de Banach. Soit p un nombre réel fini $\geqslant 1$
<u>On dit qu'une fonction $\vec{f}$ **définie** sur X</u> , <u>**à** valeurs dans? $\vec{F}$</u> , <u>appartient a $\mathcal{L}^p(X , \mu ; \vec{F})$</u> , <u>ou simplement $\mathcal{L}^p(\vec{F})$ **si** aucune confusion n'est à craindre</u> (*), <u>ou qu'elle est de **puissance** p-ième intégrable</u>, <u>si, quel que soit</u> $\varepsilon > 0$, <u>**il** existe une fonction $\vec{g}$</u> **,** <u>**défi**nie sur X</u> **,** <u>à valeurs dans $\vec{F}$</u> , <u>**étagée** a support compact</u>, <u>telle que **l'on** ait l'inégalité</u>

* Par contre , **il** ne faut pas laisser **disparaître** $\vec{F}$. On **réserve** le nom de $\mathcal{L}^p(X, \mu)$ ou $\mathcal{L}^p$ au cas où $\vec{F}$ est le corps des scalaires.

$N_p(\|\vec{f}-\vec{g}\|) \leq \varepsilon$; ou encore , s'il existe une suite de fonctions $\vec{f}_0, \vec{f}_1, \ldots, \vec{f}_n, \ldots,$ étagées à support **compact**, telles que la quantité $N_p(\|\vec{f} - \vec{f}_n\|)$ converge vers 0, pour n tendant vers l'infini. On dit alors que la suite des $\vec{f}_n$ est une suite d'approximations de $\vec{f}$ dans $\mathcal{L}^p(X,\mu;\vec{F})$, par des fonctions étagées.* Par abus de langage, on se permet de dire que $\vec{f}$ appartient à $\mathcal{L}^p(\vec{F})$, si elle est seulement définie presque partout, et si, prolongée n'importe comment là où elle n'est pas définie, elle est dans $\mathcal{L}^p(\vec{F})$.

Z L'espace $\mathcal{L}^1(X,\mu;\vec{F})$ n'est autre que l'espace des fonctions intégrables à valeurs dans $\vec{F}$. Nous généralisons donc l'espace des fonctions intégrables. Cependant, pour $p \neq 1$, il ne sera Pas question d'intégrale d'une fonction appartenant à $\mathcal{L}^p(X,\mu;\vec{F})$.

Théorème 41 - Si $\vec{f}$ appartient à $\mathcal{L}^p(\vec{F})$, alors la fonction ≥ 0 $\|\vec{f}\| : x \to \|\vec{f}(x)\|$, appartient à $\mathcal{L}^p$.

Démonstration - La démonstration se fait comme celle du théorème 23, correspondant à $p = 1$.

Théorème 42 - L'espace $\mathcal{L}^p(\vec{F})$ est un espace vectoriel, et, sur cet espace, la fonction N_p définie par

$$N_p(\vec{f}) = N_p(\|\vec{f}\|) = \left(\int^* \|\vec{f}\|^p\right)^{\frac{1}{p}} \qquad (\text{IV},4;49)$$

est une seml-norme.

* Par une démonstration analogue à celle du théorème 32, s'appuyant sur le lemme 2 du théorème 30, mais cette fois avec $\delta = \left(\frac{\varepsilon}{2NM}\right)^p$, on verra qu'on peut remplacer les fonctions étagées par des fonctions continues, à support compact, décomposables.

Démonstration - Soient $\vec{f}_n$ et $\vec{g}_n$ des suites d'approximation de $\vec{f}$ et de $\vec{g}$ dans $\mathcal{L}^p(\vec{F})$; alors l'inégalité de Minkowski montre" que les $\vec{f}_n + \vec{g}_n$ forment une suite'd'approxlmation de $\vec{f} + \vec{g}$, qui par conséquent appartient aussi à $\mathcal{L}^p(F)$.

Démonstration évidente relativement à $k\vec{f}$, k constante scalaire. On appelle **semi-norme** sur un espace vectoriel, une fonction N à valeurs ≥ 0 finies, vérifiant les relations

(IV,4;50)
$$\begin{cases} N(k\vec{X}) = |k| \, N(\vec{X}) \\ N(\vec{X}+\vec{Y}) \leq N(\vec{X}) + N(\vec{Y}) \; . \end{cases}$$

Si, en outre, $N(\vec{X}) = 0$ entraîne $\vec{X} = \vec{0}$, c'est une norme, comme nous l'avons déjà vu. Le fait que N_p est une semi-norme résulte alors trivialement du théorème 40 de Minkowski; par contre ce n'est manifestement pas une norme; si en effet on a $N_p(\vec{f}) = 0$, c'est-à-dire si $\int^* \|\vec{f}\|^p = 0$, cela prouve simplement que la fonction $\vec{f}$ est presque partout nulle (théorème 26), et ce n'est donc pas nécessairement l'élément nul de l'espace vectoriel $\mathcal{L}^p(\vec{F})$.

Théorème 43 - Soit $\vec{u}_n$ une suite de fonctions définies presque partout sur X à valeurs dans F appartenant à $\mathcal{L}^p(\vec{F})$ telles que la série $\sum_{n=0}^{\infty} N_p(\vec{u}_n)$ soit convergente. Alors la série $\sum_{n=0}^{\infty} \vec{u}_n(x)$ est presque partout absolument convergente, et, si l'on appelle $\vec{S}(x)$ sa somme, $\vec{S}(x) = \sum_{n=0}^{\infty} \vec{u}_n(x)$, $\vec{S}$ est une fonction définie presque partout, qui appartient à $\mathcal{L}^p(\vec{F})$; en outre on a la relation

$$N_p(\vec{S}) = N_p\left(\sum_{n=0}^{\infty} \vec{u}_n\right) \leq \sum_{n=0}^{\infty} N_p(\vec{u}_n) \qquad \text{(IV,4;51)}$$

(inégalité de convexité dénombrable);

$N_p(\vec{S} - \vec{S}_m) = N_p(\vec{R}_m)$ tend vers 0 pour m infini, et $N_p(\vec{S}_m)$ tend vers $N_p(\vec{S})$.

Démonstration analogue à celle du cas $p = 1$ (théorème 37).

Bien que N_p soit une semi-norme et non une norme, et que par conséquent $\mathcal{L}^p(\vec{F})$ soit seulement un espace vectoriel semi-normé, on emploie souvent le langage des espaces normés. Par exemple, si $\vec{f}_n$ est une suite de fonctions de $\mathcal{L}^p(\vec{F})$, et si $N_p(\vec{f} - \vec{f}_n)$ tend vers 0, on dira que " $\vec{f}_n$ tend vers $\vec{f}$ dans $\mathcal{L}^p(\vec{F})$ "; mais ici la limite n'est pas unique, si $\vec{f}_n$ "tend vers" $\vec{f}$, elle tend aussi vers $\vec{g}$ dès que $\vec{g}$ est presque partout égale à $\vec{f}$. Alors, dans le théorème 43, on dira que $\vec{S}_m$ tend vers $\vec{S}$ dans $\mathcal{L}^p(\vec{F})$, puisque $N_p(\vec{S} - \vec{S}_m)$ tend vers 0 pour m infini; on dira aussi que la série $\sum_{n=0}^{\infty} \vec{u}_n$ est convergente dans $\mathcal{L}^p(\vec{F})$, et même absolument convergente

(C'est d'autant plus un abus de langage que $\vec{s}$ n'est, définie que presque partout, et que l écriture $\vec{s} \in \mathcal{L}(F)$ est déjà un abus de langage).

Corollaire - Si une suite de fonctions $\vec{f}_n$ converge vers $\vec{f}$ dans $\mathcal{L}^p(\vec{F})$, pour n tendant vers l'infini, alors on peut extraire de la suite considérée une suite partielle qui converge presque partout vers $\vec{f}$.

Démonstration analogue à celle du théorème 38.

Théorème 44 - Pour qu'une fonction $\vec{f}$, définie sur X, à valeurs dans l'espace de Banach $\vec{F}$, appartienne à $\mathcal{L}^p(\vec{F})$, il faut et il suffit qu'elle soit mesurable, et que $\int^* \|\vec{f}\|^p$ soit finie.

La démonstration est exactement la même que pour $p = 1$, théorème 39.

Corollaire - Si $\vec{f}$ appartient à l'espace $\mathcal{L}^p(\vec{F})$, alors la fonction $\|\vec{f}\|^p : x \to \|\vec{f}(x)\|^p$, est intégrable sur X.

Plus généralement, si $q \leq p$, la fonction $\|\vec{f}\|^q$ appartient à $\mathcal{L}^{p/q}(X,\mu) = \mathcal{L}^{p/q}$

Démonstration - Il résulte de l'hypothèse, et du théorème, que la fonction $\vec{f}$ est mesurable; il en est donc de même de la fonction $\|\vec{f}\|^q$, parce que la norme est une fonction continue (théorème 9 du chapitre II, et théorème 22); alors l'inégalité $\int^* (\|\vec{f}\|^q)^{p/q} = (N_p(\vec{f}))^p < +\infty$, donne le résultat en vertu du théorème 44.

Théorème 45 - (effet d'une application linéaire continue). Si L est une application linéaire continue de l'espace de Banach $\vec{F}$ dans l'espace de Banach $\vec{G}$, et si $\vec{f}$ est une fonction définie sur X à valeurs dans $\vec{F}$, appartenant à $\mathcal{L}^p(\vec{F})$, alors la fonction $L \circ \vec{f}$ appartient à $\mathcal{L}^p(\vec{G})$, et en outre on a l'inégalité

$$N_p(L \circ \vec{f}) \leq \|L\| \, N_p(\vec{f}) . \tag{IV,4;52}$$

Si en particulier $p = 1$, $L \circ \vec{f}$ est intégrable, et on a :

$$\int L \circ \vec{f} = L\left(\int \vec{f}\right) . \tag{IV,4;53}$$

L'inégalité est évidente, puisque $\|(L\circ\vec{f})(x)\| \leq \|L\| \, \|\vec{f}(x)\|$, le reste est évident par la considération d'une suite d'approximation. Ceci nous montre que <u>l'application linéaire continue L de $\vec{F}$ dans $\vec{G}$ définit une application linéaire continue de $\mathcal{L}^p(\vec{F})$ dans $\mathcal{L}^p(\vec{G})$, de norme $\leq \|L\|$</u> * .

<u>Théorème 46 - Effet d'applications bilinéaires continues : inégalité de Hölder. Si $\vec{f}$ appartient à $\mathcal{L}^p(\vec{F})$, et si $\vec{g}$ appartient à $\mathcal{L}^q(\vec{G})$, si B est une application bilinéaire continue de $\vec{F}\times\vec{G}$ dans $\vec{H}$, alors, si $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$, et si r ainsi déterminé est ≥ 1, la fonction $B(\vec{f},\vec{g}) : x \to B(\vec{f}(x),\vec{g}(x))$, appartient à $\mathcal{L}^r(\vec{H})$, et en outre on a l'inégalité de Hölder</u> :

$$N_r(B(\vec{f},\vec{g})) \leq \|B\| N_p(\vec{f}) \; N_q(\vec{g}) . \qquad (IV,4;54)$$

<u>Démonstration</u> -Soient d'abord f et g deux fonctions réelles ≥ 0.

Alors l'inégalité de Hölder relative au produit $f\,g$ se démontre exactement de la même manière que nous avons démontré l'inégalité de Minkowski (théorème 40), mais cette fois en se ramenant au théorème 36 du chapitre III. *On* en déduit (IV,4;54) dans tous les cas, en se ramenant immédiatement au cas du produit de deux fonctions réelles ≥ 0, puisque $\|B(\vec{f}(x),\vec{g}(x))\| \leq \|B\| \; \|\vec{f}(x)\| \; \|\vec{g}(x)\|$.

Remarquons maintenant qu'en vertu des hypothèses faîtes sur $\vec{f}$ et $\vec{g}$, ce sont des fonctions mesurables, alors le corollaire 3 du théorème 23 bis montre que $B(\vec{f},\vec{g})$ est mesurable de X dans $\vec{H}$.

Il suffit alors d'appliquer à cette fonction le théorème 44 pour voir qu'elle appartient bien à $\mathcal{L}^r(\vec{H})$.

* Avec l'abus de langage signalé page 507 , puisque $\mathcal{L}^p(\vec{F})$ est seulement semi-normé.

Remarques 1°) Si r est < 1 (par exemple si $p = 1$, $q = 1$, auquel cas $r = \frac{1}{2}$), on ne peut rien conclure.

2°) Le théoréme peut s'énoncer en disant que $(\vec{f}, \vec{g}) \to B(\vec{f}, \vec{g})$ est une application bilinéaire continue de $\mathcal{L}^p(\vec{F}) \times \mathcal{L}^q(\vec{G})$ dans $\mathcal{L}^r(\vec{H})$ (avec l'abus de langage de la page **507**).

Corollaire 1 - Soient p et p' deux exposants conjugués, c'est-a-dire vérifiant $\frac{1}{p} + \frac{1}{p'} = 1$, ou $p' = \frac{p}{p-1}$. Si alors $\vec{f}$ et $\vec{g}$ appartiennent respectivement à $\mathcal{L}^p(\vec{F})$ et $\mathcal{L}^{p'}(\vec{G})$, la fonction $B(\vec{f}, \vec{g})$ est intégrable et en outre on a l'inégalité:

$$\left\| \int B(\vec{f}, \vec{g}) \right\| \leq \|B\| \, N_p(\vec{f}) \, N_{p'}(\vec{g}) \, . \qquad (\text{IV},4;55)$$

Il suffit en effet d'appliquer le théorème, compte tenu de ce que $\left\| \int B(\vec{f}, \vec{g}) \right\| \leq \int \| B(\vec{f}, \vec{g}) \| = N_1\left(B(\vec{f}, \vec{g})\right)$

Ce corollaire est particulièrement important dans le cas où il s'agit de fonctions à valeurs réelles ou complexes, et où l'application bilinéaire est simplement le produit. On a dans ce cas l'inégalité

$$\left| \int f g \right| \leq N_p(f) \, N_{p'}(g) \qquad (\text{IV},4;56)$$

Corollaire 2 - Si $\vec{f}$ et $\vec{g}$ appartiennent respectivement à $\mathcal{L}^2(\vec{F})$ et à $\mathcal{L}^2(\vec{G})$, alors la fonction $B(\vec{f}, \vec{g})$ est intégrable, et on a l'inégalité

$$\left\| \int B(\vec{f}, \vec{g}) \right\| \leq \|B\| \, N_2(\vec{f}) \, N_2(\vec{g}) \, . \qquad (\text{IV},4;57)$$

Ce théorème est particulièrement important dans le cas de fonctions a valeurs complexes, et l'on a alors l'inégalité de Cauchy-Schwarz :

$$\left| \int f \bar{g} \right| \leq \left(\int |f|^2 \right)^{\frac{1}{2}} \left(\int |g|^2 \right)^{\frac{1}{2}} \, . \qquad (\text{IV},4;58)$$

On sait comment cette inégalité se démontre directement. On écrit tout simplement que

$$\int (f+\lambda g)(\bar{f}+\bar{\lambda}\bar{g}) = \int |f+\lambda g|^2 \geq 0 \,, \tag{IV,4;59}$$

et en développant suivant λ, on trouve que le trinôme

$$\int |f|^2 + \lambda \int g\bar{f} + \bar{\lambda}\int f\bar{g} + |\lambda|^2 \int |g|^2 \tag{IV,4;60}$$

est ≥ 0 . On en déduit que son discriminant est négatif, d'où (IV,4;58).

Cela résulte aussi tout simplement de l'inégalité de Cauchy-Schwarz (théorème 2 ter du chapitre III) appliquée à la forme sesquilinéaire $(f,g) \to \int f\bar{g}$, définie sur l'espace vectoriel $\mathcal{L}^2$. Mais encore faut-il avoir montré que, si $f \in \mathcal{L}^2$ et $g \in \mathcal{L}^2$, $f\bar{g}$ est intégrable et que $f+g \in \mathcal{L}^2$, pour pouvoir parler de l'espace vectoriel $\mathcal{L}^2$ et définir la forme sesquilinéaire ci-dessus, aussi bien oue pour pouvoir écrire (IV, 4;59 et 60). On peut le faire directement comme suit : $f+g$ et $f\bar{g}$ sont mesurables; ensuite

$$|f+g|^2 \leq 2\left(|f|^2+|g|^2\right), \quad |f\bar{g}| \leq |f|^2+|g|^2 \;; \tag{IV,4;61}$$

donc $N_2(f+g)$ et $N_1(f\bar{g})$ sont finis, et le théorème 44 donne le résultat. Dans le cas précis de (IV,4;58), il est utile de savoir quand on a l'inégalité large $\leq$, et quand on a l'inégalité stricte $<$. Si d'une part $N_2(\vec{g}) > 0$, et si d'autre part, pour toute valeur de λ, $N_2(\vec{f}+\lambda\vec{g}) > 0$, alors le trinôme (IV,4;59) ne doit jamais s'annuler, son discriminant est < 0 , et par suite l'inégalité est une inégalité stricte $<$. On voit donc qu'on a toujours l'inégalité stricte, sauf s'il existe deux constantes complexes non simultanément nulles α et β, telles que la fonction $\alpha\vec{f}+\beta\vec{g}$ soit presque partout nulle.

Corollaire 3 - Critère d'appartenance aux divers espaces $\mathcal{L}^p(\vec{F})$. Si $\|\mu\| < +\infty$, alors toute fonction appartenant à $\mathcal{L}^p(\vec{F})$ appartient à tous les $\mathcal{L}^r(\vec{F})$, pour $r \leq p$, et en outre on a les inégalités

(IV,4;61) $$N_r(\vec{f}) \leq \|\mu\|^{\frac{1}{r}-\frac{1}{p}} N_p(\vec{f}) .$$

Démonstration - Il suffit en effet d'appliquer le théorème 46 à la fonction $\vec{f}$ et à la fonction scalaire constante $g = 1$, la fonction B étant alors l'application bilinéaire $(\vec{f}, k) \to k\vec{f}$ de $\vec{F} \times \mathbb{K}$ dans?, avec q défini par $\frac{1}{q} = \frac{1}{r} - \frac{1}{p}$. *

Ce théorème est particulièrement important dans les applications. Si par exemple X est un intervalle borné de la droite reelle, et $\mu = dx$, on voit que l'appartenance à $\mathcal{L}^p$ entraîne l'appartenance à tous les $\mathcal{L}^r$, pour $r \leq p$; en particulier une fonction appartenant à l'un quelconque des $\mathcal{L}^p$ appartient à $\mathcal{L}^1$, et est nécessairement intégrable. Par contre il est bien évident que :

1°/ Il n'existe pas de relation en sens inverse entre les $\mathcal{L}^p$ • une fonction qui appartient à $\mathcal{L}^p$ n'a aucune raison d'appartenir à $\mathcal{L}^s$, pour $s > p$ Si, par exemple, sur l'intervalle compact $[0,1]$ de la droite réelle $\mathbb{R}$, pour la mesure $d\mu = dx$, nous considérons la fonction $f(x) = \frac{1}{x^\alpha}$, alors, pour $\alpha = \frac{1}{2}$, elle appartient à $\mathcal{L}^1$, mais n'appartient pas à $\mathcal{L}^2$.

2°/ Il n'existe plus aucune relation entre les appartenances aux divers espaces $\mathcal{L}^p$, lorsque la norme de la mesure μ est infinie. Si par exemple nous prenons $X = \mathbb{R}$, et $d\mu = dx$ alors l'exemple donné dans 1°) montre que l'appartenance à $\mathcal{L}^p$ n'entraîne pas l'appartenance à $\mathcal{L}^s$, $s > p$; inversement la fonction, définie par $f(x) = \frac{1}{\sqrt{1+x^2}}$, appartient alors à $\mathcal{L}^2$ et n'appartient pas à $\mathcal{L}^1$.

* Si on cherche seulement à montrer que $\vec{f} \in \mathcal{L}^r(\vec{F})$, sans s'occuper de (IV,4;61), c'est tout à fait évident. Car $\vec{f}$ est mesurable; ensuite $\|\vec{f}(x)\|^r \leq \|\vec{f}(x)\|^p$ si $\|\vec{f}(x)\| \geq 1$, et ≤ 1 si $\|\vec{f}(x)\| \leq 1$. Donc $\|\vec{f}(x)\|^r \leq 1 + \|\vec{f}(x)\|^p$, et l'intégrale supérieure de $\|\vec{f}\|^r$ est finie (comme $\|\mu\| < +\infty$, $\int^* 1$ est fini).

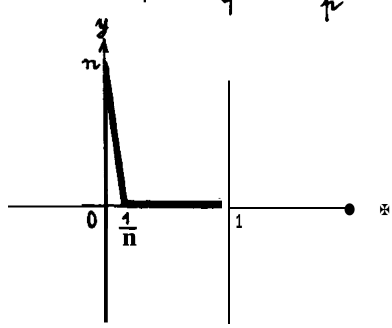
3°/ Sur l'espace $(\mathbb{C}^{[0,1]})_{cb}$ des fonctions complexes continues sur l'intervalle $[0,1]$ de $\mathbb{R}$, <u>les normes N_p relatives à la mesure dx sont toutes Inéquivalentes,</u> ce qui donne un exemple simple de normes Inéquivalentes (voir page 56, théorème 13 du chapitre II) sur un espace vectoriel de dimension infinie. Tout d'abord N_p est une norme et pas seulement une **semi-norme**, sur cet espace; en effet, si $N_p(f) = 0$, f doit être presque partout nulle; mais elle est continue; soit alors $c \in [0,1]$; pour $\varepsilon > 0$ donné, il doit exister un nombre $\eta > 0$, tel **que** $|x-c| \leqslant \eta$ **entraine** $|f(x) - f(c)| \leqslant \varepsilon$; l'ensemble de ces points x a une longueur $\geqslant \eta$, **et** f est presque partout nulle, donc il existe au moins **un tel** point x tel que $f(x) = 0$, **donc** $|f(c)| \leqslant \varepsilon$; comme ε est **arbitraire**, $f(c) = 0$, f est identiquement nulle, et N_p est bien une norme.
Alors, pour tout **entier** n, considérons la fonction $f_n(x)$ définie par :

(IV,4;61bis)
$$f_n(x) = \begin{cases} n(1-nx) & \text{pour } 0 \leqslant x \leqslant \frac{1}{n}, \\ 0 & \text{pour } x > \frac{1}{n}. \end{cases}$$

Elle est bien continue. On a

(IV,4;61ter)
$$N_p(f_n) = n^{1-\frac{1}{p}} \left(\frac{1}{p+1}\right)^{\frac{1}{p}}.$$

Faisons **tendre** n vers $+\infty$; alors, pour $q > p$, il est bien impossible qu'il existe un nombre fixe $k > 0$ tel que $N_q(f_n) \leqslant k\, N_p(f_n)$, et le théorème 12 du chapitre II montre alors que N_q et N_p sont inequivalentes.



<u>Corollaire 4 - Si $\vec{f}$ appartient à la fois à $\mathcal{L}^p(\vec{F})$ et à $\mathcal{L}^q(\vec{F})$, elle appartient à tous les $\mathcal{L}^r(\vec{F})$, $p \leqslant r \leqslant q$, et on a l'inégalité</u> :

$$N_r(\vec{f}) \leqslant \left(N_p(\vec{f})\right)^\alpha \left(N_q(\vec{f})\right)^\beta \quad , \qquad (\text{IV},4;61\text{quarto})$$

où

$$\alpha = \frac{\frac{1}{r} - \frac{1}{q}}{\frac{1}{p} - \frac{1}{q}} \quad , \quad \beta = \frac{\frac{1}{p} - \frac{1}{r}}{\frac{1}{p} - \frac{1}{q}}$$

<u>Démonstration</u> - Comme $\vec{f}$ est mesurable, il suffit (théorème 44) de vérifier (IV,4;61quarto) pour $\|\vec{f}\|$, ou encore pour une fonction g mesurable réelle $\geqslant 0$ *.

Comme $p \leqslant r \leqslant q$, on a aussi $\frac{1}{q} \leqslant \frac{1}{r} \leqslant \frac{1}{p}$; alors $\frac{1}{r}$ est barycentre de $\frac{1}{p}$ et $\frac{1}{q}$, pour des poids α et $\beta \geqslant 0$, de somme $\alpha + \beta = 1$: $\frac{1}{r} = \frac{\alpha}{p} + \frac{\beta}{q}$. Ces nombres α et β sont précisément ceux qui sont dans l'énoncé. Alors, comme $\alpha + \beta = 1$, on a $g = g^\alpha g^\beta$;
on a $g^\alpha \in \mathcal{L}^{\frac{p}{\alpha}}$, $g^\beta \in \mathcal{L}^{\frac{q}{\beta}}$, et $N_{\frac{p}{\alpha}}(g^\alpha) = (N_p(g))^\alpha$, $N_{\frac{q}{\beta}}(g^\beta) = (N_q(g))^\beta$.

Alors, si l'on applique l'inégalité de Hölder (IV,4;54) au produit $g^\alpha g^\beta = g$, avec les exposants $\frac{p}{\alpha}, \frac{q}{\beta}$, $\frac{1}{\frac{p}{\alpha}} + \frac{1}{\frac{q}{\beta}} = \frac{1}{r}$, on trouve (IV,4;61 quarto).

Comme conséquence : <u>l'ensemble des p finis $\geqslant 1$ pour lesquels une fonction donnée $\vec{f}$ appartient à $\mathcal{L}^p(\vec{F})$, est un Intervalle de $[1, +\infty[$, ouvert, semi-ouvert, ou fermé</u>. En outre, l'inégalité (IV,4;61 quarto) montre que, sur cet intervalle, la fonction $\xi \longrightarrow \log N_{\frac{1}{\xi}}(\vec{f})$ est convexe (voir page 192).

* Si on ne cherche pas à démontrer (IV,4;61quarto), mais seulement l'appartenance de $\vec{f}$ à $\mathcal{L}^r(\vec{F})$, c'est tout à fait évident. Comme en effet $\vec{f}$ est mesurable, il suffit de montrer que $N_r(\vec{f})$ est fini. Or $\|\vec{f}(x)\|^r \leqslant \|\vec{f}(x)\|^p$ si $\|\vec{f}(x)\| \leqslant 1$, et $\leqslant \|\vec{f}(x)\|^q$ si $\|\vec{f}(x)\| \geqslant 1$. On a donc $\|\vec{f}(x)\|^r \leqslant \|\vec{f}(x)\|^p + \|\vec{f}(x)\|^q$, et son intégrale supérieure est bien finie.

Les espaces $L^p(X, \mu; \vec{F})$; le théorème de Fischer-Riesz

L'étude des espaces $\mathcal{L}^p$ nous a montre de façon constante combien il était gênant d'avoir une semi-norme et non une norme, car nous ne pouvions alors appliquer aucun des théorèmes relatifs aux espaces vectoriels normés.

C'est pourquoi on va introduire un nouvel espace normé. Dans l'espace $\mathcal{L}^p(\vec{F})$, les fonctions pour lesquelles $N_p(\vec{f})=0$ forment un sous-espace vectoriel, à savoir l'ensemble des fonctions presque partout nulles; nous pourrons alors considérer le quotient de $\mathcal{L}^p(\vec{F})$ par ce sous-espace vectoriel. C'est ce quotient qu'on appelle l'espace $L^p(X, \mu; \vec{F})$, ou $L^p(\vec{F})$ si aucune confusion n'est à craindre. Un élément de $L^p(\vec{F})$ n'est donc pas une fonction définie sur X à valeursdans $\vec{F}$, mais une classe de fonctions deux a deux presque partout égales. On dit d'ailleurs d'une façon générale, que deux fonctions définies sur X sont μ -équivalentes au sens de Lebesgue, ou simplement equivalentes si aucune confusion n'est a craindre, si elles sont μ -presque partout égales. Alors $L^p(\vec{F})$ est "l'espace des fonctions appartenant à $\mathcal{L}^p(\vec{F})$, et définies à une équivalence près".

Bien entendu $L^p(\vec{F})$ est un espace vectoriel, puisque c'est un quotient d'un espace vectoriel par un sous-espace vectoriel (chapitre 1, page 15). En outre c'est un espace vectoriel normé. Si en effet nous considérons un élément $\vec{f}^{\bullet}$ de $L^p(\vec{F})$, il est constitué par un ensemble de fonctions $\vec{f}$ de $\mathcal{L}^p(\vec{F})$, qui ont toutes la même semi-norme $N_p(\vec{f})$, et cette semi-norme commune s'appelle la norme de l'élément considéré $\vec{f}^{\bullet}$ de $L^p(\vec{F})$; c est évidemment une norme, car elle vérifie trivialement les inégalités (IV,4;50); en outre, cette fois-ci, si $N_p(\vec{f}^{\bullet}) = 0$, alors $\vec{f}^{\bullet}$ est la classe constituée de toutes les fonctions resque partout nulles, c'est-a-dire l'élément nul de L^p (3)

Théorème 47 - (Théorème de Fischer-Riesz) - L'espace vectoriel normé $L^p(X, \mu; \vec{F})$, où $\vec{F}$ est un espace de Banach, est lui-même un espace de Banach; autrement dit, il est complet.

Démonstration - Pour le ddmontrer, nous pouvons nous appuyer sur le theoreme 56 du chapitre II:

Soit alors $\vec{u}_n^\bullet$ une suite d'éléments de $L^p(\vec{F})$, telle que la série $\sum_{n=0}^{\infty} N_p(\vec{u}_n^\bullet)$ soit convergente. Dans chaque classe $\vec{u}_n^\bullet$ choisissons un représentant, c'est-à-dire un élément $\vec{u}_n$ de $\vec{u}_n^\bullet$. On a alors la relation $\sum_{n=0}^{\infty} N_p(u_n) < +\infty$. Le théorème 43 montre alors que la série $\sum_{n=0} \vec{u}_n$ converge presque partout vers un élément $\vec{S}$ de $\mathcal{L}^p(\vec{F})$; et. qu'en outre $N_p(\vec{S}-\vec{S}_m)$ converge vers 0 pour m tendant vers l'infini. Si l'on appelle $\vec{S}^\bullet$ la classe de $\vec{S}$, c'est un élément de L * $(\vec{F})$, et la série $\sum_{n=0}^{\infty} \vec{u}_n^\bullet$ converge vers $\vec{S}^\bullet$ dans $L^p(\vec{F})$. Cela prouve bien que cet espace est complet.

<u>Théorème 48</u> - <u>Si $\vec{f}^\bullet$ est un élément de $L^1(\vec{F})$, on peut définir $\int \vec{f}^\bullet$ comme étant la valeur de $\int \vec{f}$, pour un élément quelconque $\vec{f}$ de $\vec{f}^\bullet$. Alors l'intégrale est une application linéaire continue de $L^1(X,\mu;\vec{F})$ dans $\vec{F}$, de norme 1. On a :</u>

(IV,4;62) $$\left\| \int \vec{f}^\bullet \right\| \leqslant N_1(\vec{f}^\bullet)$$

La démonstration est triviale, tout au moins en ce qui concerne le fait que la norme est $\leqslant 1$. Le fait que la norme est exactement 1 est plus difficile à démontrer, et nous l'admettrons: c est toutefois évident si F est le corps des scalaires, car alors, si f est une fonction $\geqslant 0$, on a exactement $\int f^\bullet = N_1(f)$.

<u>Théorème 49</u> - <u>Dans $L^p(\vec{F})$, l'ensemble des classes des fonctions étagées à support compact, ou l'ensemble des classes des fonctions continues à support compact, décomposables sont denses.</u>

En ce qui concerne les fonctions étagées, c'est évident, puisque, si $\vec{f}^\bullet \in L^p(\vec{F})$, et si $\vec{f} \in \mathcal{L}^p(\vec{F})$ est un représentant de $\vec{f}^\bullet$, alors il existe une-suite d'approximation de $\vec{f}$ formée de fonctions étagées $\vec{f}_n$; les $\vec{f}_n^\bullet$ convergent vers $\vec{f}^\bullet$ dans $L^p(\vec{F})$. Pour les fonctions continues, voir (*) page 506.

Remarque - Nous avons dit (théorème 41 du chapitre II) que tout espace vectoriel normé de dimension finie était complet; nous avons annoncé à ce moment qu'il n'en était plus de même pour les espaces vectoriels normés de dimension infinie. L'espace $L^p([0,1], dx)$ est bien complet, mais le **sous**-espace vectoriel des classes des fonctions continues est dense, donc non fermé, et par suite non complet (corollaire 1 du théorème 42 du chapitre II). On peut encore en déduire facilement que l'espace $(\mathcal{C}^{[0,1]})_{c\ell}$, muni de la norme N_p (voir page 513), n'est pas complet

Les espaces $\mathcal{L}^\infty(\vec{F})$ et $L^\infty(\vec{F})$

Définition - On dit qu'une fonction f, μ -mesurable, réelle $\geqslant 0$, ddfinie sur X, est p-essentiellement bornée, ou simplement essentiellement bornée si aucune confusion n'est à craindre, s'il existe un nombre M tel que f soit presque partout majorée par M. La borne inférieure M_0 des nombres M possédant cette propriété s'appelle la borne supérieure essentielle de f et se note $N_\infty(f)$. Elle est **caracté**risée par les deux propriétés suivantes :

1°/ Quel que soit ε 7 0 , l'inégalité $f(x) > M - \varepsilon$ est **vérifiée** pour un ensemble de valeurs de x de mesure 7 0;

2°/ L'inégalité $f(x) > M$ n'est **vérifiée** que pour un ensemble de valeurs de x de mesure nulle.

Si f n'est pas essentiellement bornée, on pose $N_\infty(f) = +\infty$.

On a trivialement $N_\infty(f+g) \leqslant N_\infty(f) + N_\infty(g)$, et $N_\infty(k\, f) = k\, N_\infty(f)$, pour $k \geqslant 0$.

On appelle alors $\mathcal{L}^\infty(X, \mu ; \vec{F})$, ou $\mathcal{L}^\infty(\vec{F})$ si aucune confusion n'est à craindre, l'ensemble des fonctions sur X à valeurs dans $\vec{F}$, mesurables, et essentiellement **bornées**, c'est-&-dire telles que la fonction $\|\vec{f}\|$ **soit** μ **-essentiellement bornée.** On définit sur cet espace la seml-norme $N_\infty(\vec{f})$ par l'égalité $N_\infty(\vec{f}) = N_\infty(\|\vec{f}\|)$.

On appelle alors $L^\infty(X, \mu ; \vec{F})$, ou $L^\infty(\vec{F})$ si aucune confusion n'est à craindre, l'espace vectoriel quotient de l'espace $\mathcal{L}^{\infty}(X, \mu\,;\vec{F})$ par le sous-espace vectoriel des fonctions μ-presque partout nulles. Alors c'est un, espace vectoriel normé, par la norme $N_\infty(\vec{f}^{\bullet}) = N_\infty(\vec{f})$, $\vec{f}$ étant l'un quelconque des éléments de la classe $\vec{f}^{\bullet}$. On démontre alors sans difficulté que les théorèmes 40, 41, 42, 43, 45, 46, 47, s'étendent lorsque certains des exposants sont égaux à $+\infty$.

Le théorème 44 disparaît, mais il est à peu près remplacé par la définition de $\mathcal{L}^{\infty}(\vec{F})$. Le corollaire du théorème 43 peut beaucoup s'améliorer : si $\vec{f}_n$ est une suite de fonctions de $\mathcal{L}^{\infty}(\vec{F})$ telles que $N_\infty(\vec{f} - \vec{f}_n)$ tende vers 0 pour n infini, alors les $\vec{f}_n$ convergent presque partout vers $\vec{f}$ (il est inutile d'extraire une suite partielle:. Soit en effet B_n l'ensemble des x tels que $\|\vec{f}(x) - \vec{f}_n(x)\| > N_\infty(\vec{f} - \vec{f}_n)$; c'est un ensemble de mesure nulle. La réunion B des B_n est encore de mesure nulle; sur le complémentaire $A = \complement B$, les $\vec{f}_n$ convergent simplement, et même uniformément, vers $\vec{f}$, puisqu'on a $\sup\limits_{x \in A} \|\vec{f}(x) - \vec{f}_n(x)\| = N_\mu(\vec{f} - \vec{f}_n)$, qui tend vers 0 pour n infini. La convergence dans $\mathcal{L}^{\infty}(\vec{F})$ est, "à un ensemble de mesure nulle près", la convergence uniforme des fonctions.

Il en résulte que le théorème 49 ne subsiste pas. Prenons en effet $X = \mathbb{R}$, $\mu = dx$. Si $\vec{g}$ est une fonction continue, $N_\infty(\vec{g})$ coïncide avec $|||\vec{g}|||$; car l'ensemble des x pour lesquels $\|\vec{g}(x)\| > N_\infty(\vec{g})$ est de mesure nulle, mais ouvert, donc vide, et $N_\infty(\vec{g})$ est la vraie borne supérieure de $\|\vec{g}\|$, c'est-à-dire $|||\vec{g}|||$. Si alors $\vec{f}$ est une fonction de $\mathcal{L}^{\infty}(\vec{F})$, limite d'une suite de fonctions continues $\vec{f}_n$, $N_\infty(\vec{f}_m - \vec{f}_n)$ tend vers 0 pour m et n infinis, donc aussi $|||\vec{f}_m - \vec{f}_n|||$; les $\vec{f}_n$ forment une suite de Cauchy dans $(\vec{F}^{\mathbb{R}})_{c\ell}$, et comme cet espace est complet (corollaire 2 du théorème 65 du chapitre II), elles ont une limite uniforme $\vec{g}$, continue et bornee. Alors $|||\vec{f}_n - \vec{g}|||$ converge vers 0, donc aussi $N_\infty(\vec{f}_n - \vec{g})$; alors $N_\infty(\vec{f} - \vec{g}) = 0$, $\vec{f}$ et $\vec{g}$ sont presque partout égales. Ainsi seules peuvent être limites de fonctions continues les fonctions de $\mathcal{L}^{\infty}(\vec{F})$ qui sont presque partout égales à des fonctions continues; la fonction caractéristique d'un intervalle $[a, b]$ n'a pas cette propriété.

Théorème 50. - Soit $\vec{f}$ une fonction mesurable sur X , à valeurs dans $\vec{F}$. L'ensemble des p de $[1, +\infty]$ ~~pour lesquels~~ $\vec{f} \in \mathcal{L}^p(\vec{F})$ est un intervalle $|p_1 , p_2|$, ouvert, semi-ouvert ou fermé. La fonction $p \longrightarrow N_p(\vec{f})$ est une fonction continue sur $[p_1 , p_2]$ à valeurs $\geqslant 0$ dans $\overline{\mathbb{R}}$; si en outre $\| \mu \| = 1$, elle est croissante, et même strictement croissante, sauf si $\| \vec{f} \|$ est presque partout égale à une même constante c , auquel cas elle est constante et égale 4 c . De toute façon, la fonction $\xi \longrightarrow \log N_{\frac{1}{\xi}}(\vec{f})$ est toujours convexe sur $\left| \frac{1}{p_2} , \frac{1}{p_1} \right|$.

Démonstration - Le fait que l'ensemble des p pour lesquels $\vec{f} \in \mathcal{L}^p(\vec{F})$ soit un intervalle3 résulte du corollaire 4 du théorème 46, étendu aux exposants infinis.

Montrons la continuité; on peut évidemment se borner au cas où f est réelle $\geqslant 0$. Soit d'abord p_0 un point intérieur à l'intervalle; soient p_1' et p_2' , tels que $p_1 < p_1' < p_0 < p_2' < p_2$. Lorsque p tend vers p_0 , f^p converge simplement vers f^{p_0} ; en vertu de la note (*) de la page 514 , on, a, pour $p_1' \leqslant p \leqslant p_2'$, la majoration $f^p \leqslant f^{p_1} + f^{p_2'}$, le second membre étant indépendant de $p \geqslant 0$ et intégrable. Alors le théorème 35 de Lebesgue dit que $\int f^p$ converge vers $\int f^{p_0}$; donc $N_p(f) = \left(\int f^p \right)^{\frac{1}{p}}$ converge vers $N_{p_0}(f) = \left(\int f^{p_0} \right)^{\frac{1}{p_0}}$, et la continuité en p_0 est prouvke. On pouvait aussi utiliser le théorème 7 bis

Soit maintenant $p_0 = p_1$. Si $p_1 \in |p_1 , p_2|$, on prendra p_2' tel que $p_1 < p_2' < p_2$, et, en remplaçant $[p_1' , p_2']$ par $[p_1 , p_2']$, on fera le même raisonnement que précédemment et la continuité sera encore prouvée. Le seul cas nouveau à voir est donc celui où $p_1 \notin |p_1 , p_2|$ c'est-a-dire où $N_{p_1}(f) = +\infty$; on doit alors montrer que $N_p(f)$ tend vers $+\infty$ lorsque p tend *vers* p_1 par valeurs $> p_1$. Soit donné $A > 0$.
Puisque $N_{p_1}(f) = +\infty$, il existe $M \geqslant 1$ et un compact K de X tel que $\int (f_{M,K})^{p_1} \geqslant 2A$,

Alors, lorsque? tend vers p_1, $(f_{M,K})^p$ converge vers $(f_{M,K})^{p_1}$, en gardant son support dans K et en restant bornée par $M^{p'_2}$ dès que $p'_1 \leqslant p \leqslant p'_2$. Donc, d'après le théorème de Lebesgue particulier (théorème 34), $\int (f_{M,K})^p$ tend vers $\int (f_{M,K})^{p_1}$ pour p tendant vers p_1. Donc il existe $\eta > 0$ tel que $p \leqslant p_1 + \eta$ entraîne $\int (f_{M,K})^p \geqslant A$, et a fortiori $\int f^p \geqslant A$, et la continuité en p_1 est complètement démontrée. La continuité en p_2 se démontre de la même manière si $p_2 < +\infty$.

Soit donc $p_2 = +\infty$, et posons $M = N_\infty(f) \leqslant +\infty$. Nous devons montrer que $N_p(f)$ tend vers M quand p tend vers $+\infty$. Soit donné $M' < M$; soit M'_1 tel que $M' < M'_1 < M$. Il existe un ensemble Y de mesure > 0 sur lequel $f(x) \geqslant M'_1$; alors $N_p(f) \geqslant M'_1 (\mu(Y))^{\frac{1}{p}}$; pour p tendant vers $+\infty$, cette dernière quantité tend vers M'_1 ; donc il existe q' telle que $p \geqslant q'$ entraîne $N_p(f) \geqslant M'$.

Ceci suffit à montrer la continuité en $p_2 = +\infty$, si $N_\infty(f) = +\infty$. Si $M = N_\infty(f) < +\infty$, il faut encore une inégalité en sens inverse. Soit alors $M'' > M$. Dans l'inégalité (IV,4;61 quarto), appliquée à p_1, ∞, p, au lieu de p, q, r, on a $\alpha = \frac{\frac{1}{p}}{\frac{1}{p_1}}$, $\beta = \frac{\frac{1}{p_1}-\frac{1}{p}}{\frac{1}{p_1}}$, on voit que α tend vers 0 et β vers 1 pour p infini ; alors le second membre correspondant de (IV,4;61 quarto) tend vers $(N_{p_1}(f))^0 (N_\infty(f))^1 = M$.

Donc il existe un entier q'' tel que $p \geqslant q''$ entraîne $N_p(f) \leqslant M''$; alors, si $q = \sup(q', q'')$, $p \geqslant q$ entraîne $M' \leqslant N_p(f) \leqslant M''$, et la continuité en p_2 est complètement démontrée.

Si $\|\mu\| = 1$, l'inégalité (IV,4;61) montre que la fonction $p \longrightarrow N_p(f)$ est croissante.

Nous admettrons qu'elle est strictement croissante, si f n'est pas presque partout égale à une constante.

Prolongement d'une mesure non ≥ 0

Nous venons de voir, dans ce paragraphe, que, pour une mesure de Radon μ réelle ≥ 0 , on pouvait prolonger la fonction $\varphi \longrightarrow \mu(\varphi)$, initialement définie sur $\mathcal{C}(X)$, en une fonction $f \longrightarrow \int f\, d\mu$ définie sur un espace vectoriel beaucoup plus grand, $\mathcal{L}^1(X,\mu)$; et même nous avons défini cette intégrale pour des fonctions vectorielles $\vec{f} \in \mathcal{L}^1(X,\mu\,;\vec{F})$.

Qu'en est-il pour une mesure μ réelle non ≥ 0 , ou complexe, ou vectorielle ?

Il existe aussi une théorie du prolongement, mais nous ne pouvons pas en parler actuellement. Nous allons toutefois étudier un prolongement partiel, au moins lorsque la mesure μ est à valeurs dans un espace vectoriel de dimension finie.

Appelons $\Gamma(X)$ l'espace vectoriel des fonctions à valeurs complexes, boréliennes, bornées à support compact sur X. $\mathcal{C}(X)$ est un sous-espace vectoriel de $\Gamma(X)$. Nous ne mettrons pas une topologie sur $\Gamma(X)$, mais une notion de convergence que nous appellerons la L - convergence (L étant l'initiale de Lebesgue) (dont on peut effectivement montrer qu'elle ne correspond pas à la convergence dans un espace topologique) :

<u>une suite $f_0, f_1, \ldots, f_n, \ldots$ de $\Gamma(X)$ est dite L - convergente vers f, si les f_n convergent vers f, pour n tendant vers l'infini, tout en restant bornées en module dans leur ensemble, et en gardant leur support dans un compact K de X</u>

Soit alors μ une mesure de Radon ≥ 0. Son prolongement de Lebesgue permet de définir une fonction $f \longrightarrow \int f\, d\mu$, définie sur $\Gamma(X)$, à valeurs complexes, et ayant les propriétés suivantes :

1°/ Cette fonction est linéaire ;

2°/ Si la suite des f_n est L--convergente vers f , alors $\int f_n\, d\mu$ converge vers $\int f\, d\mu$;

3°/ $\mathcal{O}$ étant un ouvert de X d'adhérence compacte, et $\chi_{\mathcal{O}}$ étant sa fonction caractéristique, alors, quel que soit $\varepsilon > 0$, il existe un compact $K \subset \mathcal{O}$ tel que, pour toute fonction φ de $\mathcal{C}(X)$, $0 \leq \varphi \leq 1$, égale à 1 sur un voisinage de K, et de support dans $\mathcal{O}$, on ait

$$\left| \int \chi_{\mathcal{O}} \, d\mu - \int \varphi \, d\mu \right| \leq \varepsilon \;;$$

4°/ la fonction est positive, en ce sens que, si $f \geq 0$, alors $\int f \, d\mu \geq 0$.

Si $\vec{E}$ est un Banach, toute application de $\Gamma(X)$ dans $\vec{E}$ satisfaisant aux 3 premières conditions sera dite application linéaire L-continue de $\Gamma(X)$ dans $\vec{E}$; si $\vec{E} = \mathbb{C}$, et si la 4ème condition est vérifiée, elle sera dite ≥ 0. Le prolongement de Lebesque d'une mesure de Radon $\mu \geq 0$ définit donc une forme linéaire L-continue ≥ 0 sur $\Gamma(X)$.

La condition 3°/ est équivalente a une condition analogue 3bis/: quel que soit le compact K de X, et $\varepsilon > 0$, il existe un ouvert $\mathcal{O} \supset K$ tel que, pour toute $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq 1$, égale à 1 sur un voisinage de K et de support dans $\mathcal{O}$, on ait $\left| \int \chi_K \, d\mu - \mu(\varphi) \right| \leq \varepsilon$. Si en effet on suppose 2°/, et si K et ε sont donnés, on choisit $\mathcal{U}$ voisinage ouvert d'adhérence compacte de K, et on trouve le compact H de l'ouvert $\mathcal{U} - K$ dont 2°/ affirme l'existence; alors l'ouvert $\mathcal{U} - H$ satisfait à 2 bis) pour le compact K. Et de même en sens inverse.

Si X est métrisable, toute application de $\Gamma(X)$ dans $\vec{E}$ satisfaisant a la condition 2°/, satisfait automatiquement a 3°/. Considérons en effet la suite des compacts $K_n = \{x \in \mathcal{O} \,;\, d(x, \complement\mathcal{O}) \geq \frac{1}{n}\}$. Pour $\varepsilon > 0$ donné, K_n satisfait à 3°/ pour n assez grand. Si en effet c'était inexact, on pourrait, pour tout n trouver $\varphi_n \in \mathcal{C}(X)$, $0 \leq \varphi_n \leq 1$, égale a 1 sur un voisinage de K_n et de support dans $\mathcal{O}$, telle que cependant $\left| \int \chi_{\mathcal{O}} \, d\mu - \int \varphi_n \, d\mu \right| > \varepsilon$; or la suite des φ_n converge vers $\chi_{\mathcal{O}}$, et ce serait contraire à l'hypothèse de L-continuité. Mais, si X n'est pas métrisable, la condition 3°/, qui utilise une sorte de limite d'une famille non dénombrable, ne saurait se réduire à 2°/, valable seulement pour des suites. Si nous avons introduit cette condition 3°/ ce n'est pas pour le plaisir de nous compliquer la vie : elle sera indispendable au théorème d'unicité 50 bis.

Pour toute application linéaire L-continue $\vec{\mu}$ de $\Gamma(X)$ dans $\vec{E}$, il sera commode de noter $\vec{\mu}(A)$ la valeur de $\vec{\mu}$ sur la fonction caractéristique d'une partie A de X, borélienne, d'adhérence compacte.

Théorème **50** bis - Soit $\vec{E}$ un espace de Banach. Toute application de $\Gamma(X)$ dans $\vec{E}$, linéaire et L-continue nulle sur le sous-espace vectoriel $\mathcal{C}(X)$ est identiquement nulle.

Démonstration - Il résulte de la condition de continuité 3°/ que cette fonction est nécessairement nulle sur les fonctions caractéristiques d'ouverts d'adhérence compacte. Nous aurons alors besoin d'un lemme :

Lemme 1 - Soit $\mathcal{U}$ un espace topologique et soit $\mathcal{F}$ le plus petit ensemble de parties de $\mathcal{U}$ ayant les propriétés suivantes:

1°/ Stabilité par différence :

Si $A \in \mathcal{F}, B \in \mathcal{F}, B \subset A$, alors $A - B \in \mathcal{F}$;

2°/ Stabilité par limite croissante :

Si les A_n forment une suite croissante de parties appartenant à $\mathcal{F}$, alors leur réunion appartient à $\mathcal{F}$;

3°/ Stabilité par limite décroissante :

Si les A_n forment une suite décroissante de parties appartenant à $\mathcal{F}$, leur intersection appartient à $\mathcal{F}$. Si alors $\mathcal{F}$ contient toutes les parties ouvertes de $\mathcal{U}$, $\mathcal{F}$ est la tribu des parties boréliennes de $\mathcal{U}$.

Démontrons ce lemme.

Remarquons que les propriétés précédentes ne permettent pas à priori d'affirmer que $\mathcal{F}$ est une tribu. Nous ne savons pas que l'intersection $A \cap B$ de deux parties de $\mathcal{F}$ appartient à $\mathcal{F}$. Si nous le démontrons, alors nous saurons que $\mathcal{F}$ est une tribu; en effet la stabilité par rapport à l'intersection de deux parties entraîne la stabilité par rapport à l'intersection d'un nombre fini de parties, puis la stabilité pour les intersections dénombrables, en vertu de la condition sur les suites décroissantes; d'autre part, la stabilité par différence entraîne que le complémentaire d'une partie de $\mathcal{F}$ appartient à $\mathcal{F}$, puisque $\mathcal{U}$ étant ouvert, appartient à $\mathcal{F}$. alors il y aura stabilité aussi Pour toutes les réunions dénombrables, et $\mathcal{F}$ sera bien une tribu.

Considérons alors, pour une partie ouverte $\mathcal{O}$ de $\mathcal{U}$, l'ensemble $\mathcal{F}_{\mathcal{O}}$ de toutes les parties A de $\mathcal{U}$ telles que $A \cap \mathcal{O}$ appartienne à $\mathcal{F}$. C'est évidemment un ensemble de parties, stable par différence et pour le.5 réunions de suites croissantes et les intersections de suites décroissantes; d'autre Part il contient tous les ouverts, puisque $\mathcal{F}$ contient tous les ouverts et que l'intersection de deux ouverts est un ouvert. Comme $\mathcal{F}$ est supposé être le plus petit ensemble de parties de $\mathcal{U}$ ayant les 3 propriétés indiquées, l'ensemble $\mathcal{F}_{\mathcal{O}}$ contient $\mathcal{F}$. Autrement dit, pour toute partie A de $\mathcal{F}$ et tout ouvert $\mathcal{O}$ de $\mathcal{U}$ $A \cap \mathcal{O}$ appartient à $\mathcal{F}$. Soit alors A une partie de $\mathcal{F}$, et considérons l'ensemble $\mathcal{F}_A$ des parties B de $\mathcal{U}$ telles que $A \cap B$ appartiennent à $\mathcal{F}$. C'est encore une famille de parties stable pour les différences, les réunions de suites croissantes et les Intersections de suites décroissantes. D'après ce que nous venons de voir, elle contient toutes les parties ouvertes de $\mathcal{U}$; nécessairement elle contient toute la famille $\mathcal{F}$. On voit donc bien que, pour A et B dans $\mathcal{F}$, $A \cap B$ est dans $\mathcal{F}$, ce qui démontre que $\mathcal{F}$ est une tribu. Comme alors $\mathcal{F}$ contient tous les ouverts, $\mathcal{F}$ contient la tribu borélienne; mais la tribu borélienne a les 3 stabilités requises, et $\mathcal{F}$ est le plus petit ensemble de parties ayant cette propriété. Donc $\mathcal{F}$ est la tribu borélienne, et le lemme 1 est démontré.

Continuons maintenant la démonstration du théorème. L'ensemble $\mathcal{M}$ des parties boréliennes A, contenues dans un ouvert $\mathcal{U}$ d'adhérence compacte de X pour lesquelles $\vec{\mu}(A) = 0$, est évidemment stable par différence d'après la linéarité de $\vec{\mu}$, et stable pour les réunions de suites croissantes et les intersections de suites décroissantes d'après la L-continuité de $\vec{\mu}$. D'autre part, d'après ce que nous venons de démontrer avant le lemme 1, $\mathcal{M}$ contient toutes les parties ouvertes de $\mathcal{U}$ Le lemme 1 montre alors que $\mathcal{M}$ contient toutes les parties boréiiennes de $\mathcal{U}$. Mais nous devons faire attention à une petite difficulté : il n'est pas sûr qu'une partie A de $\mathcal{U}$, bordlienne relativement a $\mathcal{U}$ soit la même chose qu'une partie A de $\mathcal{U}$, borélienne relativement à X. Nous démontrerons à ce sujet un deuxième lemme, qui étend aux parties boréliennes les théorèmes 5 et 6 du chapitre II.

Lemme 2 - <u>Soit $\mathcal{U}$ une partie de X. Une partie A de $\mathcal{U}$, borélienne dans X, est borelienne dans $\mathcal{U}$, et réciproquement, si $\mathcal{U}$ est borelienne dans X</u>.

Considérons en effet toutes les parties B de X dont les intersections avec $\mathcal{U}$ et $\complement\,\mathcal{U}$ sont boréliennes dans les espaces topologiques $\mathcal{U}$ et $\complement\,\mathcal{U}$. On voit aussitôt qu'elle forment une tribu de parties de X. Comme cette tribu contient tous les ouverts de X (puisque ceux-ci coupent $\mathcal{U}$ suivant des les ouverts de $\mathcal{U}$), elle contient nécessairement toutes parties boréliennes de X. il en résulte que toute partie B de X, borélienne dans X et contenue dans $\mathcal{U}$, est borélienne dans $\mathcal{U}$.

Considérons maintenant l'ensemble des parties C de $\mathcal{U}$ qui, considérées dans X, sont boréliennes; on voit immédiatement que cet ensemble est une tribu de parties de $\mathcal{U}$, si $\mathcal{U}$ est borélienne dans X (toutes les propriétés exigées d'une tribu sont évidemment vérifiées, Sauf peut être celle relative au complémentaire. Mais, si $A \subset \mathcal{U}$ est borélienne dans X, $\complement_{\mathcal{U}} A = \mathcal{U} - A = X - \complement\mathcal{U} - A$ est borélienne dans X si $\complement\mathcal{U}$ est borélienne dans X, donc si $\mathcal{U}$ est borélienne dans X). Comme alors il contient tous les ouverts de $\mathcal{U}$ (car un ouvert de $\mathcal{U}$ est son intersection avec un ouvert de X, donc est borélien dans X puisque $\mathcal{U}$ et un ouvert sont boréliens dans X), il contient toutes les parties boréliennes de $\mathcal{U}$; donc, si C est une partie de $\mathcal{U}$, borélienne dans $\mathcal{U}$, elle est borélienne dans X, ce qui démontre le lemme 2.

A la suite de ce deuxième lemme, nous savons maintenant que la fonction $\vec{\mu}$ est nulle sur toutes les parties boréliennes de X d'adhérence compacte (c'est-à-dire sur toutes leurs fonctions caractéristiques). Tout sera alors démontré par un dernier lemme :

<u>Lemme 3</u> - <u>Une application linéaire L - continue de $\Gamma(X)$ dans $\vec{E}$, nulle sur les fonctions caractéristiques d'ensembles, est nulle; une forme linéaire ($\vec{E} = \mathbb{C}$) L - continue sur $\Gamma(X)$, ≥ 0 sur les fonctions caractéristiques d'ensembles est ≥ 0</u>.

La fonction μ est nulle (resp. ≥ 0), sur toutes les fonctions boréliennes étagées (resp. étagées ≥ 0) à support compact de $\Gamma(X)$.

Comme alors, d'après la remarque qui suit le théorème 23bis, toute fonction borélienne sur X à support compact est limite uniforme d'une suite de fonctions boréliennes étagées, bornées dans leur ensemble, à support dans un compact fixe, ≥ 0 si elle est ≥ 0, l'application de la L -continuité montre que $\vec{\mu}$ est nulle sur $\Gamma(X)$ tout entier. (resp. ≥ 0 sur $\Gamma^+(X)$).

<u>Corollaire 1</u> - <u>Si $\vec{\mu}_1$ et $\vec{\mu}_2$ sont deux applications linéaires L - continues de $\Gamma(X)$ dans $\vec{E}$, qui coïncident sur $\mathcal{C}(X)$ ou sur les fonctions caractéristiques d'ensembles, elles coïncident.</u>

Corollaire 2 - Si μ est une forme linéaire **L-continue** sur $\Gamma(X)$, si d'autre part elle est Positive sur $\mathscr{C}_+(X)$ ou sur les fonctions caractéristiques d'ensembles, elle est positive sur $\Gamma_+(X)$, et elle est le prolongement de Lebesgue à $\Gamma(X)$ d'une mesure de Radon ≥ 0.

La positivité sur les fonctions caractéristiques entraîne, d'après le **lemme** 3, la positivité sur $\Gamma_+(X)$ donc sur (2+(X); on a donc, dans les 2 hypothèses, positivité Sur $\mathscr{C}_+(X)$.

Alors, considérée comme forme linéaire sur $\mathscr{C}(X)$, il résulte de la L-continuité qu'elle est continue sur chaque $\mathscr{C}_K(X)$, donc que c'est une mesure de Radon, et par conséquent une mesure de Radon ≥ 0. Alors cette forme linéaire et le prolongement de Lebesgue de cette mesure positive sont deux formes linéaires L- continues sur $\Gamma(X)$, qui cofncldent sur $\mathscr{C}(X)$ donc coïncident sur $\Gamma(X)$.

Soit alors $\vec{\mu}$ une mesure de Radon sur X à valeurs dans un espace de Banach $\vec{E}$. On dit que $\vec{\mu}$ admet un prolongement borélien, s'il existe une application linéaire L-continue de $\Gamma(X)$ dans $\vec{E}$, coïncidant avec $\vec{\mu}$ sur $\mathscr{C}(X)$. 11 résulte du théorème 50bis que, si ce prolongement existe, il est unique.

Il est visible que ce prolongement n'existe pas toujours. En effet il existe au moins une condition nécessaire de prolongement : si une suite φ_n de fonctions de $\mathscr{C}(X)$ L-converge vers 0, **la suite ries** $\vec{\mu}(\varphi_n)$ doit converger vers 0 dans $\vec{E}$. Prenons par exemple 'la "mesure identique" définie Page 456, qui n'admettait pas de majorante absolue : X compact, $\vec{E}=\mathscr{C}(X)$, $\mu(\varphi)=\varphi\ [\varphi\in\mathscr{C}(X),\ \vec{\mu}(\varphi)\in\vec{E}=\mathscr{C}(X)]$. Alors la L-convergence des φ_n vers 0 n'entraîne évidemment Pas la convergence des $\vec{\mu}(\varphi_n)$ dans $\vec{E}$, c'est-a-dire la convergence uniforme des φ_n (Exemple: $X=[0,1]$; $\varphi_n(x)=0$ pour $x=0$ et $x\geq\frac{1}{n}=1$ pour $x=\frac{1}{2n}$, et affine dans chaque intervalle $\left[0,\frac{1}{2n}\right]$, $\left[\frac{1}{2n},\frac{1}{n}\right]$).

Théorème 50 ter - L'ensemble& $(X;\vec{E})$ des mesures de Radon définie sur **X**, à valeurs dans l'espace de Banach $\vec{E}$, admettant un prolongement borélien, est un espace vectoriel, et l'application $(\vec{\mu}, f) \longrightarrow \int f \, d\vec{\mu}$ est une application bilinéaire de $\mathcal{M}(X;\vec{E}) \times \Gamma(X)$ dans $\vec{E}$. Si $\vec{E}$ est de dimension finie toute mesure à valeurs dans $\vec{E}$ admet un prolongement borélien.

Démonstration - Si $\vec{\mu}$ et $\vec{\nu}$ sont deux mesures admettant un prolongement borélien, alors $\vec{\mu} + \vec{\nu}$ admet un prolongement borélien, ainsi que $k\vec{\mu}$ pour tout scalaire k. En effet il suffit précisément de définir ce prolongement par les formules

$$\int f \, d(\vec{\mu}+\vec{\nu}) = \int f \, d\vec{\mu} + \int f \, d\vec{\nu}$$

$$\int f \, d(k\vec{\mu}) = k \int f \, d\vec{\mu} \, ;$$

on définit la en effet des applications linéaires L-continues de $\Gamma(X)$ dans $\vec{E}$, qui coïncident bien respectivement avec $\vec{\mu} + \vec{\nu}$ et $k\vec{\mu}$ sur $\mathcal{C}(X)$. Cela prouve à la fois que $\mathcal{M}(X;\vec{E})$ est un espace vectoriel, et que l'application est bilinéaire. $\mathcal{M}(X;\mathbb{C})$ contient les mesures de Radon ≥ 0, grâce à la théorie de Lebesgue, donc les mesures réelles par différence (théorème 18), donc les mesures complexes ($\mu = \mu_1 + i\mu_2$, μ_1 et μ_2 reelles). Alors, si μ est une mesure scalaire, $\vec{e}$ un élément de E, la mesure $\mu\vec{e}$, définie par $(\mu\vec{e})(\varphi) = \vec{e}\mu(\varphi)$, admettra le prolongement défini par $\int f \, d(\mu\vec{e}) = \vec{e} \int f \, d\mu$

Si donc $\vec{E}$ est de dimension finie, et si $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, en est une base, toute mesure $\vec{\mu}$ a valeurs dans $\vec{E}$ s'écrit $\vec{\mu} = \sum_{i=1}^{n} \mu_i \vec{e}_i$ donc est prolongeable.

Théorème 50 quarto - Soit μ mesure réelle sur X. Les formules démontrées pour $\mu^+, \mu^-, |\mu|$, au théorème **18**, subsistent pour les prolongements boréliens : on a

$$(\mathrm{IV},4;63) \quad \left\{ \begin{aligned} \int f \, d\mu^+ &= \sup_{\substack{g \in \Gamma(X) \\ 0 \leq g \leq f}} \int g \, d\mu \, , \text{ pour } f \in \Gamma(X), f \geq 0 \, . \\ \int f \, d|\mu| &= \sup_{\substack{g \in \Gamma(X) \\ |g| \leq f}} \int g \, d\mu \, , \text{ pour } f \in \Gamma(X), f \geq 0 \, . \end{aligned} \right.$$

Pour toute partie boréllenne A de X , d'adhérence compacte, on a

$$\mu^+(A) = \operatorname*{Sup}_{B \subset A} \mu(B) \; ; \; \mu^-(A) = \operatorname*{Sup}_{B \subset A} (-\mu(B)) \tag{IV,4;64}$$

Démonstration - Dans tous les cas, il est bien évident que le premier membre est au moins égal au deuxième. Il faut donc démontrer des inégalités en sens inverse. Tout d'abord, si $\mathcal{O}$ est un ouvert d'adhérence compacte il existe une fonction $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq 1$, de support dans $\mathcal{O}$, telle que $\mu^+(\varphi) \geq \mu^+(\mathcal{O}) - \frac{\varepsilon}{2}$. Il existe alors, d'après la formule (IV,2;41), une fonction $\Psi \in \mathcal{C}(X), 0 \leq \Psi \leq \varphi \leq \chi_{\mathcal{O}}$ telle que $\mu(\Psi) \geq \mu^+(\varphi) - \frac{\varepsilon}{2}$. On a donc bien $\mu(\Psi) \geq \mu^+(\mathcal{O}) - \varepsilon$, ce qui montre la première formule dans le cas où f est la fonction caractéristique d'un ouvert (et alors on voit qu'on peut même se borner aux g de $\mathcal{C}(X)$ * .

Soit maintenant A un ensemble borélien quelconque d'adhérence compacte, de fonction caractéristique χ_A . soit $\mathcal{O}$ un ouvert tel que $|\mu|(\mathcal{O} - A) \leq \frac{\varepsilon}{2}$. Soit Ψ une fonction déterminée comme précédemment, relativement à l'ouvert $\mathcal{O}$ telle que $\mu(\Psi) \geq \mu^+(\mathcal{O}) - \frac{\varepsilon}{2}$ donc $\geq \mu^+(A) - \frac{\varepsilon}{2}$. Si alors on considère la fonction (discontinue) $g = \Psi\chi_A$, $0 \leq g \leq \chi_A$ l'inégalité

$$\int \Psi\chi_A \, d\mu = \mu(\Psi) - \int \Psi\chi_{\mathcal{O}-A} \, d\mu \geq \mu(\Psi) - |\mu|(\mathcal{O}-A) \geq \mu(\Psi) - \frac{\varepsilon}{2} \geq \mu^+(A) - \varepsilon ,$$

prouve encore la première formule si f est la fonction caractéristique de A . Par combinaisons linéaires, la première formule est donc aussi vraie pour toute fonction f étagée borélienne ≥ 0 à support compact.

Soit enfin f une fonction ≥ 0 quelconque de $\Gamma(X)$ soit M sa borne supérieure, K son support. Elle est limite uniforme d'une suite de fonctions étagées boréliennes ≥ 0 bornées par M , à support dans K . On peut donc quel que soit $\varepsilon > 0$, déterminer une fonction h étagée borélienne ≥ 0, majorée par M , à support dans K, telle que $|f - h| \leq \delta = \frac{\varepsilon}{4\mu^+(K)}$

* si f est semi-continue inférieurement, on peut se borner à des g continues, mais pas si f est arbitraire, même pour μ déjà ≥ 0 (voir remarque après théorème 39 bis).

SI nous posons $k = (h-\delta)^+$, on a certainement $0 \leq k \leq f$, et $f - k \leq 2\delta$. On a donc $\int k\, d\mu^+ \geq \int f d\mu^+ - 2\delta\mu^+(K) = \int f d\mu^+ - \frac{\varepsilon}{2}$.

Comme alors nous venons de voir qu'il existe une fonction $g \in \Gamma(X)$, $0 \leq g \leq k$ donc $\leq f$, vérifiant $\int g\, d\mu \geq \int k\, d\mu^+ - \frac{\varepsilon}{2}$ donc $\geq \int f\, d\mu^+ - \varepsilon$, la 1ère formule (IV,4;63) est complètement démontrée. La 2ème formule (IV,4;63) se démontre de la même manière. Montrons maintenant (IV,4;64).

Soit A une partie borélienne de X, d'adhérence compacte. soit $\varepsilon > 0$ donné. Soient $\delta, \eta > 0$ arbitraires; nous les particulariserons plus loin en fonctions de ε. D'après (IV,4;63) il existe une fonction $g \in \Gamma(X)$, $0 \leq g \leq \chi_A$, telle que $\mu^+(A) \geq \int g\, d\mu \geq \mu^+(A) - \delta$. On en déduit donc aussitôt $\mu^+(A) \geq \int g\, d\mu^+ \geq \mu^+(A) - \delta$ donc $\int g\, d\mu^- = \int g\, d\mu^+ - \int g\, d\mu \leq \delta$.

Appelons B l'ensemble boréllen des points x où $g(x) \geq \eta$. La fonction $\eta\chi_B$ est comprise entre 0 et g ; on a donc nécessairement $\eta\,\mu^-(B) = \int \eta\chi_B\, d\mu^- \leq \int g\, d\mu^- \leq \delta$ donc $\mu^-(B) \leq \frac{\delta}{\eta}$. SI au contraire nous considérons la fonction k égale à 1 sur B, à η sur $A - B$, à 0 sur $\complement A$, elle est comprise entre g et χ_A, on a donc nécessairement $\mu^+(B) + \eta\,\mu^+(A-B) = \int k\, d\mu^+ \geq \int g\, d\mu^+ \geq \mu^+(A) - \delta$ donc $\mu^+(B) \geq \mu^+(A) - \delta - \eta\,\mu^+(A-B) \geq \mu^+(A) - \delta - \eta\,\mu^+(A)$

On a finalement

$$\mu(B) = \mu^+(B) - \mu^-(B) \geq (\mu^+(A) - \delta - \eta\,\mu^+(A)) - \frac{\delta}{\eta}.$$

B dépend du choix de δ et η, mais si nous avons choisi δ, η, tels que $\delta + \eta\,\mu^+(A) + \frac{\delta}{\eta} \leq \varepsilon$, on aura un $B \subset A$ borélien tel que $\mu(B) \geq \mu^+(A) - \varepsilon$, et la formule (IV,4;64) sera démontrée. Or un tel choix est possible; choisissons d'abord $\eta = \frac{\varepsilon}{2\mu^+(A)}$, donc $\eta\,\mu^+(A) = \frac{\varepsilon}{2}$, et ensuite δ tel que $\delta(1 + \frac{1}{\eta}) = \frac{\varepsilon}{2}$. ou $\delta = \frac{\varepsilon}{2(1+\frac{1}{\eta})}$.

Ainsi le théorème 50 quarto est démontré.

Il est alors utile d'introduire la notion suivante

soit μ une mesure réelle. On dit qu'elle est complètement positive (resp. complètement négative; complètement nulle) sur une partie A de X , si, pour toute partie borélienne B de A d'adhérence compacte, on a $\mu(B) \geqslant 0$ (resp $\leqslant 0$; $= 0$).

On écrira $\mu(A) \gg 0$ (resp. $\mu(A) \ll 0$) si μ est complètement positive (resp.négative) sur A si μ est complètement positive sur A , elle l'est a fortiori sur toute partie de A • si μ est complètement positive sur une infinité dénombrable de parties boréllennes elle l'est sur leur réunion **A** (si en effet $B \subset \bigcup_{n=0}^{\infty} A_n$, on a $B = \bigcup_{n=0}^{\infty} B_n$, où les **B** sont disjointes boréliennes et $B_n \subset A_n$: $B_0 = B \cap A_0$, $B_1 = B \cap A_1 \cap \complement A_0, \dots B_n A_n \cap \complement A_{n+1} \dots \cap \complement A_0$; alors $\mu(B) = \sum_{n=0}^{\infty} \mu(B_n) \geq 0$.

Pour que $\mu(A) \gg 0$,il faut et il suffit que **A** soit de mesure intérieure nulle pour μ^- ; alors, pour toute partie $B \subset$ **A** boréllenne d'adhérence compacte, on a $\mu(B) = \mu^+(B)$, $\mu^-(B) = 0$ (Supposons en effet **A** de mesure intérieure nulle pour μ^- . Pour tout B borélien $\subset A$ d'adhérence compacte, on aura $\mu^-(B) = 0$, donc $\mu(B) = \mu^+(B) \geqslant 0$, et μ sera complètement positive sur **A** . Inversement supposons $\mu(A) \gg 0$. Alors $\mu(K) \geqslant 0$ pour toute partie compacte $K \subset A$. La formule (IV,4;64) appliquée à K et à μ^- donne $\mu^-(K) = \operatorname{Sup}_{B \subset K} (-\mu(B)) = 0$; donc $\mu^-_*(A) = 0$.

Théorème 50 quinto - Si μ est une mesure réelle sur X , espace localement compact dénombrable à l'infini, on peut décomposer X en réunion de deux ensembles boréliens complémentaire X^+ et X^- , tels que μ soit complètement positive sur X^+ et complètement négative sur X^-.

Démonstration - Soit d'abord K un compact de X . On a vu au théorème précédent que $\mu^+(K) = \operatorname{Sup}_{A \in K} \mu(A)$; il est donc possible de déterminer une suite de parties boréllennes de X contenues dans **K** : $A_0, A_1, \dots, A_n, \dots$ telles que l'on ait $\mu(A_n) \geqslant \mu^+(K) - \frac{1}{2^n}$.

Posons alors $B_n = \bigcup_{p \geq n} A_p$. On a immédiatement les inégalités

$$\text{(IV,4;65)} \quad \mu^+(A_n) \geq \mu^+(K) - \frac{1}{2^n} \text{, donc } \mu^+(B_n) = \mu^+(K) ; \ \mu^-(A_n) = \mu^+(A_n) - \mu(A_n) \leq \frac{1}{2^n} \text{, donc } \mu^-(B_n) \leq \frac{1}{2^{n-1}} .$$

Si alors nous appelons K^+ l'intersection $\bigcap_{n \geq 0} B_n$ et K^- l'ensemble $K - K^+$, on a aussitôt les inégalitds

$$\text{(IV,4;66)} \quad \begin{aligned} \mu^+(K^+) &= \mu^+(K) , & \mu^-(K^+) &= 0 ; \\ \mu^+(K^-) &= 0 , & \mu^-(K^-) &= \mu^-(K) . \end{aligned}$$

Nous venons ainsi de faire la détermination des deux parties cherchées relativement au compact K . Si alors $K_0, K_1, \ldots, K_n, \ldots$ est une suite croissante de compacts de réunion X , on peut, pour chacun d'eux K_n , déterminer deux parties K_n^+ et K_n^- comme nous venons de le faire; il suffira de poser $X^+ = \bigcup_{n \geq 0} K_n^+$, et $X'^- = \bigcup_{n \geq 0} K_n^-$. Puisque μ est complètement positive sur K_n^+ , elle est complètement positive sur la réunion X^+, et, puisqu'elle est complètement négative sur K_n^-, elle est complètement négative sur X'^-.

En fait il n'est pas sûr que l'intersection de X^+ et de X'^- soit nulle; mais, si nous désignons par Y leur intersection, μ est, sur Y , à la fois complètement positive et complètement négative, donc complètement nulle; en remplaçant alors X^+ et X'^- par X^+ et $X^- = X'^- - Y$, on a répondu aux conditions du problème.

<u>Remarque</u> - On peut se demander quel est le degré d'indétermination du système X^+ , X^- , que nous venons de trouver. Soit Y^+, Y^- , une autre solution du même problème. L'ensemble $X^+ \cap Y^-$ des points qui appartiennent à X^+, et n'appartiennent pas à Y^+ , ou l'ensemble $X^- \cap Y^+$ de ceux qui appartiennent a Y^+ et n'appartiennent pas à X^+ , sont des parties de X sur lesquelles μ est à la fois complètement positive et complètement négative donc complètement nulle; on peut donc dire, en un certain sens, que X^+ et Y^+ diffèrent

d'un ensemble de mesure complètement nulle pour μ, c'est-à-dire de mesure nulle pour $|\mu|$; et de même X^- et Y^-. Inversement il est clair que si l'on modifie X^+ et X^- par des ensembles sur lesquels μ est complètement nulle, on ne change rien à leurs propriétés.

Corollaire - Pour que 2 mesures $\mu, \nu, \geqslant 0$, soient étrangeres (c'est-a-dire $\operatorname{Inf}(\mu, \nu) = 0$), il faut et il suffit qu'on puisse décomposer X en réunion de 2 ensembles X_μ, X_ν, tels que $\nu(X_\mu) = 0$ et $\mu(X_\nu) = 0$ et on peut alors les choisir boréliens complémentaires.

Démonstration - Si une telle décomposition est possible, on a, pour toute mesure $\lambda \geqslant 0$, majorée a la fois par μ et ν, $\lambda(X_\mu) \leqslant \nu(X_\mu) = 0$, $\lambda(X_\nu) \leqslant \mu(X_\nu) = 0$, donc $\lambda(X) = 0$ et par suite $\lambda = 0$; donc $\operatorname{Inf}(\mu, \nu)$, qui est $\geqslant 0$, est 0, et μ et ν sont étrangères. Inversement, soient μ et $\nu \geqslant 0$ étrangères. Posons $\lambda = \mu - \nu$ De $\operatorname{Inf}(\mu, \nu) = 0$, on déduit $\operatorname{Sup}(\mu, \nu) = \mu + \nu$ (voir page 456) Alors $\lambda^+ = \operatorname{Sup}(\mu - \nu, 0) = \operatorname{Sup}(\mu, \nu) - \nu = \mu + \nu - \nu = \mu$ et de même $\lambda^- = \nu$; le théorème donne donc le résultat, avec $X_\mu = X^+$ et $X_\nu = X^-$, boréliens complémentaires.

Remarque - On dit qu'une mesure μ est portée par une partie A de X, ou concentrée sur A, si $\complement A$ est de μ- mesure nulle. Cela n'a pas de rapport direct avec le support, si ce n'est que le support est le plus petit ensemble fermé pouvant porter μ. Mais A n'est pas nécessairement fermé; par exemple, si les a_ν sont les rationnels de $\mathbb{R}$, et si la série $\sum_\nu c_\nu$ converge, $c_\nu > 0$, la mesure $\sum_\nu c_\nu \delta_{(a_\nu)}$ est portée par $\mathbb{Q}$ mais son support est $\mathbb{R}$. Il n'y a d'ailleurs pas un plus petit ensemble portant μ, sauf dans le cas des mesures atomiques (page 427), car on peut toujours retirer à A n'importe quel ensemble de μ- mesure nulle; par exemple la mesure dx sur $\mathbb{R}$ est portée, si on le désire, par l'ensemble des irrationnels. Alors le corollaire précédent peut s'écrire:

Pour que μ et $\nu, \geqslant 0$, soient étrangères, il faut et il suffit qu'il existe 2 ensembles disjoints de X, portant respectivement μ et 3 : $\complement X_\mu$ porte 3 et $\complement X_\nu$ porte μ, (c'est-a-dire, si X_μ et X_ν sont complémentaires, X_μ porte μ et X_ν porte ν). Par exemple, si μ est une mesure $\geqslant 0$ atomique sur X (c'est-a-dire de la forme $\sum c_\nu \delta_{(a_\nu)}$ avec $\sum_{a_\nu \in K} c_\nu < +\infty$ pour tout compact K ; les a_ν sont en infinité dénombrable sur K , donc sur X réunion dénombrable de compacts) et si ν est une mesure diffuse (c'est-à-dire pour laquelle tout point est de mesure nulle, donc aussi tout ensemble dénombrable), μ et ν sont étrangères, puisque respectivement portées par l'ensemble des a_ν et son complémentaire.

On peut encore donner une condition équivalente, apparemment dissymétrique :

Pour que μ et ν , ≥ 0 , soient étrangères, il faut et il suffit que ν soit μ- singulière, c'est-à-dire, portée par un ensemble de μ- mesure nulle.

En effet, dire que A , de p-mesure nulle, porte ν , c'est dire que $\mu(A) = 0$ et $\nu(\complement A) = 0$.

§ 5 MULTIPLICATION D'UNE MESURE PAR UNE FONCTION

Produit d'une mesure vectorielle par une fonction continue scalaire

Soit $\vec{\mu}$ une mesure sur l'espace localement compact X , à valeurs dans un espace vectoriel normé $\vec{E}$. Soit d'autre part p une fonction continue scalaire sur X . Le produit de la mesure $\vec{\mu}$ par la fonction p , qu'on appelle la mesure $p\vec{\mu}$ ou $d(p\vec{\mu})$ ou $p\,d\vec{\mu}$, est, par définition, la mesure $\vec{\nu}$ sur X , à valeurs dans $\vec{E}$, définie par la formule

$$\vec{\nu}(\varphi) \text{ ou } p\vec{\mu}(\varphi) = \vec{\mu}(p\varphi), \quad \varphi \in \mathscr{C}(X). \tag{IV,5;1}$$

Cette formule définit bien une mesure.

En effet, si φ appartient à $\mathcal{C}(X)$, comme p est continue, $p\varphi$ est encore continue, et son support, contenu dans celui de φ, est encore compact, donc $p\varphi$ est elle aussi dans $\mathcal{C}(X)$, de sorte que le dernier membre de (IV,5;1) a bien un sens. Par ailleurs, ce dernier membre dépend linéairement de φ, donc $\vec{\nu}$ est une application linéaire de $\mathcal{C}(X)$ dans $\vec{E}$.

Il reste à démontrer sa continuité. Supposons que φ ait son support dans un compact K de X ; on a alors l inégalité

(IV,5;2) $$\|\vec{\nu}(\varphi)\| = \|\vec{\mu}(p\varphi)\| \leq \|\vec{\mu}\|_K \sup_{x\in K} |p(x)| \; |||\varphi|||$$

ce qui prouve la continuité de $\vec{\nu}$, et en même temps donne la majoration

(IV,5;3) $$\|\vec{\nu}\|_K \leq \|\vec{\mu}\|_K \sup_{x\in K} |p(x)|$$

Cette définition montre aussi pourquoi on ne peut pas multiplier une mesure par une fonction qui ne soit pas continue scalaire. En effet le dernier membre de (IV,5;1) ne possède un sens que si $p\varphi$ est une fonction continue scalaire à support compact, et cela n'est toujours vrai que si p est continue scalaire.

<u>Exemple 1</u> - Si $\vec{\mu}$ est la mesure $\sum \vec{c}_\nu \, \delta_{(a_\nu)}$, alors on a la formule

(IV,5;4) $$p\Big(\sum_\nu \vec{c}_\nu \, \delta_{(a_\nu)}\Big) = \sum_\nu \vec{c}_\nu \, p(a_\nu) \, \delta_{(a_\nu)} .$$

En particulier

$$p\,\delta_{(a)} = p(a)\,\delta_{(a)} .$$

<u>Exemple 2</u> - Si $\vec{\mu}$ est la mesure que nous avions appelée $\vec{f}\,dx$, et si q est une fonction continue scalaire, la mesure $q(\vec{f}\,dx)$ n'est autre que la mesure appelée, avec la même dénomination, $(q\vec{f})\,dx$. En particulier, <u>si f est continue scalaire, $f\,dx$ est bien la mesure produit de la fonction continue scalaire f par la mesure scalaire dx</u>.

Propriétés élémentaires

Le support de $\varphi \vec{\mu}$ est contenu trivialement dans l'intersection du support de φ et du support de $\vec{\mu}$.

Il peut être strictement plus petit que cette intersection. Prenons par exemple pour μ la mesure réelle de Dirac δ sur $\mathbb{R}$, et pour φ la fonction x. Alors on a $x\delta = 0$ (d'après (IV,5;4)).

Or les supports sont respectivement l'origine et la droite entière, dont l'intersection est l'origine, alors que le produit, qui est nul, a un support vide.

L'application $(\varphi, \vec{\mu}) \longrightarrow \varphi \vec{\mu}$, qui, au couple d'une fonction continue scalaire et d'une mesure à valeurs dans $\vec{E}$, fait correspondre une mesure à valeurs dans $\vec{E}$, est évidemment bilinéaire.

On a enfin la règle d'associativité suivante : si φ et q sont 2 fonctions continues scalaires,

(IV,5;5) $$\varphi(q\vec{\mu}) = (\varphi q)\vec{\mu}\,.$$

Cas où μ est une mesure réelle $\geqslant 0$

Dans le cas particulier où μ est une mesure de Radon réelle $\geqslant 0$, alors on peut définir des produits de façon beaucoup plus étendue, puisque l'on connait la valeur de μ sur des fonctions qui ne sont pas nécessairement continues à support compact; on peut même alors multiplier μ par des fonctions à valeurs vectorielles :

Définition - Soit μ une mesure de Radon réelle $\geqslant 0$ sur un espace X localement compact dénombrable à l'infini. Soit $\vec{\varphi}$ une fonction définie μ-presque partout sur X, à valeurs dans l'espace de Banach $\vec{E}$, et localement μ-intégrable, c'est-à-dire μ-intégrable sur tout compact de X *. Alors on appelle produit $\vec{\varphi}\mu$ la mesure $\vec{\nu}$ sur X à valeurs dans $\vec{E}$, définie par la formule

(IV,5,6) $$\vec{\varphi}\mu(\psi) = \int \vec{\varphi}\,\psi\, d\mu \in \vec{E}$$

* La fonction φ est localement intégrable, si et seulement si elle est μ-mesurable, et si, pour tout compact K de X.

$$\int_K^* \|\vec{\varphi}\|\, d\mu < +\infty$$

Montrons que cette formule définit bien une mesure sur X à valeurs dans $\vec{E}$, Si φ appartient à $\mathcal{C}(X)$, elle est en particulier μ -mesurable et bornée, et son support est contenu dans un compact K de X. Comme alors la fonction $\vec{f}$ est supposée intégrable sur K, la fonction $\vec{f}\varphi$ est k-mesurable (corollaire 2 du théorème 23 bis), et l'intégrale supérieure $\int^* \|\vec{f}\varphi\| d\mu \leq |||\varphi||| \int_K \|\vec{f}\| d\mu$ est finie, donc elle est intégrable, et le dernier membre de (IV,5;6) a un sens.

Naturellement ce dernier membre dépend linéairement de φ, et définit bien une application linéaire de $\mathcal{C}(X)$ dans $\vec{E}$, Quant à la continuité, elle résulte trivialement de l'inégalité ci-dessus, qui montre en outre que la norme de $\vec{f}\mu$ sur le compact K admet la majoration

(IV,5;7) $$\| \vec{f}\mu \|_K \leq \int_K \| \vec{f} \| d\mu .$$

En prenant la borne supérieure pour tous les compacts K, on en déduit :

(IV,5;7bis) $$\| \vec{f}\mu \| \leq \int \| \vec{f} \| d\mu ;$$

on montre que l'inégalité est une égalité si $\vec{f}$ est à valeurs scalaires ou dans un espace vectoriel de dimension 1, mais ne l'est pas, en général, si $\vec{f}$ est à valeurs dans un espace $\vec{E}$ de dimension ≥ 2.

Si f est une fonction à valeurs complexes (resp. réelles, resp. réelles ≥ 0) alors il en est de même du produit $f\mu$. Ici encore le support de $f\mu$ est contenu dans l'intersection des supports de f et de μ. D'autre part la dénomination $f\,dx$ utilisée antérieurement pour certaines mesures sur la droite, est maintenant pleinement justifiée.

Théorème 51 - Si μ est une. mesure réelle ≥ 0 sur un espace localement compact X dénombrable à l'infini, si f est une fonction réelle ≥ 0 localement μ -intégrable sur X, si $\vec{f}$ est une fonction définie sur X à valeurs dans un espace de Banach $\vec{F}$, alors la fonction$_4$ est $f\mu$-intégrable, si et seulement si la fonction $f\vec{f}$ est μ -intégrable. En outre

(IV,5;8) $$\int \vec{f}\, d(f\mu) = \int (f\vec{f})\, d\mu \quad ,^* \qquad \text{ou}$$
$$\int \vec{f}\,(f\,d\mu) = \int (\vec{f} f)\, d\mu$$

* On voit pourquoi la notation $f\,d\mu$ est recommandée; $d(f\mu)$ est en effet identique a $f\,d\mu$ dans (IV,5;8). Nous écrivions déjà $f\,dx$.

Nous admettrons ce théorème . Sa démonstration est basée sur l'égalité relative aux intégrales supérieures : si f est quelconque ≥ 0 ,

$$\int^* f\,(\vec{p}\,d\mu) = \int^* (f\,\vec{p})\,d\mu\,. \qquad \text{(IV,5;9)}$$

Ce théorème est évidemment particulièrement utile, et on l'use constamment dans la pratique sans même s'en apercevoir.

En particulier, si l'on a, pour des fonctions à valeurs réelles ≥ 0 , une expression $\int fg\ d\mu$, on peut la considérer indifféremmment comme signifiant l'intégrale de fg par rapport à $d\mu$, ou l'intégrale de f par rapport à la mesure produit $g\ d\mu$, ou l'intégrale de g par rapport à la mesure produit $f\,d\mu$; ceci naturellement à condition que ces mesures-produit aient un sens, c'est-à-dire que f et g soient localement μ- intégrables.

Corollaire. Si $\vec{p}$ est une fonction sur X à valeurs dans $\vec{E}$, g une fonction réelle ≥ 0 localement μ-intégrable, alors $\vec{p}$ est localement $g\,\mu$-intégrable, si et seulement si $\vec{p}\,g$ est localement μ-intégrable, et les mesures produits $\vec{p}\,(g\,\mu)$ et $(\vec{p}\ \ g)\ \mu$ coïncident.

Elles ont en effet la même valeur sur toute $\varphi \in \mathscr{C}(X)$.
On dit que la mesure $\vec{p}\,\mu$ est de densité $\vec{p}$ par rapport à la mesure μ réelle ≥ 0 .

Mesures de base μ , mesures de base ≥ 0 .
On dit qu'une mesure $\vec{\nu}$, définie sur X à valeurs dans l'espace de Banach $\vec{E}$, est de base μ , où μ est une mesure réelle ≥ 0 , si elle peut s'écrire sous la forme $\vec{p}\,\mu$, où $\vec{p}$ est une fonction sur X à valeurs dans $\vec{E}$, localement μ - intégrable. On dit que $\vec{\nu}$ est de base ≥ 0 , s'il existe une mesure $\mu \geq 0$ telle que $\vec{\nu}$ soit de base μ .

On peut alors se proposer de chercher :

1°/ De combien de manières une même mesure $\vec{\nu}$ peut être exprimée sous la forme $\vec{p}\,\mu$, μ donnée ;

2°/ Quelles sont les mesures qui peuvent s'exprimer sous cette forme, c'est-a-dire qui sont de base μ .

'Théorèmes 52 - 53 - 54 - (Lebesgue - Nikodym et Dunford-Pettis). *

Soit X un espace localement compact dénombrable à l'infini, et soit μ une mesure de Radon ≥ 0 sur X.

1°/ soit ν une autre mesure ≥ 0 sur X. Si ν est de base μ, toute fonction ou partie μ- mesurable est aussi? - mesurable. En outre tout ensemble de μ- mesure nulle est aussi de ν- mesure nulle. Réciproquement, si toute partie borélienne d'adhérence compacte de μ- mesure nulle est aussi de ν- mesure nulle, ν est de base μ, $\nu = f\mu$, et f est unique, à une modification prés sur un ensemble de μ- mesure nulle (si $f_1\mu = f_2\mu$, f_1 et f_2 sont μ - presque partout égales).

Pour que la mesure $\nu \geq 0$ soit majorée par μ, il faut et il suffit que $\nu = f\mu$, $f \leq 1$ p-presque partout.

2°/ Toute mesure $\vec{\nu}$ à valeurs dans un Banach $\vec{E}$, de base ≥ 0, admet un prolongement borélien, est absolument majorée, et admet une plus petite majorante absolue. Si on suppose que $\vec{\nu}$ est de base ≥ 0, pour qu'elle soit de base μ, il faut et il suffit que toute partie borélienne d'adhérence compacte, de k-mesure nulle, soit aussi de $\vec{\nu}$- mesure nulle; alors $\vec{\nu} = \vec{f}\mu$, où f est déterminée à une modification près sur un ensemble de μ- mesure nulle; pour qu'elle soit absolument majorée par μ, il faut et il suffit que $\vec{\nu} = \vec{f}\mu$, $\|\vec{f}\| \leq 1$ k-presque partout, et alors, pour toute f borélienne bornée à support compact,

$$\left\| \int f\, d\vec{\nu} \right\| \leq \int |f|\, d\mu \; ; \tag{IV, 5 ; 9 bis}$$

pour qu'elle admette μ comme plus petite majorante absolue, il faut et il suffit que $\vec{\nu} = \vec{f}\mu$, $\|\vec{f}\| = 1$ μ-presque partout.

3°/ Si $\vec{E}$ est de dimension finie, toute mesure ν à valeurs dans $\vec{E}$ est de base 3 0. En particulier, toute mesure scalaire est de base ≥ 0.

* Nous réunissons plusieurs théorèmes en un seul La partie 1" est dûe a Lebesgue et Nikodym; la partie 2° est'un cas particulier d'un théorème plus général de Dunford et Pettis.

4°/ Si f et g sont des fonctions réelles sur X, localement μ-intégrables, on a $f\mu \leq g\mu$ si et seulement si $f \leq g$, μ-presque partout; la borne supérieure des mesures $f\mu$ et $g\mu$ est la mesure $(\mathrm{Sup}(f, g))\mu$. Si $\vec{\nu}$ est a valeurs dans $\vec{E}$ et de base ≥ 0, pour qu'elle soit absolument majorée par $g\mu$, il faut et il suffit que $\vec{\nu} = \vec{f}\mu$, $\|\vec{f}\| \leq g$ μ-presque partout.

Démonstration - A) Soit d'abord $\nu \geq 0$ $\nu = f\mu$, $f \geq 0$ localement μ-intégrable. Soit A une partie μ-mesurable d'adhérence compacte. Soit $\mathcal{O}_n$ une suite d'ouverts contenant A, tels que $\mu(\mathcal{O}_n) \leq \mu(A) + \frac{1}{n}$, et soit K_n une suite de compacts contenus dans A tels que $\mu(K_n) \geq \mu(A) - \frac{1}{n}$; soit $A^* = \bigcap \mathcal{O}_n$, $A_* = \bigcup_n K_n$. A^* et A_* sont boréliens donc mesurables pour toute mesure de Radon, $A_* \subset A \subset A^*$, et $\mu(A^* - A_*) = 0$. Alors $\nu(A^* - A_*) = \int_{A^*-A_*} f\, d\mu = 0$; donc $\nu(A_*) \leq \nu_*(A) \leq \nu^*(A) \leq \nu(A^*)$, donc A est aussi ν-mesurable; si A n'est pas d'adhérence compacte, la mesurabilité de A étant équivalente à celle de ses intersections avec les compacts, la conclusion subsiste. Et si $\mu(A) = 0$, $\nu(A) = \int_A f\, d\mu = 0$. La mesurabilité des fonctions se ramenant à celle des ensembles, toute fonction μ-mesurable est ν-mesurable.

Inversement, supposons que tout ensemble borélien d'adhérence compacte de μ-mesure nulle soit aussi de ν-mesure nulle. Nous allons construire une fonction $f \geq 0$, localement μ-intégrable, telle que $\nu = f\mu$, donc ν sera bien de base μ; si $\nu \leq \mu$, la construction montrera directement que $f \leq 1$. La partie 1°/ du théorème ne sera cependant pas encore démontrée, car il restera a prouver que f est déterminée à une modification arbitraire près sur un ensemble de μ-mesure nulle.

Construisons donc cette fonction f.

Soient k et n des entiers ≥ 0. SI nous considérons la mesure $\nu - \frac{k}{2^n}\mu$, on peut d'après le théorème 50 quinto décomposer X en réunion de deux ensembles boréliens disjoints $X^+_{k,n}$ et $X^-_{k,n}$, tels que, sur le premier, $\nu - \frac{k}{2^n}\mu$ soit complètement positive, et, sur le deuxième, complètement négative. Pour $k = 0$, nous pourrons prendre $X^+_{0,n} = X$, $X^-_{0,n} = \emptyset$

Posons alors $A_{k,n} = X^+_{k,n} \cap X^-_{k+1,n}$. Soit $h < k$; il n'est pas certain que $X^-_{h,n} \subset X^-_{k,n}$; mais, si nous appelons $N_{h,k,n}$ l'ensemble borélien $X^-_{h,n} \cap X^+_{k,n}$ des points qui appartiennent au premier sans appartenir au deuxième, il est de μ- mesure nulle; en effet pour toute partie borélienne B d'adhérence compacte de cet ensemble, on a les inégalités $\nu(B) \leq \frac{h}{2^n}\mu(B)$ et $\geq \frac{k}{2^n}\mu(B)$, ce qui n'est possible que si $\mu(B) = 0$, donc $\mu_*(N_{h,k,n})=0$ et par suite $\mu(N_{h,k,n}) = 0$ puisque c'est un ensemble borélien donc μ- mesurable. Si nous appelons N_n la ré-union de tous les $N_{h,k,n}$ correspondant à toutes les valeurs h, k, $h < k$, c'est un ensemble de μ mesure nulle. Si x n'appartient pas à cet ensemble exceptionnel N_n et, si $x \in X^-_{h,n}$ on a aussi $x \in X^-_{k,n}$ pour tout $k \geq h$. Il en résulte que l'intersection de $A_{h,n}$ et de $A_{k,n}$, $h \neq k$, est contenue dans N_n, donc de μ- mesure nulle (en effet supposons par exemple $h < k$; $(A_{h,n} \cap A_{k,n}) \subset (X^-_{h+1,n} \cap X^+_{k,n}) \subset N_n$).

Par ailleurs la réunion de tous les $A_{k,n}$ pour $k = 0$, $1, 2, \dots, m, \dots$ et n fixé, contient presque tout X. Soit en effet $x \in X$. Supposons d'abord qu'il existe au moins un entier h tel que $x \in X^-_{h,n}$; il existe alors un premier entier h (nécessairement $h \geq 1$ puisque $X^-_{0n} = \emptyset$) à avoir cette propriété; alors nécessairement%@ $X^-_{h-1,n}$ et donc $x \in X^+_{h-1,n}$ et par suite 2 $\in A_{h-1,n}$. Mais si nous appelons M_n l'ensemble des points x qui n'appartiennent a aucun des $X^-_{k,n}$, pour $k = 0, 1, 2, \dots$, c'est l'ensemble des points qui appartient à tous les $X^+_{k,n}$. par suite pour toute partie borélienne à support compact B' de cet ensemble M_n $\nu(B) \geq \frac{k}{2^n}\mu(B)$ quel que soit k, ce qui impose $\mu(B) = 0$, de sorte que M_n est de μ-mesure intérieure nulle. donc de μ-mesure nulle puisqu'il est borélien. On peut résumer ceci en disant que à des ensembles de μ- mesure nulle près, les $A_{k,n}$, pour n fixé et $k = 0, 1, 2, \dots$, sont deux à deux disjoints et de réunion X. Appelons alors f_n la fonction égale, dans $A_{k,n}$, à $\frac{k}{2^n}$. Elle n'est pas définie aux points d'intersection de deux des $A_{k,n}$ ou aux points qui n'appartiennent pas à leur réunion; c'est-à-dire sur l'ensemble $N_n \cup M_n$ borélien de μ - mesure nulle. On pourra par exemple, prendre $f_n = 0$ sur cet ensemble; f_n est alors borélienne. Elle est en outre localement μ - intégrable; car, pour tout compact K, $\int_K f_n \, d\mu = \sum_{n\geq 0} \int_{K\cap A_{k,n}} \frac{k}{2^n} d\mu \leq \sum_{n\geq 0} \nu(K\cap A_{k,n}) \leq \nu(K) < +\infty$.

La mesure $f_n \mu$ est donc bien définie. Dans chaque ensemble $A_{k,n}$ on a $f_n \mu \ll \nu \ll (f_n + \frac{1}{2^n})\mu$. Ces mêmes inégalités sont donc encore exactes sur $\bigcup_{k\geq 0} A_{k,n}$; mais aussi sur M_n, car il est de μ - mesure nulle donc aussi de ν - mesure nulle d'après l'hypothèse faite sur ν et μ (que nous utilisons ici seulement), donc sur X entier:

Faisons maintenant varier n. Une fois tous les choix faits pour n, nous les ferons pour $n+1$, en prenant $X^{\pm}_{2k\, n+1} = X^{\pm}_{k,n}$, ce qui est possible puisqu'ils ont simplement à satisfaire aux mêmes inégalités

$$\nu - \frac{k}{2^n}\mu = \nu - \frac{2k}{2^{n+1}}\mu \ll 0 \quad \text{ou} \gg 0$$

Alors, à des ensembles de μ - mesure nulle près, $A_{2k,n+1}$ et $A_{2k+1,n+1}$ sont 2 ensembles disjoints de réunion $A_{k,n}$. On a donc μ-presque partout : $f_n \leq f_{n+1} \leq f_n + \frac{1}{2^n}$.
Si nous modifions les f_n sur les ensembles boréliens de μ - mesure nulle en les rendant nulles sur l'ensemble borélien de μ - mesure nulle $M \cup N$, $M = \bigcup_{n\geq 0} M_n$, $N = \bigcup_{n\geq 0} N_n$, elles continuent à vérifier les mêmes propriétés, mais on a partout $f_n \leq f_{n+1} \leq f_n + \frac{1}{2^n}$. La suite des f_n est donc croissante, et a une limite f avec $f_n \leq f \leq f_n + \frac{1}{2^n}$; donc f est borélienne et localement μ - intégrable, donc définit une mesure $f\mu$. Pour toute φ de $\mathscr{C}(X)$, $\varphi \geq 0$, on a

$$\int f_n \varphi \, d\mu \leq \int \varphi \, d\nu \leq \int f_n \varphi \, d\mu + \frac{1}{2^n}\int \varphi \, d\mu$$

$$\int f_n \varphi \, d\mu \leq \int f \varphi \, d\mu \leq \int f_n \varphi \, d\mu + \frac{1}{2^n}\int \varphi \, d\mu ,$$

donc
$$\left| \int f \varphi \, d\mu - \int \varphi \, d\nu \right| \leq \frac{1}{2^n}\int \varphi \, d\mu ,$$

donc $= 0$, n étant arbitraire. Donc $f\mu$ et ν coïncident sur $\mathscr{C}_+(X)$ et par suite sur $\mathscr{C}(X)$ et sont égales; c'est ce que nous voulions démontrer. Si en outre $\nu = \mu$, on peut prendre $X^-_{2^n,n} = X$ et $X^+_{2^n,n} = \emptyset$ donc $f_n \leq 1$ et $f \leq 1$.

B) Soit $\vec{\nu}$ une mesure à valeurs dans un Banach $\vec{E}$, de base $\lambda \geq 0$; on a donc $\vec{\nu} = \vec{q}\lambda$, $\vec{q}$ localement λ-intégrable. Elle est alors trivialement absolument majorée par la mesure $\|\vec{q}\|\lambda \geq 0$. Donc elle admet une plus petite majorante absolue d'après ce qui suit le théorème 18. Par ailleurs elle admet un prolongement borélien par

$$\int f\, d\vec{\nu} = \int f\vec{q}\, d\lambda \tag{IV,5,9 ter}$$

Nous définissons bien là en effet une application linéaire de $\Gamma(X)$ dans ? , prolongeant l'application définie par $\vec{\nu}$ sur $\mathcal{C}(X)$ car, si f est borélienne donc λ-mesurable, $f\vec{q}$ est λ-mesurable; mais $|f|$ est bornée par $M \geq 0$ et à support compact K, et $\vec{q}$ est localement λ-intégrable donc $\chi_K \vec{q}$ est λ-intégrable; donc $\|f\vec{q}\| \leq M\chi_K \|\vec{q}\|$ a une λ-intégrale supérieure finie, donc $f\vec{q}$ est bien λ-intégrable. Il reste à démontrer la L-continuité de cette application. Supposons que des f_n convergent vers f, en restant bornées en module par M et en gardant leur support dans le compact K ; les $f_n\vec{q}$ convergent vers $f\vec{q}$, en restant majorées en norme par la fonction λ-intégrable fixe $\chi_K M\|\vec{q}\|$, donc le théorème 35 de convergence de Lebesgue montre que $\int f_n\vec{q}\, d\lambda$ converge vers $\int f\vec{q}\, d\lambda$. Mais il y a aussi à vérifier la condition 3°/ de la continuité (voir page 521). Soit donc $\mathcal{O}$ un ouvert d'adhérence compacte. La fonction $\|\vec{q}\|$ est intégrable sur $\mathcal{O}$; il existe donc $M \geq 0$ tel que $\int_{\mathcal{O}} (\|\vec{q}\| - \|\vec{q}\|_M)\, d\lambda \leq \frac{\varepsilon}{2}$, où il existe ensuite un compact $K \subset \mathcal{O}$ tel que $\lambda(\mathcal{O} - K) \leq \frac{\varepsilon}{2M}$. Le compact K répond a la condition 3°/ pour $\vec{\nu}$. Si en effet $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq 1$, φ égale a 1 sur un voisinage de K à support dans $\mathcal{O}$, on a

$$\|\vec{\nu}(\varphi) - \vec{\nu}(\mathcal{O})\| \leq \int_{\mathcal{O}-K} \|\vec{q}\|\, d\lambda \leq \int_{\mathcal{O}-K} \|\vec{q}\|_M\, d\lambda$$

$$+ \int_{\mathcal{O}-K} (\|\vec{q}\| - \|\vec{q}\|_M)\, d\lambda \leq M \cdot \frac{\varepsilon}{2M} + \frac{\varepsilon}{2} = \varepsilon$$

Remarque - Nous avons donné page 456-9 l'exemple d'une mesure non absolument majorée, et n'ayant pas de prolongement borélien; ce que nous venons de dire prouve qu'elle n'est pas de base $\geqslant 0$. On peut aussi donner un exemple d'une mesure ayant une majorante absolue (donc une plus petite majorante absolue), mais qui n'est pas de base $\geqslant 0$. Les deux notions ne sont donc pas équivalentes, l'une est plus forte que l'autre.

C) Il est évident qu'une mesure de la forme $\vec{f}\mu$, avec $\|\vec{f}\| \leqslant 1$ μ- presque partout, est absolument majorée par μ . Nous devons donc démontrer la réciproque; nous nous bornerons à le faire si $\vec{E}$ est séparable. Nous ne démontrerons d'abord qu'une réciproque partielle : si $\vec{\nu}$ est déjà de la forme $\vec{f}\mu$, et si elle est absolument majorée par μ , alors $\|\vec{f}\| \leqslant 1$ μ-presque partout. Supposons que l'ensemble $\{x \in X ; \|\vec{f}(x)\| > 1\}$ soit de μ- mesure > 0 ; comme il est la réunion des ensembles $A_\varepsilon = \{x \in X ; \|\vec{f}(x)\| \geqslant 1+2\varepsilon\}$, pour $\varepsilon = 1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{n}, \cdots$, il existe un $\varepsilon > 0$ pour lequel cet ensemble A_ε est de μ mesure > 0 . Nous prendrons $\varepsilon < \frac{1}{2}$; on verra plus loin pourquoi.

Soit a un point de la sphère unité S de $\vec{E}$, et soit $B(a,\varepsilon)$ la boule de centre a et de rayon ε ; elle est fermée convexe. Appelons $K(a,\varepsilon)$ la réunion des homothétiques de cette boule, de centre origine et de rapport réel $\geqslant 1$. On voit alors immédiatement que $K(a,\varepsilon)$ est aussi un ensemble convexe fermé (soient en effet $\alpha\vec{x}$ et $\beta\vec{y}$ deux points de cet ensemble, avec $\alpha \geqslant 1$, $\beta \geqslant 1$, $\vec{x}$ et $\vec{y}$ dans $B(a,\varepsilon)$; alors le point $t\alpha\vec{x} + (1-t)\beta\vec{y}$ peut aussi s'écrire $\gamma(s\vec{x} + (1-s)\vec{y})$, avec $\gamma = t\alpha + (1-t)\beta \geqslant 1$ $s = \frac{t\alpha}{t\alpha + (1-t)\beta}$, $1-s = \frac{(1-t)\beta}{t\alpha+(1-t)\beta}$, $s\vec{x} + (1-s)\vec{y} \in B(a,\varepsilon)$, donc il appartient encore à $K(a,\varepsilon)$, qui est bien convexe.

Supposons maintenant qu'une suite de points $k_n \vec{x}_n$ de cet ensemble convexe vers une limite $\vec{y}$ de $\vec{E}$. Comme les normes des $\vec{x}_n$ sont $\geqslant 1-\varepsilon$, on voit que les k sont bornés supérieurement, et ils sont aussi bornés inférieurement par 1, on peut donc extraire de la suite une suite partielle, que nous continuerons à appeler du même nom, pour laquelle les k_n convergent vers une limite $k \geqslant 1$. Mais alors les $\vec{x}_n$ convergent vers la limite $\vec{x} = \frac{\vec{y}}{k}$, qui appartient nécessairement à $B(a,\varepsilon)$ fermée; donc $\vec{y} = k\vec{x}$, $k \geqslant 1$, $\vec{x} \in B(a,\varepsilon)$ donc $\vec{y}$ appartient à $K(a,\varepsilon)$, qui est bien fermé).

Par ailleurs, et c'est là qu'intervient le fait que $\vec{E}$ est séparable, il existe une infinité dénombrable de points a, soit $a_1, a_2, \dots, a_n, \dots$, tels que les B (a_n, ε) recouvrent la sphère S (*). Alors les ensembles convexes fermés $K(a_n, \varepsilon)$ couvrent l'ensemble des points de norme $\geqslant 1$ de $\vec{E}$ et les ensembles $(1+2\varepsilon)K(a_n,\varepsilon)$ couvrent l'ensemble des points de $\vec{E}$ de mesure $\geqslant 1+2\varepsilon$.

A_ε est alors réunion de l'ensemble de parties $A_n = A_\varepsilon \cap \vec{f}^{-1}((1+2\varepsilon)K(a_n,\varepsilon))$ qui sont tous des ensembles μ-mesurables; l'un au moins d'entre eux à donc une μ-mesure > 0 soit A_1 ; soit $K_1 \subset A_1$ un compact de μ-mesure > 0 . L'inégalité de convexité ( ) donne alors, puisque $\vec{f}(K_1) \subset (1+2\varepsilon)K(a_1,\varepsilon)$: $\int_{K_1} \vec{f}\, d\mu \in \mu(K_1)(1+2\varepsilon)K(a_1,\varepsilon)$,

donc $\left\| \int_{K_1} \vec{f}\, d\mu \right\| \geqslant \mu(K_1)(1+2\varepsilon)(1-\varepsilon) > \mu(K_1)$,

puisque $\mu(K_1) \neq 0$ et $\varepsilon < \frac{1}{2}$.

Mais ceci est impossible car $\vec{f}\mu$ est supposée absolument majorée par μ ; on a donc, pour toute $\varphi \geqslant 0$ de $\mathcal{C}(X)$ $\left\| \int \varphi \vec{f}\, d\mu \right\| \leqslant \mu(\varphi)$; et la condition 3bis de la L - continuité (voir page 521) montre que pour tout compact K_1, $\left\| \int \chi_{K_1} \cdot \vec{f}\, d\mu \right\| \leqslant \mu(K_1)$. On a donc bien $\|\vec{f}\| \leqslant 1$ presque partout si $\vec{f}\mu$ est absolument majorée par μ .

D) Il en résulte en particulier que si la mesure $\vec{f}\mu$ est nulle, alors, quel que soit le nombre réel $k > 0$, la mesure $k\vec{f}\mu$ est absolument majorée par μ , donc $k\|\vec{f}\|$ est μ-presque partout majorée par 1, c'est-à-dire $\|\vec{f}\|$ est μ-presque partout majorée par $\frac{1}{k}$; en faisant tendre k vers l'infini, on voit que $\|\vec{f}\|$ est μ-presque partout nulle. Donc si $\vec{f}_1\mu$ et $\vec{f}_2\mu$ représentent la même mesure, $\vec{f}_1$ et $\vec{f}_2$ sont μ-presque partout égales. Ceci démontre l'unicité qui figure tant dans le 1° que dans le 2° de l'énoncé du théorème, et le 1°, de l'énoncé est complètement démontré.

* Tout sous-ensembles d'un espace métrique séparable E est lui-même séparable. Si en effet la suite des b_n de E est dense, et si a_n est choisi dans S tel que $d(a_n,b_n) \leqslant d(b_n,S)+\frac{1}{n}$, la suite des a_n est dense dans S .

E) Supposons maintenant que $\vec{\nu} = \vec{f}\mu$ admette μ comme plus petite majorante absolue; alors d'une part on a $\|\vec{f}\| \leqslant 1$ μ - presque partout; mais par ailleurs la mesure $\vec{f}\mu$ est toujours absolument majorée par $\|\vec{f}\|\mu \leqslant \mu$; si donc μ est la plus petite majorante absolue, on a $\|\vec{f}\|\mu = \mu$, donc nécessairement $\|\vec{f}\|$ est μ - presque partout égale à 1. Inversement supposons que $\|\vec{f}\|$ soit μ - presque partout egale a 1; alors d une part $\vec{f}\mu$ est absolument majorée par μ ; montrons que μ est sa plus petite majorante absolue.

Soit en effet λ cette dernière; on a $\lambda \leq \mu$. Supposons que $\lambda < \mu$; il existe alors ε, $0 < \varepsilon \leq \frac{1}{2}$, et un ensemble A borélien d'adhérence compacte A tel que $\mu(A) > (1+2\varepsilon)\lambda(A)$ (si en effet, pour tout $\varepsilon > 0$, on a $\mu(A) \leq (1+2\varepsilon)\lambda(A)$, on a, d'après le corollaire 2 du théorème 50 bis, $\mu \leq (1+2\varepsilon)\lambda$; celà donnerait $\mu \leq \lambda$, contrairement à l'hypothèse). Mais $\|\vec{f}\| = 1$, et la sphère unité de E , supposé séparable, est recouverte par une infinité dénombrable de boules (fermées convexes) $B(a_n, \varepsilon)$.

On a donc nécessairement, pour l'un au moins des $A_n = A \cap \vec{f}^{-1}(B(a_n, \varepsilon))$, $\mu(A_n) > (1+2\varepsilon)\lambda(A_n)$, par exemple pour $n = 1$. Il existe alors un compact $K_1 \subset A_1$ tel que $\mu(K_1) > (1+2\varepsilon)\lambda(K_1)$. Mais, sur K_1 , la fonction $\vec{f}$ prend ses valeurs dans l'ensembfe convexe fermé $B(a_1, \varepsilon)$; on a donc la relation de convexité $\int_K \vec{f}\,d\mu \in \mu(K_1)(a_1, \varepsilon)$, ou $\|\int_{K_1} \vec{f}\,d\mu\| \geqslant (1-\varepsilon)\mu(K_1) > (1-\varepsilon)(1+2\varepsilon)\lambda(K_1) \geqslant \lambda(K_1)$

$$\text{donc} \quad \left\|\int_K \vec{f}\,d\mu\right\| > \lambda(K_1)$$

ce qui est contradictoire avec l'hypothèse suivant laquelle $\vec{f}\mu$ est absolument majorée par λ , comme nous l'avons déjà vu. Donc $\vec{f}\mu$ admet μ comme plus petite majorante 'absolue, si et seulement si $\|\vec{f}\|$ est μ - presque partout égale à 1.

Soit alors $\vec{f}\mu$ une mesure quelconque, $\vec{f}$ étant localement μ - intégrable. Puisqu'elle peut s'écrire

$$\frac{\vec{f}}{\|\vec{f}\|}\left(\|\vec{f}\|\ \mu\right), \quad \text{ou} \quad \left\|\frac{\vec{f}}{\|\vec{f}\|}\right\| = 1 \ ^{*}$$

Sa plus petite majorante absolue est donc la mesure $\|\vec{f}\|\mu$.

F) Soit $\vec{\nu}$ une mesure de base $\geqslant 0$, absolument majorée par μ ; si nous savons que $\vec{\nu} = \vec{f}\mu$, alors C) nous prouve que $\|\vec{f}\| \leqslant 1$ μ-presque partout. Montrons donc que $\vec{\nu}$ est de base μ Elle est de base $\geqslant 0$, donc $\vec{\nu} = \vec{q}\lambda$; nous venons alors de voir que sa plus petite majorante absolue est $\|\vec{q}\|\lambda$. Puisqu'elle est absolument majorée par μ , on a donc $\|\vec{q}\|\lambda \leqslant \mu$; d'après la partie A), on sait alors que $\|\vec{q}\|\lambda = g\mu$, $0 \leqslant g \leqslant 1$; donc $\vec{\nu} = \vec{q}\lambda = \frac{\vec{q}}{\|\vec{q}\|}(\|\vec{q}\|\lambda) = \frac{\vec{q}}{\|\vec{q}\|} g\mu = \vec{f}\mu$, ce qui prouve bien notre affirmation, De la même manière, si $\vec{\nu}$ est de base $\geqslant 0$ et admet μ comme plus petite majorante absolue, elle admet en particulier μ comme majorante absolue, donc est de la forme $\vec{f}\mu$, et alors E) nous dit que $\|f\| = 1$ μ - presque partout. Nous avons donc bien trouvé la condition nécessaire et suffisante pour qu'une mesure $\vec{\nu}$ de base $\geqslant 0$ soit absolument majorée par μ (resp. pour qu'elle ait μ comme plus petite majorante absolue): c'est que $\vec{\nu} = \vec{f}\mu$, avec $\|\vec{f}\| \leqslant 1$ (resp. $=1$) μ - presque partout.

L'inégalité (IV,5.9 bis), vraie par hypothèse si f est dans $\mathcal{C}(X)$, (et déjà utilisée a 2 reprises dans la démonstration si f est la fonction caractéristique d'un compact) est alors trivialement vraie pour toute f de $\Gamma_1(X)$, puisque $\vec{\nu} = \vec{f}\mu$, avec $\|\vec{f}\| \leqslant 1$ μ - presque partout , et que l'on a (IV,5;9 ter).

* $\frac{\vec{f}}{\|\vec{f}\|}$ n'est pas définie aux points où $\vec{f} = \vec{0}$. Mais c'est là un ensemble de $\|\vec{f}\|\mu$ - mesure nulle; nous choisirons arbitrairement la valeur de $\frac{\vec{f}}{\|\vec{f}\|}$ sur cet ensemble, mais de norme 1, de sorte qu'elle est partout de norme 1.

G) Si $\vec{\nu}$ est de base μ, alors tout ensemble borélien d'adhérence compacte de μ- mesure nulle est aussi de $\vec{\nu}$- mesure nulle; montrons la réciproque, ν étant supposée de base ≥ 0; posons donc! $\vec{\nu} = \vec{q}\lambda$ avec $\| \vec{q} \| \equiv 1$, ce qui est possible si nous prenons pour λ sa Plus Petite majorante absolue.

Soit A un ensemble borélien d'adhérence compacte de μ- mesure nulle; alors il est de $\vec{\nu} = \vec{q}\lambda$ - mesure nulle, ainsi que tous ses sous-ensembles boréliens, donc la mesure $\chi_A \vec{q}\lambda$, nulle sur tous les boréliens d'adhérence compacte, est nulle; donc $\chi_A \vec{q}$ est λ-presque partout nulle, donc aussi sa norme χ_A, autrement dit A est de λ- mesure nulle. D'après A), λ est donc de base μ, et par suite aussi $\vec{\nu} = \vec{q}\lambda$

Maintenant les parties 1°/ et 2°/ de l'énoncé sont complètement démontrées.

H) L'ensemble des mesures sur X à valeurs dans $\vec{E}$, et de base ≥ 0, est un espace vectoriel. Si en effet $\vec{\nu}_1$ et $\vec{\nu}_2$ sont de base λ_1 et λ_2, $\vec{\nu}_1 = \vec{q}_1\lambda_1$, $\vec{\nu}_2 = \vec{q}_2\lambda_2$, λ_1 et λ_2 sont toutes deux majorées par $\lambda = \lambda_1 + \lambda_2$, donc $\lambda_1 = g_1\lambda$, $\lambda_2 = g_2\lambda$, g_1 et g_2 comprises entre 0 et 1; donc $\vec{\nu}_1 + \vec{\nu}_2 = (\vec{q}_1 g_1 + \vec{q}_2 g_2)\lambda$ est de base ≥ 0 De même, si $\vec{\nu}$ est de base ≥ 0, et k un scalaire, $k\vec{\nu}$ est de base ≥ 0. Toute mesure réelle, étant différence de 2 mesures ≥ 0, est alors de base ≥ 0 (plus précisément, $\mathfrak{Z}$ est de base $|\mathfrak{Z}|$, qui est sa plus petite majorante absolue); toute mesure complexe $\nu = \nu_1 + i\,\nu_2$ est de base ≥ 0. Si alors $\vec{E}$ est de dimension finie, et si $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, est une base de $\vec{E}$, on a $\vec{\nu} = \sum_{i=1}^{n} \vec{e}_i\, \nu_i$, ν_i étant une mesure scalaire; ν_i est de base ≥ 0, donc trivialement aussi $\vec{e}_i\,\nu_i$ donc aussi $\vec{\nu}$. Donc la partie 3° de l'énoncé est démontrée.

I) Toute la partie 4°/ de l'énoncé est maintenant triviale. Soit f réelle, localement μ- intégrable. On sait que $|f\mu|$ est la plus petite majorante absolue de $f\mu$ (voir apres théorème 18); et la fin de E) nous dit que $|f|\,\mu^{f}$ est cette plus petite majorante absolue; donc $|f\mu| = |f|\,\mu$

Mais alors $(p\mu)^+ = \dfrac{|p\mu| + p\mu}{2} = \dfrac{|p|+p}{2}\mu = p^+\mu$,
et de même $(p\mu)^- = p^-\mu$.

Si alors p et q sont réelles, localement μ-intégrables, $\mathrm{Sup}(p\mu, q\mu) = p\mu + ((q-p)\mu)^+ = (p+(q-p)^+)\mu = (\mathrm{Sup}(p,q))\mu$; et de même $\mathrm{Inf}(p\mu, q\mu) = (\mathrm{Inf}(p,q))\mu$. En particulier $p\mu \leq q\mu$ exprime que leur borne supérieure est $q\mu$, donc que $\mathrm{Sup}(p,q) = q$ μ-presque partout ou $p \leq q$ μ-presque partout (et $p\mu \geq 0$ équivaut à : $p \geq 0$ μ-presque partout).

Si enfin $\vec{p}$ est à valeurs dans $\vec{E}$, q à valeurs réelles ≥ 0, et si $\vec{p}\mu$ est absolument majorée par $q\mu$, comme sa plus petite majorante absolue est $\|\vec{p}\|\mu$ d'après E) on a $\|\vec{p}\|\mu \leq q\mu$, donc $\|\vec{p}\| \leq q$ μ-presque partout; mais si $\vec{\nu}$, à valeurs dans $\vec{E}$, est de base ≥ 0 et absolument majorée par $q\mu$, on sait d'après F) que $\vec{\nu}$ est de base $q\mu$, donc de base μ, donc $\vec{\nu} = \vec{p}\mu$ et alors $\|\vec{p}\| \leq q$ μ-presque partout. Ceci achève de démontrer 4°/ et le théorème.

Remarque 1 - Dans tous les énoncés relatifs à une mesure $\vec{\nu}$ à valeurs dans $\vec{E}$, on suppose; de base ≥ 0 (par exemple : si $\vec{\nu}$ de base ≥ 0, est absolument majorée par μ, alors $\vec{\nu} = \vec{p}\mu$, $\|\vec{p}\| \leq 1$ μ-presque partout). Des contre-exemples montrent que cette hypothèse est inévitable sauf évidemment si $\vec{E}$ est de dimension finie, puisqu'alors toute mesure à valeurs dans $\vec{E}$ est de base ≥ 0. Le théorème général de Dunford-Pettis donne une largo catégorie d'espaces $\vec{E}$ pour lesquels on n'a pas besoin de supposer $\vec{\nu}$ de base ≥ 0.

Remarque 2 - Les physiciens interprètent souvent la mesure de Dirac δ sur la droite, comme si elle pouvait s'écrire $\delta(x)\,dx$ comme produit de dx par une fonction. Nous voyons maintenant qu'une telle écriture est impossible. En effet, l'origine est un ensemble de mesure nulle pour dx et non pour δ, donc δ n'est pas de base dx.

Corollaire 1 - Soient μ et ν deux mesures ≥ 0 sur X. Il existe une plus grande mesure ν_1, majorée par ν et de base μ, donnée par la formule :

$$\nu_1 = \operatorname*{Sup}_{n \in \mathbb{N}} (\mathrm{Inf}(\nu, n\mu)) \qquad \text{(IV,5;9 quarto)}$$

Il existe une plus grande mesure ν_2, majorée par ν et μ-singulière *. On a $\nu = \nu_1 + \nu_2$, et c'est la seule décomposition de ν en sommes de deux mesures ≥ 0, la première de base μ, la deuxième μ-singulière. Enfin ν_1 et ν_2 sont étrangères.

Demonstration - Tout d'abord Inf $(\nu, n\mu)$ est majorée par $n\mu$ donc, d'après le 1°/ du théorème, de la forme $f_n \mu$, $f_n \leq n$ μ-presque partout. Mais, lorsque n croît, les mesures ainsi déterminées croissent, donc $f_n \mu \leq f_{n+1} \mu$, ou $f_n \leq f_{n+1}$ μ-presque partout. La suite des fonctions f_n est donc, à des ensembles de μ-mesure nulle près, croissante. Soit f la limite, définie μ-presque partout. Les f_n étant μ-mesurables f l'est aussi; d'autre part, pour tout borélien A d'adhérence compacte, $\int_A f_n \, d\mu \leq \nu(A)$

donc, d'après le théorème 36 de Fatou, $\int_A^* f \, d\mu \leq \nu(A)$,

donc f est localement μ-intégrable, et définit une mesure $f\mu = \nu_1 \geq 0$, majorée par ν et de base μ. On a, pour toute $\varphi \geq 0$ de $\mathscr{C}(X)$, $\int \varphi f \, d\mu = \lim_{n \to \infty} \int \varphi f_n \, d\mu$

(Fatou ou Lebesgue) donc $= \operatorname{Sup}_{n \in \mathbb{N}} \int \varphi f_n \, d\mu$, donc ν_1 est exactement (d'après la formule (IV,2;4I5sexto)) la borne supérieure des $f_n \mu$, donc vérifie la formule (IV,5;9quarto). Si une mesure $\lambda \geq 0$ quelconque est majorée par ν et de base μ, elle est de la forme $q\mu$; si $q_n = \operatorname{Inf}(q, n)$, le 4°/ du théorème donne $\operatorname{Inf}(\lambda, n\mu) = \operatorname{Inf}(q\mu, n\mu) = q_n \mu$; mais $\lambda \leq \nu$ entraîne $q_n \mu = \operatorname{Inf}(\lambda, n\mu) \leq \operatorname{Inf}(\nu, n\mu) \leq \nu_1$; mais, de $q_n \mu \leq \nu_1$ ou $\int \varphi q_n \, d\mu \leq \nu_1(\varphi)$ pour toute $\varphi \geq 0$ de $\mathscr{C}(X)$, on déduit, par Fatou ou Lebesgue, $\int \varphi q \, d\mu \leq \nu_1(\varphi)$ ou $\lambda = q\mu \leq \nu_1$: ν_1 est bien la plus grande mesure ≥ 0, majorée par ν et de base μ.

* Voir page 521, • cela veut dire que ν_2 est portée par un ensemble de μ-mesure nulle, ou que ν_2 et μ sont étrangères.

La mesure $3_1 = 3 - \nu$ est alors $\geqslant 0$ et majorée par ν : Elle est μ- singulière; si en effet $\theta = \mathrm{Inf}(\nu_2, \mu) \geqslant 0$, on voit que $3_1 + \theta$ est encore majorée par $\nu_1 + \nu_2 = \nu$, et encore de base μ puisque $\theta \leqslant \mu$; donc nécessairement $\nu_1 + \theta = \nu_1$ et $\theta = 0$ ν_2 et μ sont bien étrangères. Il existe alors (corollaire du théorème 50 quinto) un ensemble portant μ et de ν_2- mesure nulle; mais cet ensemble porte aussi $\int\!\mu$, donc ν_2 et $\int\!\mu = \nu_1$ sont aussi étrangères (plus généralement, ce raisonnement montre que, si deux mesures $\geqslant 0$, $\lambda_1 \lambda_2$, sont étrangères, deux mesures de bases respectives λ_1 et λ_2 sont aussi étrangères). Soit alors λ une mesure $\geqslant 0$, majorée par ν et étrangère à μ. Elle est aussi, comme nous venons de le voir, étrangère à 3_1 . De $\mathrm{Inf}(\lambda, \nu_1) = 0$ on déduit $\mathrm{Sup}(\lambda, \nu_1) = \lambda + \nu_1$; mais λ et ν_1 sont majorées par ν, donc aussi $\mathrm{Sup}(\lambda, \nu_1)$, donc $\lambda + 3_1$; donc $\lambda \leqslant \nu - \nu_1 = \nu_2$, et 3_1 est bien la plus grande mesure $\geqslant 0$ majorée par ν et étrangère à μ .

Il est alors évident que $3 = \nu_1 + \nu_2$ est la seule décomposition de ν en sommes de deux mesures $\geqslant 0$ respectivement de base μ et μ- singulière, Si en effet $\nu = \theta_1 + \theta_2$ est une telle décomposition, on a, d'après les propriétés maximales de ν_1 et ν_2 , $\theta_2 \leqslant \nu_2$, donc $\theta_1 \geqslant \nu_1$, mais aussi $\theta_1 \leqslant \nu_1$, donc $\theta_1 = \nu_1$, et $\theta_2 = \nu_2$.

Soit maintenant $\vec{\nu}$ une mesure vectorielle de base $\geqslant 0$. Nous dirons que $\vec{\nu}$ est portée par une partie A de X , si $\complement A$ est de mesure nulle pour la plus petite majorante absolue $|\vec{\nu}|$. Alors $\vec{\nu}$ est dite μ- singulière ou étrangère à μ , si elle est portée par un ensemble de μ- mesure nulle. Alors :

<u>Corollaire 2 - Soit μ une mesure $\geqslant 0$ sur X , et $\vec{\nu}$ une mesure vectorielle de base $\geqslant 0$. Elle admet une décomposition unique $\vec{\nu} = \vec{\nu}_1 + \vec{\nu}_2$, où $\vec{\nu}_1$ est de base μ , et $\vec{\nu}_2$ est μ- singulière. En outre la décomposition correspondante de la plus petite majorante absolue $|\vec{\nu}|$ est $|\vec{\nu}| = |\vec{\nu}_1| + |\vec{\nu}_2|$.</u>

<u>Démonstration</u> - $\vec{\nu}$ est de base $\geqslant 0$ soit $\vec{\nu} = \vec{q}\lambda$. soit $\lambda = \lambda_1 + \lambda_2$, la décomposition de λ par rapport à μ : λ_1 est de base μ , λ_2 est μ- singulière (corollaire). Comme λ_1 et λ_2 sont $\leqslant \lambda$, la fonction $\vec{q}$ localement λ-intégrable, est a fortiori localement λ_1- intégrable et localement λ_2- intégrable, et $\vec{q}\lambda = \vec{q}\lambda_1 + \vec{q}\lambda_2$. Mais $\vec{\lambda}_1$ est

de base μ , donc aussi $\vec{q}\lambda_1$; et λ_2 est étrangère à μ donc aussi $\vec{q}\lambda_2$, comme nous l'avons vu dans la démonstration du corollaire précédent. Nous avons donc obtenu une décomposition du type demandé. Comme $|\vec{\nu}| = \|\vec{q}\|\lambda$, $\vec{\nu}_1 = \|\vec{q}\|\lambda_1$ et $\|\vec{\nu}_2\| = \|\vec{q}\|\lambda_2$, on a évidemment $|\vec{\nu}| = |\vec{\nu}_1| + |\vec{\nu}_2|$. L'unicité revient à dire ceci : si $\vec{\nu}_1 + \vec{\nu}_2 = \vec{0}$, $\vec{\nu}_1$ de base μ, $\vec{\nu}_2$ μ-singulière $\vec{\nu}_1$ et $\vec{\nu}_2$ sont nulle% Mais si **A** est un ensemble de μ-mesure nulle, portant ν_2, **A est** de $|\vec{\nu}_1|$-mesure nulle puisque $|\vec{\nu}_1|$ est de base μ, et $\complement A$ est de $|\vec{\nu}_1| - |\vec{\nu}_2|$ - mesure nulle puisque **A** porte $\vec{\nu}_2$. Donc X est de $|\nu_1|$- mesure nulle, donc $\vec{\nu}_1 = \vec{0}$ donc aussi $\vec{\nu}_2 = \vec{0}$.

<u>Corollaire **3** - Soit $\vec{\nu}$ une mesure de base $\geqslant 0$, μ une mesure $\geqslant 0$. Pour que $\vec{\nu}$ soit de base μ , il faut et il suffit que, pour tout compact K et tout nombre $\varepsilon > 0$, il existe un nombre $\delta(K,\varepsilon) = \delta > 0$ tel que, pour A borélienne $\subset K$, $\mu(A) \leqslant \delta$ implique $\|\overrightarrow{\nu(A)}\| \leqslant \varepsilon$.</u>

<u>Démonstration</u> - Soit d'abord $\vec{\nu} = \vec{f}\mu$. Puisque $\vec{f}$ est intégrable sur K , il existe un nombre $M \geqslant 0$ tel que $\int_K (\|\vec{f}\| - \|\vec{f}\|_M)\, d\mu \leqslant \frac{\varepsilon}{2}$, où $\|\vec{f}\|_M = \mathrm{Inf}(\|\vec{f}\|, M)$ (application du théorème 35 de Lebesgue pour $\|\vec{f}\| - \|\vec{f}\|_M$, lorsque M tend vers $+\infty$). Il suffit alors de prendre $\delta = \frac{\varepsilon}{2M}$. En effet, si $\mu(A) \leqslant \delta$, on a

$$\|\vec{\nu}(A)\| = \|\int_A \vec{f}\, d\mu\| \leqslant \int_A \|\vec{f}\|\, d\mu = \int_A \|\vec{f}\|_M\, d\mu + \int_A (\|\vec{f}\| - \|\vec{f}\|_M)\, d\mu \leqslant M\delta + \frac{\varepsilon}{2} .$$

Inversement supposons la propriété de l'énoncé vérifiée. Alors tout ensemble borélien d'adhérence compacte de μ- mesure nulle est aussi de $\vec{\nu}$- mesure nulle, et le théorème de Lebesgue-Nikodym assure que $\vec{\nu}$ est de base μ .

<u>Corollaire 4</u> - <u>Soient $\vec{\nu}$ de base $\geqslant 0$, $\mu \geqslant 0$. Pour que $\vec{\nu}$, soit de base μ , il faut et il suffit que, pour tout ouvert@' d'adhérence compacte et tout $\varepsilon > 0$, il existe $\eta(\mathcal{O},\varepsilon) = \eta > 0$ tel que, pour $\varphi \in \mathcal{C}(X)$ à support dans $\mathcal{O}$, $0 \leqslant \varphi \leqslant 1$, $\mu(\varphi) \leqslant \eta$</u> implique $\|\vec{\nu}(\varphi)\| \leqslant \varepsilon$.*

* Attention, cette implication ne subsisterait pas si l'on ne supposait pas $\varphi \leqslant 1$.

L'intérêt de ce corollaire est qu'il permet de reconnaître si $\vec{\nu}$ est de base μ sans utiliser le prolongement de Lebesgue; les deux mesures sont considérées seulement comme fonctions sur $\mathcal{C}_+(X)$.

<u>Démonstration</u> - supposons $\vec{\nu} = \vec{f}\mu$. Soient δ et M les nombres déterminés par le même procédé qu'au corollaire précédent, relativement à $\mathcal{O}$ et ε : $\int_{\mathcal{O}}(\|\vec{f}\| - \|\vec{f}\|_M)\,d\mu \leq \frac{\varepsilon}{2}$. et $\delta = \frac{\varepsilon}{2M}$. Alors, si $\varphi \in \mathcal{C}(X)$ a son support dans $\mathcal{O}$, $0 \leq \varphi \leq 1$, et si $\mu(\varphi) \leq \delta$, on a bien

$$\|\vec{\nu}(\varphi)\| \leq \int \varphi \|\vec{f}\|\,d\mu = \int \|\vec{f}\|_M\,\varphi\,d\mu + \int (\|\vec{f}\| - \|\vec{f}\|_M)\varphi\,d\mu \leq M\delta + \frac{\varepsilon}{2} = \varepsilon ;$$

$\eta = \delta$ vérifie la propriété de l'énoncé pour $\mathcal{O}$ et ε .

Inversement, supposons que $\vec{\nu}$ de base ≥ 0 , vérifie la propriété de l'énoncé. Soit d'abord $\nu \geq 0$. Soit A un ensemble d'adhérence compacte tel que $\mu(A) = 0$. Soit $\mathcal{O}$ un ouvert d'adhérence compacte, contenant A . A tout $\varepsilon > 0$, l'énoncé fait correspondre un $\eta = \eta(\mathcal{O}, \varepsilon)$. Soit $\mathcal{a}$ un ouvert tel que $A \subset \mathcal{a} \subset \mathcal{O}$, $\mu(\mathcal{a}) \leq \eta$. Pour toute φ de $\mathcal{C}(X)$, $0 \leq \varphi \leq 1$, à support dans $\mathcal{a}$, on a $\mu(\varphi) \leq \mu'(\mathcal{a}) \leq \eta$, donc $\nu(\varphi) \leq \varepsilon$; en prenant la borne supérieure pour toutes les φ , on a $\nu(\mathcal{a}) \leq \varepsilon$; donc $\nu(A) \leq \varepsilon$; ε étant arbitraire, on a $\nu(A) = 0$. Donc tout borélien d'adhérence compacte de μ - mesure nulle est aussi de ν - mesure nulle; le théorème de Lebesgue-Nikodym assure que 3 est bien de base μ .

Soit maintenant 3 réelle. Soit $\eta = \eta(\mathcal{O}, \varepsilon)$ attribué par l'énoncé à $\mathcal{O}$ et ε . Si $\varphi \in \mathcal{C}(X)$, $0 \leq \varphi \leq 1$, φ à support dans $\mathcal{O}$, $\mu(\varphi) \leq \eta$, on a pour toute $\Psi \in \mathcal{C}(X), 0 \leq \Psi \leq \varphi$, l'inégalité $\mu(\Psi) \leq \eta$ donc $|\nu(\Psi)| \leq \varepsilon$; en prenant la borne supérieure pour toutes les Ψ (formule (IV,2;41)), on a $\nu^+(\varphi) \leq \varepsilon$ · de même $\nu^-(\varphi) \leq \varepsilon$. Donc ν^+ et ν^- vérifient la propriété de l'énoncé et sont ≥ 0 , donc elles sont de base μ , donc ν aussi.

Soit maintenant $\vec{\nu}$ à valeurs dans $\vec{E}$ de dimension finie. Soit (e_i) $i = 1, \ldots, N$, une base de $\vec{E}$ sur le corps $\mathbb{R}$ On a, en tant que $\mathbb{R}$ - mesure : $\vec{\nu} = \sum_{i=1}^{N} \nu_i \vec{e}_i$.

L'application " i-ième coordonnée" est $\mathbb{R}$ - linéaire, continue de $\vec{E}$ dans $\mathbb{R}$, donc il existe ε' tel que $\|\nu(\varphi)\| \leq \varepsilon'$

entraîne $|\nu_i(\varphi)| \leq \varepsilon$; alors, si $\eta = \eta(\vartheta, \varepsilon')$ est attribué a $\vec{\nu}$, ϑ, ε' , on voit que $\mu(\varphi) \leq \eta$ entrainerait $\|\vec{\nu}(\varphi)\| \leq \varepsilon'$ donc $|\nu_i(\varphi)| \leq \varepsilon$; donc chaque ν_i est de base μ et par suite aussi $\vec{\nu}$. Nous ne donnerons pas la démonstration si $\vec{E}$ est un Banach quelconque.

Application au prolongement d'une mesure à valeurs vectorielles

Jusqu'à présent nous n'avons défini d'intégrales $\int f\,d\mu$, où f n'était pas une fonction de $\mathscr{C}(X)$, que si μ était une mesure réelle ≥ 0 ou si f était borélienne bornée à support compact (page 521). Il est maintenant possible de le faire dans des cas plus généraux. Soit $\vec{\mu}$ une mesure sur X , à valeurs dans l'espace de Banach $\vec{E}$. Supposons qu'elle soit de base ≥ 0 , c'est-a-dire puisse se mettre sous la forme $\vec{p}\,\mu_0$ où μ_0 est une mesure réelle ≥ 0 , et $\vec{p}$ une fonction localement μ_0- intégrable a valeurs dans $\vec{E}$ Alors nous serons tentés de donner la définition suivant; :

<u>Une fonction f , définie sur X , à valeurs scalaires, est $\vec{\mu}$-intégrable, si et seulement si la fonction $\vec{p}f$ est μ_0- intégrable, et alors, par définition</u>

$$\int f\,d\vec{\mu} = \int (\vec{p}f)\,d\mu_0 \tag{IV,5;10}$$

De même, si μ est une mesure scalaire, elle peut s'écrire, au moins d'une manière, sous la forme $p\mu_0$, μ_0 mesure ≥ 0, p fonction sur X à valeurs scalaires, localement μ_0- intégrable , on pourra dire que la <u>fonction $\vec{f}$, définie sur X , à valeurs dans l'espace de Banach $\vec{F}$, est μ-intégrable, si et seulement si $p\vec{f}$ est μ_0- intégrable, et poser</u>

$$\int \vec{f}\,d\mu = \int (p\vec{f})\,d\mu_0 . \tag{IV,5;11}$$

On peut même définir une opération qui généralise les 2 précédentes. Soit $\vec{\mu}$ une mesure sur X à valeurs dans $\vec{E}$

exprimable sous la forme $\vec{p}\,\mu_0$, μ_0 mesure ≥ 0, $\vec{p}$ fonction sur X, à valeurs dans $\vec{E}$, localement μ_0-intégrable; soit $\vec{f}$ une fonction sur X à valeurs dans $\vec{F}$; soit B une application bilinéaire continue de $\vec{E} \times \vec{F}$ dans un espace de Banach $\vec{G}$. Peut-on donner un sens à l'expression $\int B(d\vec{\mu}, \vec{f})$?

Si $\vec{f}$ est F-étagée, c'est-à-dire de la forme $\vec{f} = \sum_i \vec{f}_i \varphi_{X_i}$ où les $\vec{f}_i$ sont des constantes, et les φ_{X_i} des fonctions caractéristiques d'ensembles X_i $\vec{\mu}$-mesurables et de p-mesure finie (c'est-à-dire tels que $\vec{\mu}(X_i) = \int \varphi_{X_i}\, d\vec{\mu}$ ait un sens d'après (IV,5;10)), on pourra poser

(IV,5;12) $$\int B(d\vec{\mu}, \vec{f}) = \sum_i B\left(\vec{\mu}(X_i), \vec{f}_i\right)$$

On dira alors, si $\vec{f}$ est quelconque, que $\int B(d\vec{\mu}, \vec{f})$ a un sens, si et seulement si la fonction $B(\vec{p}, \vec{f}) : x \to B(\vec{p}(x), \vec{f}(x))$, est μ_0-intégrable, et on posera

(IV,5;13) $$\int B(d\vec{\mu}, \vec{f}) = \int B(\vec{p}, \vec{f})\, d\mu_0 .$$

Ces définitions sont justifiées par les deux motifs suivants : d'une part, si la mesure μ est elle-même ≥ 0, le théorème 51 montre que la nouvelle définition (IV,5;11) coïncide avec l'ancienne; d'autre part :

Théorème 54 bis

Si une mesure $\vec{\mu}$ peut s'exprimer sous la forme $\vec{p}_1\,\mu_1$ et sous la forme $\vec{p}_2\,\mu_2$, où μ_1 et μ_2 sont des mesures ≥ 0, et où $\vec{p}_1$ (resp. $\vec{p}_2$) est localement μ_1-intégrable (resp. μ_2-intégrable), les définitions de $\int B(d\vec{\mu}, \vec{f})$ données par (IV,5;13) coïncident; autrement dit $\int B(\vec{p}_1, \vec{f})\, d\mu_1$ a un sens si et seulement si $\int B(\vec{p}_2, \vec{f})\, d\mu_2$ a un sens, et leurs valeurs sont égales.

Appelons en effet θ la mesure somme $\mu_1 + \mu_2$. Alors on a $\mu_1 \leq \theta$, et, d'après le théorème 53, il existe une fonction g_1 réelle ≥ 0

telle que la mesure μ_1 soit égale à $g_1\theta$; alors, d'après le corollaire du théorème 51, la mesure $\vec{p}_1\mu_1 = \vec{p}_1(g_1\theta)$ est égale à la mesure $(\vec{p}_1 g_1)\theta$.

Ensuite, d'après (IV,5;13), $\int B(d\vec{\mu}, \vec{f})$ a un sens, relativement à la représentation de $\vec{\mu}$ comme $\vec{p}_1\mu_1$, si et seulement si $\int B(\vec{p}_1, \vec{f})\, d\mu_1$ a un sens; et d'après le théorème 51, cette dernière expression a un sens si et seulement si $\int (B(\vec{p}_1, \vec{f})g_1)d\theta = \int B(\vec{p}_1 g_1, \vec{f})d\theta$ a un sens, et on a

(IV,5;14) $$\int B(d\vec{\mu}, \vec{f}) = \int B(\vec{p}_1\, g_1, \vec{f})\, d\theta .$$

Si maintenant on définit le premier membre à partir de la représentation de $\vec{\mu}$ comme $\vec{p}_2\mu_2$, $\mu_2 = g_2\theta$, on trouve :

(IV,5;15) $$\int B(d\vec{\mu}, \vec{f}) = \int B(\vec{p}_2\, g_2, \vec{f})\, d\theta .$$

Mais, comme $(\vec{p}_1\, g_1)\theta$ et $(\vec{p}_2\, g_2)\theta$ représentent la même mesure $\vec{\mu}$, il résulte du théorème 52 que $\vec{p}_1\, g_1$ et $\vec{p}_2\, g_2$ sont θ - presque partout égales, de sorte que les 2 valeurs obtenues pour $\int B(d\vec{\mu}, \vec{f})$ existent bien en même temps et sont alors les mêmes. Il en résulte que nous savons prolonger une mesure vectorielle $\vec{\mu}$, toutes les fois qu'elle est de base ≥ 0 .

<u>Corollaire</u> . <u>Si $\vec{E}$ est un espace vectoriel de dimension finie, $\vec{\mu}$ une mesure sur X à valeurs dans $\vec{E}$, de norme finie, si $\vec{f}$ est une fonction sur X à valeurs dans un espace de Banach $\vec{F}$, continue et bornée, et si B est une application bilinéaire continue de $\vec{E} \times \vec{F}$ dans un espace de Banach $\vec{G}$, $\int B(d\vec{\mu}, \vec{f})$ a un sens.</u>

Démonstration - D'après le théorème 54, $\vec{\mu} = \vec{q}\,\mu_0$, où $\mu_0 = |\vec{\mu}|$ est aussi de norme finie, et où $\|\vec{q}\| \equiv 1$.
Alors $B(\vec{q}, \vec{f})$ est μ_0 -mesurable (corollaire 3 du théorème 23 bis), et bornée, donc μ_0 -intégrable puisque μ_0 est de norme finie (corollaire 2 du théorème 39).

Application à l'intégrabilité d'une fonction par rapport à plusieurs mesures

Théorème 54 ter - Soient μ et ν deux mesures ≥ 0 sur X. Si une fonction $\vec{f}$, définie sur X, à valeurs dans un espace de Banach $\vec{F}$, est intégrable à la fois pour μ et pour3 , elle est integrable pour μ + 3 , et on a :

(IV,5;16) $$\int \vec{f}\, d(\mu+\nu) = \int \vec{f}\, d\mu + \int \vec{f}\, d\nu .$$

En effet, d'après le théorème 53, on a $\mu = \varphi(\mu+\nu)$, ν = q(μ +Y), où φ etq sont $(\mu+\nu)$ -mesurables et majorées par 1; en outre μ + 3 = $(\varphi+q)(\mu+\nu)$, donc, d'après le théorème 52, $\varphi + q$ est $(\mu+\nu)$ -presque partout égale à 1. Comme alors $\vec{f}$ est μ -intégrable, le théorème 51 nous dit que $\vec{f}\varphi$ est $(\mu+\nu)$ -intégrable, et en outre

(IV,5;17) $$\int \vec{f}\, d\mu = \int \vec{f}\varphi\, d(\mu+\nu) .$$

De même $\vec{f}q$ est $(\mu+\nu)$ -intégrable, et

(IV,5;18) $$\int \vec{f}\, d\nu = \int \vec{f} q\, d(\mu+\nu) .$$

Donc $\vec{f}(\varphi+q)$ est $(\mu+\nu)$ -intégrable, et comme $\varphi + q = 1$ $(\mu+\nu)$-presque partout, $\vec{f}$ aussi est $(\mu+\nu)$-intégrable, et une addition de (IV,5;17) et (IV,5;18) donne (IV,5;16)

<u>Exemple</u> : Soit $\vec{E}$ un espace euclidien de dimension finie sur $\mathbb{R}$, et $(\vec{X} \mid \vec{Y})$ le produit scalaire sur $\vec{E}$, forme bilinéaire continue. Alors, si $\vec{\mu}$ est une mesure sur X à valeurs dans $\vec{E}$, $\vec{f}$ une fonction sur X avaleurs dans $\vec{E}$, on pourra dire si une expression $\int (d\vec{\mu} \mid \vec{f})$ a un sens, et ce sera alors un nombre réel. Par rapport à une base orthonormale $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, on aura

$$\vec{\mu} = \sum_{i=1}^{n} \mu_i \vec{e}_i \ , \quad \vec{f} = \sum_{i=1}^{n} f_i \vec{e}_i \ ;$$

si alors

$$\sum_{i=1}^{n} \int f_i \, d\mu_i$$

a un sens, on pourra affirmer que $\int (d\vec{\mu} \mid \vec{f})$ en a un et a la même valeur. On pourra aussi affirmer que $\int (d\vec{\mu} \mid \vec{f})$ a un sens, si $\vec{\mu}$ est de norme finie, et $\vec{f}$ borélienne bornée.

<u>Remarque</u> . Nous avons supposé B bilinéaire sur $\vec{E} \times \vec{F}$ et notée $B(\vec{X},\vec{Y})$, $\vec{X} \in \vec{E}$, $\vec{Y} \in \vec{F}$, et avons alors noté l'intégrale $\int B(d\vec{\mu}, \vec{f})$ pour $\vec{\mu}$ a valeurs dans $\vec{E}$ et $\vec{f}$ à valeurs dans $\vec{F}$. Si B est noté d'une autre manière, on note l'intégrale de façon correspondante. C'est ce que nous avons fait avec le produit scalaire; il était noté $(\vec{X} \mid \vec{Y})$, donc nous avons écrit $\int (d\vec{\mu} \mid \vec{f})$; on pouvait aussi écrire le même produit scalaire $(\vec{Y} \mid \vec{X})$, donc noter la même intégrale $\int (\vec{f} \mid d\vec{\mu})$. Mais supposons que $\vec{E} = \vec{F} = \mathbb{R}^3$, et soit B le produit vectoriel, application bilinéaire continue $(X,Y) \longrightarrow X \wedge Y$ de $\mathbb{R}^3 \times \mathbb{R}^3$ dans $\mathbb{R}^3$. Alors $(\vec{X},\vec{Y}) \longrightarrow \vec{Y} \wedge \vec{X}$ est l'application bilinéaire opposée. On aura donc deux vecteurs opposés : $\int d\vec{\mu} \wedge \vec{f}$ et $\int \vec{f} \wedge d\vec{\mu}$. On fait, en physique, un usage constant de telles intégrales.

Dualité entre L^p et $L^{p'}$

Soit X un espace localement compact muni d'une mesure de Radon $\mu \geq 0$. Nous abrégerons une fois pour toutes par L^p l'espace $L^p(X ; \mu)$; d'autre part, pour $1 \leq p \leq +\infty$, nous poserons, comme au corollaire 1 du théorème 46, $p' = \frac{p}{p-1}$, de sorte que l'on a $\frac{1}{p} + \frac{1}{p'} = 1$. Soit $\dot{h} \in L^{p'}$ (nous l'appelons h, c'est une classe d'équivalence, donc un ensemble de fonctions $h \in \mathcal{L}^{p'}$ deux a deux presque partout égales). Elle définit alors une forme linéaire continue sur L^p d'après la formule

$$(\mathrm{IV},5;19) \qquad \langle u_{\dot{h}}, \dot{f} \rangle = \int_X h(x) f(x)\, d\mu(x) \quad , f \in \dot{f} \;;$$

en effet le deuxième membre a un sens d'après le corollaire 1 du théorème 46, de sorte que $u_{\dot{h}}$ est bien une forme linéaire sur L^p ; sa continuité résulte de l'inégalité de Hölder, et l'on a

$$(\mathrm{IV},5;20) \qquad \| u_{\dot{h}} \| = \frac{\mathrm{Sup}\, |\langle u_{\dot{h}}, \dot{f} \rangle|}{N_p(\dot{f})} \leq N_{p'}(\dot{h})$$

L'application $\dot{h} \longrightarrow u_{\dot{h}}$ est alors une application linéaire de $L^{p'}$ dans le dual $(L^p)'$ de L^p .

<u>Théorème **54** quarto</u> - <u>L'application $\dot{h} \longrightarrow u_{\dot{h}}$ précédemment définie est une bijection isométrique de $L^{p'}$ sur $(L^p)'$ pour $1 \leq p < +\infty$; pour $p = +\infty$, c'est une injection isométrique de $L^{p'}$ dans $(L^p)'$, qui n'est jamais surjective si le support de μ contient une infinité de Points.</u>

Avant de démontrer ce théorème, précisons bien ce qu'il signifie : nous entendons par isométrie que l'on a exactement $\| u_{\dot{h}} \| = N_{p'}(\dot{h})$. Cela signifie aussi naturellement que $\dot{h} \longrightarrow u_{\dot{h}}$ est une injection, puisque l'on a $u_{\dot{h}} = 0$ si et seulement si $\dot{h} = 0$. D'autre part quand nous disons que c'est une surjection pour $p < +\infty$, nous voulons dire que toute forme linéaire continue sur L^p peut s'exprimer (d'une manière unique) comme une $u_{\dot{h}}$, avec $\dot{h} \in L^{p'}$. **On** peut donc identifier le dual $(L^p)'$ de L^p avec $L^{p'}$, pour $p < +\infty$; mais L^1 est seulement en général un sous-espace du dual de L^∞ (bien distinguer: le dual de L^1 est L^∞ , le dual de L^∞ n'esp pas L^1).

<u>Démonstration</u> - Nous allons d'abord montrer l'isométrie (donc l'injectivité), pour $p = +\infty$, de $\dot{h} \longrightarrow u_{\dot{h}}$. Appelons f la fonction $\frac{h}{|h|}$ (en lui donnant par exemple la valeur 1 là où $h = 0$). On a $f \in \mathcal{L}^\infty$, et $|f| = 1$. On a

les inégalités

$$\| u_{\dot{h}} \| \geqslant \frac{< u_{\dot{h}} , \dot{f} >}{N_\infty(\dot{f})} = \frac{\int h \frac{\bar{h}}{|h|} d\mu}{1} \geqslant \int |h| d\mu = N_1(\dot{h})$$

qui donnent bien $\| u_{\dot{h}} \| \geqslant N_1(\dot{h})$

donc $= N_1(\dot{h})$ ce qui prouve l'isométrie pour $p = +\infty$
On pourrait faire de même pour $p < +\infty$, mais là nous montreront à la fois la surjectivité et l'isométrie. Soit donc $1 \leq p < +\infty$ et soit u une forme linéaire continue sur L^p, avec

(IV,5;21) $$| < u, \dot{f} > | \leqslant \|u\| N_p(\dot{f}) .$$

Nous considérerons que u définit une forme linéaire sur $\mathcal{L}^p$, nulle sur toutes les fonctions μ- presque partout nulles. Elle définit alors à fortiori une forme linéaire sur $\mathcal{C}(X)$ qui est manifestement continue sur tout $\mathcal{C}_K(X)$, K compact de X ; car, si φ converge vers 0 dans $\mathcal{C}_K(X)$, elle converge a fortiori vers 0 dans $\mathcal{L}^p$. Donc u définit une mesure de Radon 3 sur X .

si $p = 1$, $p' = +\infty$, on a tout de suite

(IV,5;22) $$| \nu(\varphi) | = | < \mu , \varphi > | \leq \|u\| N_1(\varphi) = \|u\| \mu (|\varphi|);$$

donc ν est absolument majorée par $\|u\| \mu$, donc, d'après le théorème 52-53-54, il existe h, μ- mesurable $|h| \leqslant \|u\|$ donc $N_\infty(h) \leq \|u\|$, telle que $3 = h\mu$. Mais même si $1 < p < +\infty$ pour montrer que 3 est de base μ, nous pourrons appliquer le théorème 52 de Lebesgue Nikodym. La mesure ν admet, comme nous l'avons, signalé page un prolongement borélien à $\Gamma(X)$. Mais par ailleurs u admet un prolongement à $\mathcal{L}^p$: or $\Gamma(X) \subset \mathcal{L}^p$. Il est donc normal de se demander si le prolongement de ν à $\Gamma(X)$ est la restriction de u à $\Gamma(X)$ •,il suffit pour cela de démontrer que la restriction de u à $\Gamma(X)$, qui est continue pour la topologie induite par $\mathcal{L}^p$, est aussi L-continue. Il est d'abord évident que, si une suite de fonctions $f_n \in \Gamma(X)$ convergent simplement vers 0 en restant bornées et en gardant leur support dans un compact fixe, elles convergent vers 0 dans $\mathcal{L}^p$ (théorème de Lebesgue particulier 34); u étant continue sur $\mathcal{L}^p$ les $< u, f_n >$ convergent bien vers 0 . Par ailleurs soit $\mathcal{O}$ un ouvert de X d'adhérence compacte et soit $\varepsilon > 0$; on peut alors trouver un compact **K** tel que $\mu(\mathcal{O} - K) \leqslant \left(\frac{\varepsilon}{\|u\|}\right)^{p}$ (<u>là, il est essentiel que $p < +\infty$</u>); dans ces conditions, Pour toute fonction φ de $\mathcal{C}(X)$ $0 \leqslant \varphi \leqslant 1$, égale à 1 sur un voisinage de K et à support dans $\mathcal{O}$.

$$\left|\langle u, \chi_{\mathcal{O}}\rangle - \langle u, \varphi\rangle\right| \leqslant \|u\| N_p(\chi_{\mathcal{O}} - \varphi) \leqslant \|u\| N_p(\chi_{\mathcal{O}-K}) = \|u\| (\mu(\mathcal{O}-K))^{\frac{1}{p}} \leqslant \varepsilon,$$

ce qui prouve bien la L – continuité de u sur $\Gamma(X)$, et par conséquent l'identité de u et du prolongement borélien de ν sur $\Gamma(X)$. Mais alors, si A est un ensemble borélien d'adhérence compacte et de μ-mesure nulle, sa fonction caractéristique est μ- presque partout nulle, et la valeur de u sur cette fonction caractéristique est donc nulle; on a donc aussi ν **(A) = 0**. D'après le théorème de Lebesgue-Nikodym (théorème 52), la mesure ν est donc de base μ : il existe bien une fonction h, localement μ- intégrable, telle que $\nu = h\mu$.

Soit **M** un nombre $\geqslant 0$ et **K** un compact de X Appelons $h_{M,K}$ la fonction, définie comme page **474** lorsque h était3 $0^{,K}$, par $h_{M,K}$ est égale à h en tout point de K pour lequel $|h(x)| \leqslant M$, et à 0 partout ailleurs.

Soit alors f la fonction $|h_{M,K}|^{p'-2}\,\overline{h_{M,K}}$, nulle là où $h_{M,K}$ est nulle; son module est $|h_{M,K}|^{p'-1}$, et $hf = |h_{M,K}|^{p'}$. Comme nous avons vu que toute fonction mesurable est μ- presque partout égale à une fonction borélienne, nous pouvons, toujours, quitte à changer f sur un ensemble de μ- mesure nulle, supposer f borélienne, et par conséquent appartenant à $\Gamma(X)$, car $|f| \leqslant M^{p-1}$. On a alors les inégalités

$$(\text{IV},5;23)\quad \left\{\begin{aligned} \|u\| &\geqslant \frac{|\langle u, f\rangle|}{N_p(f)} = \frac{\left|\int f h\, d\mu\right|}{N_p(f)} \\ &= \frac{\int |h_{M,K}|^{p'} d\mu}{\left(\int |h_{M,K}|^{(p'-1)p} d\mu\right)^{\frac{1}{p}}} = \left(\int |h_{M,K}|^{p'} d\mu\right)^{\frac{1}{p'}} = N_{p'}(h_{M,K}) \end{aligned}\right.$$

Ceci étant vrai quels que soient M et K, on a, en prenant la borne supérieure :

$$(\text{IV},5;24)\qquad N_{p'}(h) \leqslant \|u\|, \quad \text{et} \quad h \in \mathcal{L}^{p'}$$

Alors la forme linéaire initiale u, et la forme u_h

définie a partir de $h \in \mathcal{L}^{p'}$ coïncident sur $\mathcal{C}(X)$, qui est dense dans $\mathcal{L}^{p}$ (encore une fois parce que $p < +\infty$; théorème 49) et par conséquent coïncident sur $\mathcal{L}^{p}$. Cela prouve la surjectivité, en même temps les indgalités (IV,5;20) et (IV,5;24) prouvent l'isométrie $\|u_h\| = N_{p'}(h)$, ce qui achève la démonstration des parties positives du théorème. Quant à la partie négative, a savoir que, toutes les fois que le support de μ contient une infinité de points, l'application $h \longrightarrow u_h$ n'est pas surjective, et que par conséquent le dual de L^{∞} ne peut pas être identifié à L^1, nous l'admettrons sans démonstration.

Remarque : Soit maintenant $\vec{E}$ un Banach quelconque. Tout élément $\overleftarrow{h}$ de $L^{p'}(\overleftarrow{E'})$ définit alors une forme linéaire continue $u_{\overleftarrow{h}}$ sur $L^{p}(\vec{E})$ par la formule

(IV,5;25) $$\langle u_{\overleftarrow{h}}, \vec{f} \rangle = \int \langle \overleftarrow{h}, \vec{f} \rangle \, d\mu$$

et en outre on a, par la même inégalité de Hölder, l'inégalité $\|u_{\overleftarrow{h}}\| \leq N_{p'}(\overleftarrow{h})$.

On définit donc ici encore une application $\overleftarrow{h} \longrightarrow u_{\overleftarrow{h}}$ de $L^{p'}(\overleftarrow{E'})$ dans le dual $(L^{p}(\vec{E}))'$ de $L^{p}(\vec{E})$. On peut encore montrer que, si $\vec{E}$ est de dimension finie et $p < +\infty$, cette application est une bijection isométrique, ce qui permet encore d'identifier le dual de $L^{p}(\vec{E})$ à $L^{p'}(\overleftarrow{E'})$. Nous l'admettrons sans démonstration. Par contre ce n'est pas exact sans autre hypothèse, si $\vec{E}$ est un Banach de dimension infinie.

§ 6 IMAGE D'UNE MESURE PAR UNE APPLICATION

Si $\vec{\mu}$ est une mesure définie sur l'espace localement compact X , à valeurs dans un espace vectoriel normé $\vec{E}$, et si L est une application linéaire continue de $\vec{E}$ dans un espace vectoriel normé $\vec{F}$, on définit de toute évidence une nouvelle mesure $L \circ \vec{\mu}$ par la formule

$$(L \circ \vec{\mu})(\varphi) = L(\vec{\mu}(\varphi)) \tag{IV,6;1}$$

En effet, $L \circ \vec{\mu}$ est trivialement une application linéaire de $\mathscr{C}(X)$ dans $\vec{F}$; et, si φ a son support dans le compact K , on a la majoration

$$\|(L \circ \vec{\mu})(\varphi)\| \leqslant \|L\| \, \|\vec{\mu}\|_K \, \|\varphi\| \text{ ; d'où } \|L \circ \vec{\mu}\|_K \leqslant \|L\| \, \|\vec{\mu}\|_K \, . \tag{IV,6;2}$$

Cette mesure s'appelle image directe de $\vec{\mu}$ par l'application linéaire continue L .

Mais nous voulons parler d'une image d'une autre nature et qui joue un rôle extrêmement important dans la pratique. Soit $\vec{\mu}$ une mesure, a valeurs dans $\vec{E}$, sur l'espace localement compact X ; et soit H une application de X dans un espace localement compact Y . Nous allonsmontrer que, <u>si H vérifie des propriétés convenables, il existe une mesure image notée $H\vec{\mu}$, qui est une mesure sur l'espace localement compact Y , encore à valeurs dans $\vec{E}$</u>.

On sait que, si φ appartient a $\mathscr{C}(Y)$, elle admet par H une image réciproque $H^*\varphi$, qui est une fonction sur X définie par la formule

$$H^*\varphi = \varphi \circ H \text{ , ou } (H^*\varphi)(x) = \varphi(H(x)) \, . \tag{IV,6;3}$$

Nous sommes donc naturellement amenés à définir l'image directe $H\vec{\mu}$ par la formule

$$(H\vec{\mu})(\varphi) = \vec{\mu}(H^*\varphi) = \int \varphi(H(x)) \, d\vec{\mu}(x) \, . \tag{IV,6;4}$$

Mais cette définition ne possèdera un sens que si le deuxième membre de (IV,6;4) possède un sens, et si en outre il définit bien une mesure sur Y à valeurs dans $\vec{E}$.

Nous allons examiner successivement deux cas où la chose se produit :

1er cas : H est une application continue propre.

Définition - On dit qu'une application continue H d'un espace localement compact X dans un espace localement compact Y est propre, ou continue à l'infini, si l'image réciproque par H de tout compact de Y est un compact de X.

Si X est compact, toute application continue H de X dans Y est propre. Car tout compact de Y est fermé, donc son image réciproque par H est fermée dans X, et comme X est compact, elle est compacte (théorèmes 21 et 22 du Chapitre II).

Si X n'est pas compact, une application constante de X dans Y n'est jamais propre. Dans le plan $\mathbb{R}^2$, la projection orthogonale sur l'un des axes de coordonnées n'est pas une application propre.

Théorème 55 - Si X et Y sont des espaces métriques, dans lesquels toutes les boules fermées sont compactes, alors une application continue H est propre, si et seulement si l'image réciproque par H de toute partie bornée de Y est une partie bornée de X.

On sait en effet que l'image réciproque d'un fermé est fermée, puisque H est continue, et d autre part que les compacts sont les fermés bornés, puisque les boules fermées sont compactes * .

Corollaire - Si X et Y sont des espaces métriques où les boules fermées sont compactes, une application continue H de X dans Y est propre, si et seulement si l image par H de toute suite de points s'éloignant indéfiniment dans X est une suite de points s'éloignant indéfiniment dans Y.

Dans un espace métrique, on dit qu'une suite de points x_n s'éloignent indéfiniment pourn infini, si la distance de x_n a un point fixe a tend vers t ∞, auquel cas la distance a n'importe quel autre point fixe b tend aussi vers + ∞, puisque $d(b, x_n) \geq d(a, x_n) - d(a, b)$.

* - Une partie compacte d'un espace métrique quelconque est fermée et bornée. Inversement, dans un espace métrique où les boules fermées sont compactes, une partie fermée bornée est contenue dans une boule fermée, c'est-à-dire dans un compact, elle y est fermée, donc elle est compacte. C'est la démonstration de 1° et 2" a) du théorème 23 du chapitre II.

Démonstration 1°/ Supposons d'abord H propre. Alors, si une suite de points x_n s'éloignent indéfiniment sur X, les $H(x_n)$ s'éloignent indéfiniment sur Y, sans quoi on pourrait en trouver une suite partielle qui resterait contenue dans une boule, c'est-à-dire dans une partie bornée de Y, alors, H étant propre, la suite partielle correspondante des x_n devrait rester bornée, d'après le théorème, ce qui est contraire à l'hypothèse, puisque les x_n s'éloignent indéfiniment.

2°/ Inversement, si cette propriété est réalisée, c'est-à-dire si H transforme toute suite s'éloignant indéfiniment dans X en une suite s'éloignant indéfiniment dans Y, alors nécessairement H est propre. En effet, si B est une partie bornée de Y son image réciproque est nécessairement bornée; sans quoi'on pourrait en extraire une suite de points s'éloignant indéfiniment, dont les images devraient alors aussi s'éloigner indéfiniment, ce qui est contraire à l'hypothèse puisque ces Images sont dans B. Donc H est propre, d'après le théorème.

Exemple - Les applications $x \to x^n$, n entier $\geqslant 1$ quelconque, sont des applications propres de la droite réelle $\mathbb{R}$ dans elle-même.

C'est à cause de ce corollaire qu'une application propre peut s'appeler "continue à l'infini". Si on ajoute à X un "point à l'infini", de même à Y, et qu'on prolonge H en disant que l'image par H du point à l infini de X est le point à l'infini de Y, alors H est propre si et seulement si l'application ainsi prolongée est continue au point à l'infini.

Théorème 56 - Si H est une application propre d'un espace localement compact X dans un espace localement compact Y, l'image directe par H de toute partie fermée de X est une partie fermée de Y *.

* Si H est continue, l'image réciproque d'un fermé est fermée, et l'image directe d un compact est compacte. Si en outre H est propre, l'image réciproque d'un compact est compacte, et l'image directe d'un fermé est fermée.

Démonstration Soit A une partie fermde de X, posons $B = H(A)$; nous voulons montrer que B est fermée. Soit donc& un point de Y adhérent à B, et nous allons montrer que b appartient'8 B Appelons β une boule compacte de centre b. Comme b est adhérent à B, cette boule rencontre nécessairement B, et en outre b est adhérent à l'intersection $\beta \cap B$ *. Mais, comme H est supposée propre, l'image réciproque $H^{-1}(\beta)$ est une partie compacte de X • elle coupe donc A suivant *une* partie compacte K. Alors: si x appartient à K, son image appartient nécessairement à la fois à β et à B, donc à $\beta \cap B$. Inversement, si un point y appartient à $\beta \cap B$, comme il appartient à B, il est l'image d'au moins un point x_0 de A, et comme il appartient à β, x_0 appartient à $H^{-1}(\beta)$, donc à K. Donc l'image $H(K)$ est exactement $\beta \cap B$.

Mals, comme H est continue et que K est compact, on en déduit que $\beta \cap B$ est un compact, donc est fermé. Alors b, qui lui est adherent, lui appartient nécessairement, et b appartient à B, ce qu'il fallait démontrer.

Les applications continues propres vont nous donner un cas où l'image directe $H\mu$ a un sens.

Théorème 57 - Si X et Y sont des espaces localement compacts, $\vec{E}$ un espace vectoriel normé, $\vec{\mu}$ une mesure sur X à valeurs dans $\vec{E}$, de support A, et si H est une application continue de X dans Y, dont la restriction à A est une application propre de A dans Y, alors la formule (IV,6;4) définit $H\vec{\mu}$ comme une mesure sur Y, appelée image directe de $\vec{\mu}$ par H.

Naturellement ces conditions sont a fortiori vérifiées, si H est-continue et propre de X dans Y, ou si H est continue et μ à support compact.

Démonstration - Soit en effet φ une fonction de $\mathscr{C}(Y)$, de support compact K. Son image réciproque $H^*\varphi = \varphi \circ H$ est continue, comme composée de deux applications continues. Cherchons quel est son support. Soit Ω l'ensemble des points y tel que $\varphi(y) \neq 0$. On sait que $\bar{\Omega} = K$, par définition du support de φ. Alors l'ensemble des points x où la fonction $H^*\varphi$ est $\neq 0$ est exactement l'image réciproque $H^{-1}(\Omega)$ Comme l'image réciproque $H^{-1}(K)$ est nécessairement fermke et contient $H^{-1}(\Omega)$, on voit que le support de $H^*\varphi$ est nécessairement contenu dans $H^{-1}(K)$.

* Rappelons que b est adhérent à un ensemble, si toute boule de centre b coupe l'ensemble.

Naturellement cette image réciproque $H^{-1}(K)$ n'est pas nécessairement compacte. Elle le sera sûrement si l'application H de X dans Y est propre, mais cette hypothèse est un peu trop restrictive. Nous supposons seulement que la restriction de H à A est propre; alors cela signifie que l'intersection de $H^{-1}(K)$ et de A est une partie compacte de A. Mais alors nous avons vu au théorème 16 qu'il est possible de donner un sens à $\vec{\mu}(H^*\varphi)$, lorsque $H^*\varphi$ est une fonction scalaire continue et que l'intersection du support de $\vec{\mu}$ et du support de $H^*\varphi$ est un compact. Nous voyons bien que l'expression qui est au deuxième membre de (IV,6;4) possède un sens, et dépend linéairement de φ . Si maintenant φ garde son support dans un compact fixe K de Y et converge uniformément vers 0 , alors $H^*\varphi$ garde son support dans un ensemble fermé fixe $H^{-1}(K)$, dont l'intersection avec A est compacte, et elle converge uniformément vers 0 ; il résulte de ce qui a été vu au théorème 16 que $\vec{\mu}(H^*\varphi)$ converge alors vers 0 . Ceci démontre que le deuxième membre de, (IV,6;4) définit bien une mesure sur Y à valeurs dans E, et par conséquent $H\mu$ existe *

<u>Théorème 58 - 1°/ Dans les conditions du théorème 57 on a l'inégalité relative aux normes</u> :

$$\|H\vec{\mu}\| \leq \|\vec{\mu}\| \leq +\infty \tag{IV,6;5}$$

2°/ <u>Si μ est complexe (resp. réelle, resp. réelle ≥ 0</u>), <u>il en est de même de $H\mu$.</u>

3°/ <u>Le support de $H\vec{\mu}$ est contenu dans l'image directe par H du support de $\vec{\mu}$</u>

4°/ <u>Si $\vec{\mu}_1$ et $\vec{\mu}_2$ ont leurs supports dans un ensemble fermé A tel que la restriction de H à A soit propre, on a</u> :

$$\begin{cases} H(\vec{\mu}_1+\vec{\mu}_2) = H\vec{\mu}_1 + H\vec{\mu}_2 \\ H(k\vec{\mu}) = k(H\vec{\mu}) \ ; \end{cases} \tag{IV,6;6}$$

* Nous ne donnons des exemples que plus loin (page 546); on aura Intérêt à les regarder dès maintenant.

autrement dit, l'application $\vec{\mu} \longrightarrow H\vec{\mu}$ est linéaire * .

Démonstration 1°) Supposons $\varphi \in \mathcal{C}(Y)$, telle que $|\varphi| \leq 1$. Alors on a aussi $|H^*\varphi| \leq 1$. On en déduit (a cause de (IV,2;33 bis)) la majoration

$$(IV,6;7) \qquad \|H\vec{\mu}(\varphi)\| = \|\vec{\mu}(H^*\varphi)\| \leq \|\vec{\mu}\| \quad ;$$

qui prouve l'inégalité (IV,6;5)

2°) Evident.

3°) Soit A le support de $\vec{\mu}$ dans X , et soit B son Image H(A) . D'après le théorème 56, B est fermé . Soit alors φ une fonction de $\mathcal{C}(Y)$, de support K sans point commum avec B . Alors le support de $H^*\varphi$ qui est contenu dans l'image réciproque $H^{-1}(K)$ par H du'support de φ , ne peut pas rencontrer A ; par suite $\vec{\mu}(H^*\varphi)$ est nulle, donc $(H\vec{\mu})(\varphi)$ est nulle. Autrement dit $H\vec{\mu}$ est nulle dans le complémentaire de B , et son support est bien contenu dans B . Ce support peut naturellement être strictement plus petit que B . Néanmoins on peut facilement montrer que, si μ est réelle ≥ 0 , le support de $H\mu$ est exactement B ** .

4") est évident.

2ème cas - La mesure μ est réelle ≥ 0 .

Théorème 59 - Soient X et Y des espaces localement compacts dénombrables à l'infini. Soit μ une mesure réelle ≥ 0 sur X. Soit H une application de X dans y, ~-mesurable, et telle que l'image réciproque par H de tout compact de Y ait une mesure finie relativement a μ . Dans ces conditions, (IV,6;4) définit $H\mu$ comme une mesure' ≥ 0 sur Y . En outre, si K est un compact de Y , on a la majoration

* Noter que H est continue et propre, mais n'a aucun caractère de linéarité; d'ailleurs X et Y , espaces localement compacts, n'ont rien de vectoriel. Mais $\vec{\mu} \to$ t-i $\vec{\mu}$ est linéaire !

** Cela résulte du corollaire 1 du théorème 60·Soit en effet $B_1 \subset B$ le support de $H\mu$. Si $B_1 \neq B$ alors $\complement B_1$ est *un ouvert* de H p-mesure nulle ,rencontrant B ; 'donc $H^{-1}(\complement B_1)$ sera un ouvert de p-mesure nulle, rencontrant le support A de μ , ce qui est absurde.

(IV,6;8) $$\|H\mu\|_K \leq \int_{H^{-1}(K)} d\mu \text{ , d'où } \|H\mu\| \leq \|\mu\|$$

Nous dirons que l'application H est μ-propre, si elle vérifie les conditions précédentes. Ces conditions sont toujours vérifiées si H est mesurable, et μ de norme finie.

Démonstration - Tout d'abord, l'affirmation de la fin est évidente : si H est mesurable, et μ de norme finie, c'est-à-dire si $\mu(X) < +\infty$, alors, pour tout compact K, $H^{-1}(K)$ est mesurable et de mesure finie.

Soit maintenant φ une fonction de $\mathcal{C}(Y)$; l'image réciproque $H^*\varphi = \varphi \circ H$ est, d'après le théorème 22, une fonction complexe μ-mesurable sur X. Par ailleurs, elle est bornée par $|||\varphi|||$. Comme alors l'ensemble des points où elle est $\neq 0$ est certainement contenu dans l'image réciproque par H du support K de φ, c'est-à-dire dans un ensemble mesurable de mesure finie relativement à μ, on voit que $\int^*|(H^*\varphi)|\,d\mu$ est nécessairement finie. Il résulte alors du théorème 39 que $H^*\varphi$ est μ-intégrable. On a donc pu donner un sens au deuxième membre de (IV,6;4). Ce deuxième membre dépend linéairement de φ, et il est ≥ 0 si $\varphi \geq 0$.

D'autre part, si φ a son support dans K, $|||H^*\varphi|||$ est égal à $|||\varphi|||$, et alors on a la majoration

(IV,6;9) $$|(H\mu)(\varphi)| \leq |||\varphi||| \int_{H^{-1}(K)} d\mu \quad ,$$

ce qui montre que $H\mu$ est continue sur $\mathcal{C}_K(Y)$, donc est une mesure, et donne la majoration (IV,6;8).

Théorème 60 -

Dans les conditions du théorème 59, soit $\vec{f}$ une application de Y dans un espace de Banach $\vec{F}$. Pour que $\vec{f}$ soit $H\mu$-intégrable (resp. $H\mu$-mesurable) il faut et il suffit que l'image réciproque $H^*\vec{f} = \vec{f} \circ H$ soit une fonction μ-intégrable sur X à valeurs dans $\vec{F}$ (resp. μ-mesurable), et dans ce cas on a l'égalité

(IV,6;10) $$\int \vec{f}\, d(H\mu) = \int (H^*\vec{f})\, d\mu = \int \vec{f}(H(x))\, d\mu(x).$$

Nous admettrons ce théorème sans démonstration.

On passe naturellement par l'égalité suivante, relative à des fonctions f réelles $\geqslant 0$:

$$\int^* f\, d(H\mu) = \int^* (H^* f)\, d\mu \leqslant +\infty . \tag{IV,6;11}$$

On en déduit facilement :

Corollaire 1 - Dans les conditions du théorème 59, soit B une partie de Y . Alors B est $H\mu$ -mesurable, si et seulement si son Image réciproque $H^{-1}(B)$ est μ -mesurable, et dans ce cas on a l'égalité

$$H\mu(B) = \mu\left(H^{-1}(B)\right) \leqslant +\infty . \tag{IV,6;12}$$

Corollaire 2 - On a

$$\| H\mu \| = \| \mu \| \leqslant +\infty . \tag{IV,6;13}$$

Il suffit en effet d'appliquer le corollaire 1 à $B = Y$.

Théorème 61 (Transitivité des images de mesures).

Soient X, Y, Z trois espaces localement compacts, H une application de X dans Y , K une application de Y dans Z , $\vec{\mu}$ une mesure sur X , à valeurs dans $\vec{E}$.

1°/ Si H et K sont continues, et si l'application composée $K \circ H$ de X dans Z, restreinte au support de $\vec{\mu}$, est propre, alors $K(H\vec{\mu})$ et $(K \circ H)\vec{\mu}$ ont un sens suivant le théoreme 57, et sont égales.

2°/ Si μ est réelle $\geqslant 0$, si X , Y, Z , sont dénombrables à l'infini, et si $K(H\mu)$ existe suivant le théorème 59, alors $(K \circ H)(\mu)$ existe en vertu du théorème, et on a encore l'identité $K(H\mu) = (K \circ H)\mu$.

Démonstration -

1er cas : Soit A le support de $\vec{\mu}$, soit $B = H(A)$. Soit d'abord C' une partie compacte de Z . Alors l'intersection $H^{-1}(K^{-1}(C')) \cap A$ a pour fmage par H exactement l'intersection $K^{-1}(C') \cap B$. Comme la première est compacte, puisque $K \circ H$ est supposée propre sur A, la deuxième est donc compacte comme image directe d'un compact par une application continue: cela prouve que la restriction de K à B est propre.

soit maintenant B' une partie compacte de Y, et soit C' son image par K. Alors l'intersection $H^{-1}(B') \cap A$ est contenue dans l'intersection $H^{-1}(K^{-1}(C') \cap A$, qui est supposée compacte. Ainsi l'image réciproque de tout compact de Y par la restriction de H à A est un ensemble fermé contenu dans un compact de X, il est donc lui aussi compact, ce qui <u>prouve que H est propre sur le support A de $\vec{\mu}$</u>. Alors $H\vec{\mu}$ a un sens, et son support est dans B (théorème 58,3°), donc K, propre sur B, est propre sur ce support, donc $K(H\vec{\mu})$ a un sens, suivant le théorème 57. Soit alors φ une fonction de $\mathscr{C}(Z)$; on a les égalités

$$(IV,6;14)\quad \left\{\begin{aligned} &\left((K\circ H)\vec{\mu}\right)(\varphi) = \int \varphi\left(K(H(x))\right)\,d\vec{\mu}\,(x) \\ &\left(K(H\vec{\mu})\right)(\varphi) = \int \varphi\left(K(y)\right)\,d\left(H\vec{\mu}(y)\right) = \int \varphi(K(H(x)))\,d\vec{\mu}\,(x). \end{aligned}\right.$$

On en déduit bien $(K\circ H)\vec{\mu} = K(H\vec{\mu})$.

<u>2ème cas</u> : $\mu \geq 0$.

Supposons donc que H soit μ-propre, et que K soit $(H\mu)$-propre. Soit C' une partie fermée de Z ; alors, l'image réciproque $K^{-1}(C')$ est $H\mu$-mesurable, puisque l'application K est $H\mu$-mesurable; mais alors, d'après le corollaire 1 du théorème 60, son image réciproque $H^{-1}(K^{-1}(C'))$ est nécessairement μ-mesurable, puisque H est μ-propre, et cela démontre que $K\circ H$ est une application μ-mesurable.

En outre, si C' est compacte, alors $K^{-1}(C')$ est de $H\mu$-mesure finie, puisque K est $H\mu$-propre; il en résulte que $H^{-1}(K^{-1}(C'))$ est de μ-mesure finie, puisque H est μ-propre; ce qui prouve que $K\circ H$ est μ-propre. soit alors φ un élément de $\mathscr{C}(Z)$.

Par définition de $(K\circ H)\mu$, on a :

$$(IV,6;15)\qquad \left((K\circ H)\mu\right)(\varphi) = \int \varphi\left(K(H(x))\right)d\mu(x)\ ;$$

alors d'après le théorème 60, cette dernière expression peut s'écrire

$$(IV,6;16)\qquad \int \varphi\left(K(y)\right)d\left(H\mu(y)\right) \text{ ou } \left(K(H\mu)\right)(\varphi)\ ,$$

ce qui démontre que $(K\circ H)\mu = K(H\mu)$

Remarque - Les conditions d'application de ces deux cas sont extrêmement différentes. Pour le premier cas, on suppose une condition réalisée pour $K\circ H$, et non pas successivement pour H et pour K, alors qu'au contraire, dans le deuxième cas, on suppose des conditions réalisées successivement pour H et pour K, et non pour $K\circ H$. On démontre facilement que ces genres de conditions ne peuvent pas être intervertis. Par exemple, dans le premier cas, il peut arriver que H soit continue,et propre sur le support de $\vec{\mu}$, et que K soit continue,et propre sur le support de $H\vec{\mu}$, sans que $K\circ H$ soit propre sur le support de $\vec{\mu}$ * ; dans le deuxième cas, il peut très bien arriver que $K\circ H$ soit μ-mesurable, sans que pour cela H soit μ-mesurable ** .

Cas où H est un homéomorphisme de X sur Y

Si H est un homéomorphisme, il vérifie évidemment toutes les conditions permettant de définir $H\vec{\mu}$ pour une mesure $\vec{\mu}$ portée par X.

Mais dans ce cas, en outre, H^{-1} existe et est lui aussi un homéomorphisme, et l'on a $H^{-1}(H\vec{\mu})=(H^{-1}\circ H)\vec{\mu}=\vec{\mu}$ (théorème 61).

D'ailleurs, dans ce cas, on peut aussi définir l'image réciproque par H d'une mesure $\vec{\nu}$ portée par Y, en posant $H^{*}\vec{\nu}=H^{-1}\vec{\nu}$, et l'image directe par H d'une fonction φ appartenant à $\mathscr{C}(X)$, en posant $H\varphi=(H^{-1})^{*}\varphi=\varphi\circ H^{-1}$.

On a alors naturellement, pour toute mesure $\vec{\mu}$ sur X et pour toute fonction φ de $\mathscr{C}(X)$, les relations

(IV,6;17)
$$(H\varphi)(y)=\varphi(H^{-1}(y))\ ;\ H\varphi(H(x))=\varphi(x)\ ;\ x\in X\ ,\ y\in Y.$$
$$H\vec{\mu}\ (\psi)=\vec{\mu}(H^{-1}(\psi))\ ;\quad H\vec{\mu}\ (H\varphi)=\vec{\mu}(\varphi)\ ;\ \varphi\in\mathscr{C}(X),\ \psi\in\mathscr{C}(Y).$$

* Prenons, par exemple, $X=Y=Z=\mathbb{R}$; H est $x\longrightarrow x^2$, qui est propre; μ est la mesure $x\,dx$; alors $H\mu(\varphi)=\int\varphi(x^2)\,x\,dx=0$, donc $H\mu=0$. Si K est constante, elle est propre sur le support de $H\mu$ (qui est vide). Mais $K\circ H$ est constante sur $\mathbb{R}$, donc n'est pas propre.

** Si on prend pour H une application qui n'est pas μ-mesurable, et pour K une application constante, $K\circ H$ est constante, donc μ-mesurable !

L'intégr_ale de $H\vec{f}$ par rapport à $H\mu$ est égale à l'intégrale de $\vec{f}$ par rapport à μ , la mesure de l'ensemble $H(A)$ **(si A c X)** par rapport à $H\mu$ est égale à la mesure de A par rapport à μ .

La Premiere formule (IV,6;17) suggère d'employer la notation $d\mu\,(H^{-1}(y))$ pour l'image directe $H\,\mu$, puisque $\varphi\left(H^{-1}(y)\right)$ est la fonction $H\varphi$. Alors (IV,6;10) s'écrit très commodément :

$$\int \vec{f}(y)\, d\mu\left(H^{-1}(y)\right) = \int \vec{f}\left(H(x)\right) d\mu(x)\,, \tag{IV,6;17 bis}$$

qu'on obtient par 'le changement de variable $y = H(x)$".

En outre, si $d\mu$ est de la forme $p\,d\lambda$, $d\mu(x) = p(x)\,d\lambda(x)$, on a

$$d\mu\,\left(H^{-1}(y)\right) = p\left(H^{-1}(y)\right) d\lambda\left(H^{-1}(y)\right) \tag{IV,6;17 ter}$$

En effet, pour $\varphi \in \mathscr{C}(Y)$:

$$\int \varphi(y)\, d\mu\left(H^{-1}(y)\right) = \int \varphi\left(H(x)\right) d\mu(x) \tag{IV,6;17 quarto}$$

$$= \int \varphi\left(H(x)\right) p(x)\, d\lambda(x) = \int \varphi\left(H(x)\right) p\left(H^{-1}\left(H(x)\right)\right) d\lambda(x)$$

$$= \int \varphi(y)\left[p\left(H^{-1}(y)\right) d\lambda\left(H^{-1}(y)\right)\right] .$$

Extension du théorème 59 lorsque μ n'est pas ≥ 0

Les méthodes développées au paragraphe précédent permettent de définir $H\vec{\mu}$, suivant le théorème 59, même si $\vec{\mu}$ est vectorielle, pourvu qu'elle admette une base ≥ 0 . On sait alors en effet, d'après ce qui a été dit au théorème 54 bis, reconnaître si une fonction scalaire f , n'appartenant pas à $\mathscr{C}(X)$, est $\vec{\mu}$ -intégrable. On dira alors que $H\vec{\mu}$ existe, si, quelle que soit $\varphi \in \mathscr{C}(X)$ la fonction $H^*\varphi$ est $\vec{\mu}$ -intégrable, et si, lorsque φ

converge uniformément vers 0 en gardant son support dans un compact fixe, $\int (H^*\varphi)\, d\vec{\mu}$ converge vers 0 . On pose alors, par définition :

$$H\vec{\mu}(\varphi) = \int (H^*\varphi)\, d\vec{\mu} = \int \varphi(H(x))\, d\vec{\mu}(x) . \tag{IV,6;18}$$

C'est ce qui arrive, en particulier, si $\vec{\mu}$ est de norme finie, $\vec{E}$ de dimension finie, H continue car on peut appliquer le corollaire du théorème 54 bis, $H^*\varphi$ étant continue et bornée. D'après le théorème 54, on a $\vec{\mu} = \vec{q}\mu_0$, μ_0 mesure ≥ 0 de norme finie, $\|\vec{q}\| = 1$; on a donc :

$$\left\| \int (H^*\varphi)\, \vec{q}\, d\mu_0 \right\| \leq \|\mu_0\| \;\; |||\varphi||| \quad ; \tag{IV,6;18 bis}$$

ce qui montre que $\|H\vec{\mu}\| \leq \|\mu_0\| = \| |\vec{\mu}| \|$. En fait, on peut même montrer qu'on a $\|H\vec{\mu}\| \leq \|\vec{\mu}\|$.

Exemples divers d'images directes de mesures

1°/ Supposons que H soit une application constante de X dans Y, l'image de X étant un point b de Y . Alors on a trivialement,

$$H^*\varphi(x) = \varphi(H(x)) = \varphi(b) , \tag{IV,6;19}$$

et par conséquent le résultat

$$(H\vec{\mu})(\varphi) = \left(\int d\vec{\mu}\right) \varphi(b) \quad \text{ou} \quad H\vec{\mu} = \left(\int d\vec{\mu}\right) \delta_{(b)} . \tag{IV,6;20}$$

L'image de $\vec{\mu}$ est donc une masse ponctuelle au point b , égale a la masse totale de $\vec{\mu}$, $\int d\vec{\mu}$. C'est tout à fait intuitif : H transporte X tout entier sur b , donc elle transporte toute la masse en b .

Ce résultat suppose évidemment qu'on soit dans les conditions précédemment définies, c'est-a-dire, par exemple, que X soit compact, ou que μ soit réelle ≥ 0 de norme finie. On voit même très bien ici les raisons d'être des restrictions relatives à H . Si par exemple H est une application constante de la droite réelle $X = \mathbb{R}$ dans Y , et si $d\mu = dx$, mesure de Lebesgue de $\mathbb{R}$, H n'est pas μ-propre. Si $H\mu$ existait, elle devrait être une masse $+\infty$ au point $b = H(\mathbb{R})$, ce qui est absurde : un point n'a jamais une mesure Infinie, puisqu'il est une partie compacte.

2°/ Supposons que $\vec{\mu}$ soit une combinaison linéaire de mesures de Dirac, à savoir $\vec{\mu} = \sum_\nu c_\nu \delta_{(a_\nu)}$. Dans ce cas, $H\vec{\mu}$, si elle est définie, est donnée par

$$H\vec{\mu} = \sum_\nu \vec{c}_\nu \delta_{H(a_\nu)} \quad \text{. En particulier } H\delta_{(a)} = \delta_{(H(a))} \ . \tag{IV,6;21}$$

<u>Cet exemple montre la signification physique concrète de l'image d'une mesure</u>. Si une mesure $\vec{\mu}$ est composéed'un certain nombre de masses portées par certains points, son image s'obtient en "transportant" ces masses, la masse portée par chaque point étant transportée sur son image par H . Il était évident a priori qu'une telle opération (IV,6;21) existait pour une mesure formde de la somme d'un nombre fini de masses ponctuelles: le résultat de ce paragraphe a été précisément de montrer que le transport d une mesure par une application H était une opération valable dans des cas beaucoup plus gdnéraux * .

3°/ Supposons que X soit un espace affine de dimension finie, et soit $\vec{h}$ un vecteur de l'espace vectoriel associé $\vec{X}$. Alors nous prendrons pour H la translation $\tau_{\vec{h}}$ de vecteur $\vec{h}$, définie par $x \longrightarrow x + \vec{h}$. C'est un homéomorphisme (et on peut donc appliquer les différentes formules (IV,6;17)).

L'image directe d'une fonction φ , définie sur X, est donnée par la formule

$$\tau_{\vec{h}}\,\varphi(x) = \varphi\left(\tau_{-\vec{h}}\,x\right) = \varphi(x - \vec{h}) \;^{**} \ . \tag{IV,6;22}$$

L'image directe d'une mesure $\vec{\mu}$ portée par X , par la translation $\vec{h}$,est la mesure définie par la formule

$$(\tau_{\vec{h}}\,\vec{\mu})(\varphi) = \mu(\tau^*_{\vec{h}}\,\varphi) = \mu\left(\tau_{-\vec{h}}\,\varphi\right) = \int \varphi(x + \vec{h})\, d\vec{\mu}(x) \, . \tag{IV,6;23}$$

* "Identifions" chaque point a de X à la mesure $\delta_{(a)}$. On identifie ainsi X à une partie de $\mathcal{C}'(X)$. De même Y s identifie à une partie de $\mathcal{C}'(Y)$. Alors l'application $H : x \longrightarrow H(x)$ devient l'application $\delta_{(x)} \longrightarrow \delta_{H(x)}$; et $\mu \longrightarrow H\mu$ est un <u>prolongement</u> de H en une application de $\mathcal{C}'(X)$ (ou d'une partie de $\mathcal{C}'(X)$) dans $\mathcal{C}'(Y)$, <u>linéaire.</u>

** Ce renvoi se trouve à la page suivante.

D'après le théorème 60, si μ est réelle $\geqslant 0$, pour qu'une fonction $\vec{f}$, à valeurs dans l'espace de Banach $\vec{F}$, soit $\tau_{\vec{h}}$ μ-intégrable, il faut et il suffit que la fonction $\tau_{-\vec{h}}\vec{f} : x \to \vec{f}(x+\vec{h})$, soit μ-intégrable, et on a

$$\int \vec{f}\, d(\tau_{\vec{h}}\,\mu) = \int \vec{f}(x+\vec{h})\, d\mu(x) \qquad (\text{IV},6;24)$$

Avec la notation (IV,6:17 bis), on écrit

$$\tau_{\vec{h}}\,\mu = d\mu(x-\vec{h})\ , \qquad (\text{IV},6;24\,bis)$$

et alors (IV,6;24) s'écrit

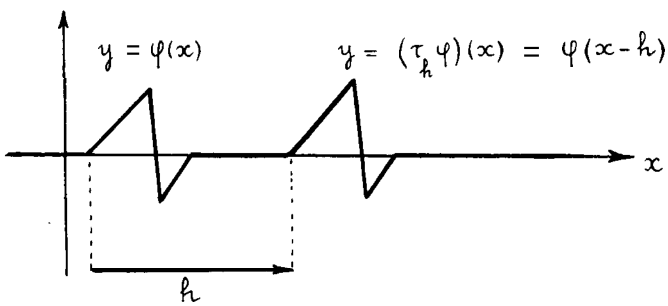
$$\int \vec{f}(x)\, d\mu(x-\vec{h}) = \int \vec{f}(x+\vec{h})\, d\mu(x)\ . \qquad (\text{IV},6;24\,ter)$$

Prenons $\vec{X} = \mathbb{R}$. Si $d\mu$ est de la forme $p(x)\,dx$, (IV,6:24 ter) s'écrit aussi, par le changement de variable $x + h = \xi$, $\int \vec{f}(\xi)\,p(\xi-h)\,d\xi$, de sorte qu on a :

$$d\mu(x-h) = p(x-h)\,dx \ ; \text{ et } \ d(x-h) = dx \ , \text{ pour } p \equiv 1. \qquad (\text{IV},6;24\,quarto)$$

** Renvoi de la page 547.

Il faut s'habituer à ces formules : la translatée de φ par $\vec{h}$ est $x \longrightarrow \varphi(x-\vec{h})$. Supposons que X soit la droite réelle $\mathbb{R}$, et construisons le graphique de $y = \varphi(x)$. Translater la fonction de h , c'est translater son graphique de h parallèlement à l'axe des x ; la valeur de $\tau_h\varphi$ en x est la valeur de φ en $\tau_{-h}\,x = x - h$.



4°/ Prenons maintenant pour H l'application de $\mathbb{R}$ sur $\mathbb{R}$ définie par l'homothétie de centre origine et de rapport k. Si $k = 0$ H est une application constante, cet exemple a déjà été vu à (IV,6;20).

Supposons donc $k \neq 0$, auquel cas H est un homéomorphisme. On a, par définition, la formule

$$(H\varphi)(x) = \varphi\left(\frac{x}{k}\right). \tag{IV,6;25}$$

Ici (IV,6;17 bis) donne

$$\int \overrightarrow{f(x)}\, d\mu\left(\frac{x}{k}\right) = \int \overrightarrow{f(kx)}\, d\mu(x). \tag{IV,6;26}$$

Si en particulier la mesure μ est la mesure que nous avons appelée $p(x)\,dx$, où p est une fonction localement intégrable par rapport à dx, on a, en anticipant sur la formule générale de changement de variable (IV,9;72), la formule

$$\left\{\begin{aligned} H\mu(\varphi) &= \int \varphi(H(x))\, d\mu(x) \\ &= \int \varphi(kx)\, p(x)\, dx \\ &= \int \varphi(y)\, p\left(\frac{y}{k}\right) \frac{dy}{|k|} \end{aligned}\right. \tag{IV,6;27}$$

, d'où

$$d\mu\left(\frac{x}{k}\right) = p\left(\frac{x}{k}\right)\frac{dx}{|k|}\ ; \quad \text{et} \quad d\left(\frac{x}{k}\right) = \frac{dx}{|k|}. \tag{IV,6;28}$$

Convergence en norme, convergence locale en norme

Quand dira-t-on qu'une suite de mesures de Radon $\vec{\mu}_n$ sur X, à valeurs dans un espace vectoriel $\vec{E}$, converge vers une mesure limite $\vec{\mu}$? On pourrait dire qu'il en est ainsi si $\|\vec{\mu}_n - \vec{\mu}\|$ converge vers 0 pour n tendant vers l'infini; c'est la <u>convergence en norme</u>. Ce n'est évidemment pas la une définition adéquate, puisque la norme d'une mesure peut être infinie, et que par conséquent, malgré son nom, ce n'est pas une véritable norme * .
Il sera plus naturel de poser la définition suivante :
On dit qu'une suite de mesures $\vec{\mu}_n$ sur X, à valeurs dans $\vec{E}$ converge vers une mesure limite $\vec{\mu}$ pour n tendant vers l'infini, si, quel que soit le compact K de X, la suite des normes $\|\vec{\mu}_n - \vec{\mu}\|_K$ converge vers 0 pour n tendant vers l'infini.

Conformément à ce qui a été remarqué page **145** à propos de la convergence définie par la norme sur les espaces d'applications linéaires continues d'un espace vectoriel normé dans un autre, cela signifie simplement que, quel que soit le compact **K** de **X**, les fonctions $\vec{\mu}_n$ sur l'espace $\mathscr{C}(X)$ convergent vers la fonction $\vec{\mu}$, uniformément sur la boule unité de $\mathscr{C}_K(X)$. Cette convergence sera la <u>convergence locale en norme</u>; on dira que les $\vec{\mu}_n$ convergent localement en norme vers $\vec{\mu}$.

<u>Exemple</u> - Soit μ une mesure de Radon réelle ≥ 0 fixe, et soit $\vec{f}_n$ une suite de fonctions à valeurs dans l'espace de Banach $\vec{E}$, localement μ-intégrables, convergeant simplement μ-presque partout vers une fonction limite $\vec{f}$ lorsque n tend vers l'infini, et telles que l'on ait une inégalité $\|\vec{f}_n(x)\| \leq g(x)$, où g est une fonction fixe ≥ 0 <u>localement</u> μ-intégrable.

Alors les mesures produit $\vec{f}_n \mu$ convergent localement en norme vers la mesure $\vec{f}\mu$, lorsque n tend vers l'infini.

Soit en effet **K** un compact de X ; on a alors l'inégalité

* C'est la circonstance déjà vue page 138, pour la convergence d'une suite d'applications d'un ensemble **E** dans un espace métrique F . Mais on peut prendre la définition précédente, sur le sous-ensemble formé de mesures de norme finie.

$$(\mathrm{IV},7;1) \qquad \|\vec{\mu}_n - \vec{\mu}\|_K \leq \int_K \|\vec{f}_n - \vec{f}\| \, d\mu \, .$$

Or les $\|f_n - f\|$ convergent simplement k-presque partout vers 0 , et sont majorées par $2g$, fonction ≥ 0 μ-intégrable sur K; d'après le théorème de Lebesgue (théorème 35), le second membre converge bien vers 0 pour n infini.

Malgré ce que cette définition a de naturel, elle est généralement trop stricte.

Considérons, par exemple, une suite de points a_n de X, tendant vers a pour n infini; il serait naturel d'espérer que les mesures de Dirac $\delta_{(a_n)}$ convergent vers la mesure de Dirac $\delta_{(a)}$ pour n infini. Or, avec la définition précédente, il n'en est rien. En effet, si par exemple X est compact, et si les a_n sont distincts de a , on a nécessairement, d'après (IV,2;7), $\|\delta_{(a_n)} - \delta_{(a)}\| = 2$, et par conséquent cette quantité ne tend pas vers 0 pour n tendant vers l'infini. D'où la nouvelle définition suivante :

Convergence vague

On dit qu'une suite de mesures de Radon $\vec{\mu}_n$, sur X , à valeurs dans $\vec{E}$, converge vaguement vers une mesure limite $\vec{\mu}$, pour n tendant vers l'infini, si, quelle que soit la fonction φ de $\mathscr{C}(X)$, la suite des vecteurs $\vec{\mu}_n(\varphi)$ converge dans $\vec{E}$ vers le vecteur $\vec{\mu}(\varphi)$ pour n tendant vers l'infini. Si $\vec{E}$ est le corps des scalaires, il s'agira de convergence des nombres $\mu_n(\varphi)$ vers le nombre $\mu(\varphi)$. Si les μ_n sont réelles (resp. ≥ 0), il en est de même pour μ . Lorsqu'on considère les mesures comme des fonctions sur $\mathscr{C}(X)$, la convergence vague n'est pas autre chose que la convergence simple de ces fonctions, et souvent on l'appelle convergence simple. Naturellement la convergence locale en norme entraîne la convergence vague *

* Il y a là une situation assez curieuse. Nous avons vu au § 15 du chapitre II que la convergence simple était généralement trop faible pour être utile, et nous avons introduit des convergences uniformes. Ici, nous signalons que les convergences uniformes sont trop fortes, et nous revenons à la convergence simple !

Exemples 1°/ Si les a_n sont une suite de points de X convergeant vers a pour n infini, les mesures de Dirac $\delta_{(a_n)}$ convergent vaguement vers la mesure de Dirac $\delta_{(a)}$.

En effet, quelle que soit φ de $\mathscr{C}(X)$, les $\varphi(a_n)$ convergent vers $\varphi(a)$, car chacune des fonctions φ est continue.

2°/ Supposons que X soit un espace affine de dimension finie. Alors, pour tout vecteur $\vec{h}$ de $\vec{X}$, on peut définir la translatée d'une mesure de Radon, $\tau_{\vec{h}}\vec{\mu}$, (formule (IV,6;23)). Si alors $\vec{h}_n$ est une suite de vecteurs de $\vec{X}$ convergeant vers $\vec{h}$, les translatées $\tau_{\vec{h}_n}\vec{\mu}$ convergent vaguement vers la translatée $\tau_{\vec{h}}\vec{\mu}$. En effet, d'après la définition, on a :

$$(\text{IV},7;2) \qquad \left(\tau_{\vec{h}_n}\vec{\mu} - \tau_{\vec{h}}\vec{\mu}\right)(\varphi) = \int\left(\varphi(x+\vec{h}_n) - \varphi(x+\vec{h})\right) d\vec{\mu}(x) .$$

Notre affirmation sera alors démontrée, si nous prouvons que les fonctions $x \longrightarrow \varphi(x+\vec{h}_n)$ convergent uniformément vers la fonction $x \longrightarrow \varphi(x+\vec{h})$. En effet, ces fonctions gardent en outre leur support dans un compact fixe K (soit K_0 le support de φ · il est compact donc, si nous mettons sur X une norme, contenu dans une boule fermée d'un certain centre a_0 et de rayon R_0 ; si alors+ est la borne supérieure de tous les $\|\vec{h}_n\|$, le support des $\tau_{-\vec{h}_n}\varphi$ est contenu dans la boule compacte de centre a_0 et de rayon $R_0 + \rho$), et par conséquent elles convergeront dans $\mathscr{C}_K(X)$; comme $\vec{\mu}$ est une fonction continue sur cet espace $\mathscr{C}_K(X)$, notre affirmation en résultera.

Montrons donc cette convergence uniforme. Comme la fonction φ est continue et à support compact, elle est uniformément continue. En effet, soit, comme tout-à-l'heure, 13 $(a_0\ ;\ R_1)$ une boule contenant le support K_0 de φ ; sur le compact $B(a_0\ ;\ R_0+1)$, la fonction φ est uniformément continue (Théorème 31 du chapitre II); donc, étant donné $\varepsilon > 0$, il existe η, $0<\eta\leqslant 1$, tel que $\|x'-x''\|\leqslant\eta$ entraîne $|\varphi(x')-\varphi(x'')|\leqslant\varepsilon$, pour x' et x'' dans $B(a_0\ ;R_0+1)$; et, si x' et x'' ne sont pas tous deux contenus dans la boule précédente, et vérifient $\|x'-x''\|\leqslant\eta$, c'est qu'ils sont tous les deux à l'extérieur de la boule $B(a_0; R_0)$, donc en dehors de K_0, alors on a $\varphi(x') = \varphi(x'') = 0$, de sorte que, finalement, on a, dans tous les cas, $|\varphi(x') - \varphi(x'')| \leq \varepsilon$; ce qui montre bien que φ est uniformément continue sur **X** .

Si alors on a $\| \vec{h}_n - \vec{h} \| \leq \eta$, on a quel que soit x de **E** , l'inégalité $|\varphi(x+\vec{h}_n) - \varphi(x+\vec{h})| \leq \varepsilon$, ce qui prouve bien que $\tau_{-\vec{h}_n}\varphi$ converge uniformément vers $\tau_{-\vec{h}}\varphi$, lorsque $\vec{h}_n$ tend vers $\vec{h}$, et nous avons donc bien montré que $\tau_{\vec{h}_n}\vec{\mu}$ converge vaguement vers $\tau_{\vec{h}}\mu$ quand $\vec{h}_n$ tend vers $\vec{h}$.

3°/ Si, dans l'exemple précédent, μ est une mesure 3 0 (ou plus généralement de base ≥ 0), la démonstration est encore plus simple. En effet, la convergence uniforme de $\varphi(x+\vec{h}_n)$ vers $\varphi(x+\vec{h})$ n'est pas nécessaire : il y a convergence simple, et les $\varphi(x+\vec{h}_n)$ sont bornées en module par $\|| Q \||$ et gardent leur support dans un compact fixe, donc le thdorème **34** de Lebesgue donne le résultat. On peut généraliser comme suit :

<u>**soit ti**, une suite d'applications continues de X dans Y , convergeant simplement pour n infini vers une application continue H . Soit μ une mesure ≥ 0 sur X , de norme finie * . Alors les mesures images $H_n\mu$ convergent vaguement, pour n infini, vers la mesure image $H\mu$</u> .

En ef'fet, on a (IV,6;4), pour les $H_n\mu$ et $H\mu$, et $\varphi \in \mathscr{C}(Y)$. Pour n infini, $\varphi(H_n(x))$ converge simplement vers $\varphi(H(x))$, en restant bornéepar la constante $\||\varphi\||$, qui est intégrable puisque μ est de norme finie; donc le théorème **35** de Lebesgue donne le résultat.

* Nous supposons, pour simplifier, μ de norme finie. On pourrait faire d'autres hypothèses, du type "applica - tions propres'.

4°/ Si $X = [a,b]$ est un intervalle borné de $\mathbb{R}$, la mesure

(IV,7;3) $$\mu_n = \frac{1}{n+1}\left(\delta_{(a)} + \delta_{\left(a+\frac{b-a}{n}\right)} + \cdots + \delta_{\left(a+k\frac{b-a}{n}\right)} + \cdots + \delta_{(b)}\right)$$

converge vaguement, pour n infini, vers la mesure $\frac{1}{b-a}\,dx$. Si en effet φ appartient à $\mathscr{C}([a,b])$, on a la formule (IV,1;40) (formule de la moyenne), qui prouve notre affirmation.

Les fonctions μ -intégrables-Riemann

Posons-nous maintenant le problème suivant :

soit $\vec{\mu}_n$ une suite de mesures de Radon, convergeant vaguement vers une mesure limite $\vec{\mu}$ pour n tendant vers l'infini. Peut-on alors affirmer que $\int \varphi\, d\vec{\mu}_n$ converge vers $\int \varphi\, d\vec{\mu}$, pour n tendant vers l'infini, pour d'autres fonctions φ que celles qui appartiennent à $\mathscr{C}(X)$? Il est bien évident que ce n'est pas vrai pour n'importe quelle fonction φ. Considérons par exemple encore une fois une suite de points a_n de X convergeant vers a pour n infini. Nous avons vu que les $\delta_{(a_n)}$ convergent vaguement vers $\delta_{(a)}$. Si alors φ est une fonction discontinue au point a, il est bien évident que les $\varphi(a_n)$ ne convergent pas nécessairement vers $\varphi(a)$, et ceci donne un contre-exemple à la suggestion précédente.

Il nous faudra donc introduire des fonctions φ particulières. Par ailleurs, nous nous bornerons au cas des mesures réelles $\geqslant 0$.

Définition - Soit μ une mesure de Radon réelle $\geqslant 0$ sur X et soit f une fonction réelle $\geqslant 0$ sur X bornée à support compact. On appelle intégrale supérieure de Riemann de f par rapport à μ, et on note $\int^{*R} f\, d\mu$, la borne inférieure des $\mu(\varphi)$ pour toutes les fonctions φ de $\mathscr{C}(X)$ qui majorent f.

On peut s'étonner de cette définition, qui ne correspond pas à celle que nous avons donnée de l'intégrale supérieure de Riemann relativement à dx au § 1. Mais on montre facilement que, lorsque μ est la mesure dx sur $\mathbb{R}$, les deux notions coïncident. Désignons par $\int^{*R_1}$ et $\int^{*R_2}$ l'intégrale supérieure de Riemann au sens du paragraphe 1 et au sens actuel, pour $\mu = dx$. Soit φ une fonction de $\mathscr{C}(X)$

majorant f ; on a immédiatement $\int^{*R_1} f\,dx \leq \int^{*R_1} \varphi\,dx = \mu(\varphi)$, parce qu'une fonction de $\mathcal{C}(X)$ est intégrable-Riemann au sens (1), d'où, en prenant la borne inférieure pour toutes les φ, l'inégalité $\int^{*R_1} f\,dx \leq \int^{*R_2} f\,dx$. Mais d'autre part on peut trouver une fonction en escalier à support compact $g \geq 0$ telle que $\int^{*R_1} f\,dx \geq \int g\,dx - \frac{\varepsilon}{2}$.
La fonction g est de la forme $\sum_{i=1}^{N} g\,\chi_{|a_i, b_i|}$, où les g_i sont des constantes et les $|a_i, b_i|$ sont N intervalles de $\mathbb{R}$. Posons $M = \operatorname{Max}_{i=1}^{N} g_i$. Posons $a' = a - \frac{\&}{4MN}$, $b' = b + \frac{\varepsilon}{4MN}$ et appelons φ_i la fonction égale a 1 sur $[a_i, b_i]$, à 0 sur $]-\infty, a'_i]$ et sur $[b'_i, +\infty[$ et affine dans chacun des intervalles $[a'_i, a_i]$, $[b_i, b'_i]$. Si alors nous appelons φ la fonction $\sum_{i=1}^{N} g_i \varphi_i$, on a d'une part $f \leq \varphi$ et d'autre part l'inégalité

$$\int g\,dx \geq \int \varphi\,dx - \frac{\varepsilon}{4MN} M\,2N = \int \varphi\,dx - \frac{\varepsilon}{2},$$

qui donne l'inégalité $\int^{*R_1} f\,dx \geq \int \varphi\,dx - \varepsilon \geq \int^{*R_2} f\,dx - \varepsilon$;

ε étant arbitraire, cela montre bien que les deux notions d'intégrale supérieure de Riemann coïncident. D'après ce qui a déjà été vu pour l'exemple de dx (page 482) l'intégrale supérieure de Riemann, qui est toujours au moins égale à l'intégrale supérieure de Lebesgue, est en général strictement plus grande. Si μ est une mesure de Radon sur la droite réelle $\mathbb{R}$, différente de dx, il y aurait deux définitions possibles de l'intégrale supérieure de Riemann l'une par les fonctions de $\mathcal{C}(X)$, et l'autre par les fonctions en escalier à support compact. Dans le § 1, nous n'avions pas le choix, car nous connaissions de façon triviale les intégraies des fonctions en escalier, mais pas encore l'intégrale d'une fonction continue à support compact; c'est seulement après la définition de l'intégrale de Riemann du § 1 que dx devenait une mesure de Radon sur $\mathbb{R}$. Au contraire, Pour une mesure de Radon quelconque μ sur $\mathbb{R}$ nous avons le choix entre les deux possibilités. Ces deux' définitions ne coïncident pas nécessairement. Si par exemple nous considérons la mesure μ sur $\mathbb{R}$, somme de la mesure dx et de la mesure δ, et si nous prenons pour f la fonction caractéristique de l'intervalle $]0,1]$, c'est une fonction en escalier, de sorte que l'on a l'égalité $\int^{*R_1} f\,d\mu = \int f\,d\mu = 1$; si par contre φ est n'importe quelle fonction de $\mathscr{C}(X)$ qui la majore, on a $\varphi(0) \geqslant 1$, et $\int \varphi\,d\mu \geqslant 2$, donc $\int^{*R_2} f\,d\mu \geqslant 2$ et en fait $= 2$.

Il est cependant très clair que la définition intéressante dans le cas général est celle qui utilise les fonctions de $\mathscr{C}(X)$: ne serait-ce parce que c'est la seule qui soit définie sur n'importe quel espace localement compact X et pas seulement sur la droite $\mathbb{R}$; mais ce sont surtout les propriétés ultérieures qui en fait justifient la définition donnée] .

Si $\vec{f}$ est une fonction à valeurs dans l'espace de Banach $\vec{F}$, bornée à support compact, on dira qu'elle est μ-intégrable-Riemann si elle possède une suite d'approximation, relativement a l'intégrale supérieure de Riemann, formée de fonctions continues décomposables a support compact : $\vec{f}_n \in \mathscr{C}(X) \otimes \vec{F}$ (voir avant le corollaire 8 du théorème 11), et $\int^{*R} \| \vec{f} - \vec{f}_n \| d\mu$ converge vers 0 pour n infini. On peut alors définir l'intégrale de Riemann de $\vec{f}$; mais les f_n forment à fortiori une suite d'approximation de $\vec{f}$ pour l'intégrale de Lebesgue, $\vec{f}$ est intégrable-Lebesgue, et son intégrale de Riemann coïncide avec son intégrale de Lebesgue.

Toute fonction continue à support compact est intégrable-Riemann, puisque (corollaire 8 du théorème 11) elle est limite uniforme de fonctions décomposables, à support dans un compact fixe. Mais, sur la droite réelle $\mathbb{R}$, une fonction réglée n'est pas forcément intégrable-Riemann, contrairement à ce qui a été dit pour $\mu = dx$ au corollaire 2 du théorème 8 : pour $\mu = dx + \delta$, la fonction caractéristique de $]0,1]$ n'est pas intégrable-Riemann, sans quoi son intégrale supérieure de Riemann, qui est 2, coïnciderait avec son intégrale de Riemann donc de Lebesgue, qui est 1.

Les théorèmes 1 à 6 sont encore valables, ainsi que les remarques qui les accompagnent, et les corollaires (sauf en ce qui concerne l'intégrale des fonctions caractéristiques d'intervalles) a la condition de remplacer les fonctions en escalier par les fonctions continues décomposables. On remarquera que, de même qu'au paragraphe 1, on n'intègre au sens de Riemann que les fonctions bornées à support compact.

<u>Théorème 62 - Si les μ_n sont une suite de mesures de Radon réelles $\geqslant 0$ sur X qui convergent vaguement vers une limite μ pour n tendant vers l'infini, si $\vec{f}$ est une fonction sur X à valeurs dans $\vec{F}$, bornée à support compact, mesurable pour toutes les mesures μ_n et intégrable-Riemann pour μ, alors les $\int \vec{f}\, d\mu_n$ convergent vers $\int \vec{f}\, d\mu$.</u>

<u>Démonstration</u> - Tout d'abord $\vec{f}$ est intégrable-Lebesgue pour toutes les μ_n et μ. D'après la remarque 2° page 408 (où l'on remplace fonctions en escalier par fonctions continues décomposables) on peut trouver deux fonctions g, continue à support compact, décomposable, à valeurs dans $\vec{F}$, et h continue réelle $\geqslant 0$ à support compact, telles que $\|\vec{f}-\vec{g}\| \leqslant h$ et μ (A) $\leqslant \frac{\varepsilon}{8}$. Comme alors les μ_n convergent vaguement vers μ, on peut déterminer un entier p tel que pour $n \geqslant p$ on ait $|\mu_n(h) - \mu(h)| \leqslant \frac{\varepsilon}{8}$, donc $\mu_n(h) \leqslant \frac{\varepsilon}{4}$.

D'autre part l'égalité $\int \vec{g}\, d\nu = \sum_{i=1}^{N} \int \vec{g}_i \varphi_i\, d\nu = \sum_{i=1}^{N} \vec{g}_i\, \nu(\varphi)$ montre que la convergence des μ_n vers μ entraine, pour toute fonction $\vec{g}$ continue décomposable à support compact la convergence des $\int \vec{g}\, d\mu_n$ vers $\int \vec{g}\, d\mu$, de sorte qu'il existe un entier q, tel que pour $n \geqslant q$ on ait

$$\left\| \int \vec{g}\, d\mu_n - \int \vec{g}\, d\mu \right\| \leqslant \frac{\varepsilon}{2}.$$

On a alors immédiatement l'inégalité suivante, pour $n \geqslant$ Max (p,q) :

$$\left\| \int \vec{f} d\mu_n - \int \vec{f} d\mu \right\| \leqslant \int \| \vec{f} - \vec{g} \| d\mu_n + \left\| \int \vec{g}\, d\mu_n - \int \vec{g}\, d\mu \right\| + \int \| g - f \| d\mu$$
$$\leqslant \mu_n(h) + \frac{\varepsilon}{2} + \mu(h) \leqslant \varepsilon,$$

qui prouve le théorème.

<u>Remarque</u> - Ce théorème est très remarquable : on suppose que $\vec{f}$ est μ-intégrable-Riemann, μ étant la mesure limite; on ne fait pas d'hypothèse analogue sur l'intégrabilité par rapport aux μ_n ; (bien entendu on suppose cependant que $\vec{f}$ est μ_n-mesurable donc μ_n-intégrable-Lebesgue).

Théorème 63 - pour qu'une fonction $\vec{f}$ définie sur X à valeurs dans $\vec{F}$, bornée à support compact, soit μ-intégrable-Riemann, il faut et il suffit que l'ensemble de ses points de discontinuité ait une μ-mesure de Lebesgue nulle.

Démonstration - 1°/ Supposons que $\vec{f}$ soit μ-intégrable-Riemann, et soient $\vec{g}$ et h des fonctions déterminées comme au théorème 62 : $\|\vec{f}-\vec{g}\| \leq h$, $\mu(h) \leq \varepsilon$. Pour tout point a de X et tout voisinage $\mathcal{V}$ de a, appelons oscillation de $\vec{f}$ dans le voisinage $\mathcal{V}$ la quantité $\omega(\vec{f};\mathcal{V}) = \operatorname{Sup}_{x,y\in\mathcal{V}} \|\vec{f}(x)-\vec{f}(y)\|$, et oscillation de $\vec{f}$ au point a la quantité $\omega(\vec{f};a) = \operatorname{Inf}_{\mathcal{V}} \omega(\vec{f};\mathcal{V})$ *. La fonction $\vec{f}$ est continue au point a, si et seulement si cette oscillation est nulle. On a alors immédiatement l'inégalité

$$\omega(\vec{f};a) \leq \omega(\vec{g};a) + \omega(\vec{f}-\vec{g};a) = \omega(\vec{f}-\vec{g};a)$$

car $\vec{g}$ est continue; par ailleurs, dans tout voisinage $\mathcal{V}$ de a, on a la majoration $\|(\vec{f}(x)-\vec{g}(x))-(\vec{f}(y)-\vec{g}(y))\| \leq h(x)+h(y)$;

comme alors h est continue, on en déduit $\omega(\vec{f}-\vec{g};a) \leq 2h(a)$.

si alors on appelle A_α l'ensemble $\{x\in X\ ;\ \omega(\vec{f};x) \geq \alpha\}$ $\alpha>0$, il est contenu dans l'ensemble $\{x \in X\ ;\ h(x) \geq \frac{\alpha}{2}\}$ et par conséquent sa μ-mesure extérieure est majorée par $\frac{2\varepsilon}{\alpha}$ puisque $\mu(h) \leq \varepsilon$.

Ceci étant vrai quel que soit $\varepsilon > 0$, sa mesure est nulle; ceci étant vrai quel que soit α, l'ensemble des points où l'oscillation de $\vec{f}$ est > 0, c'est-a-dire des points de discontinuité de $\vec{f}$, est de μ-mesure nulle.

2°/ Inversement, supposons que l'ensemble des points de discontinuité de $\vec{f}$ soit de μ-mesure nulle. Montrons d'abord que ceci entraîne que $\vec{f}$ soit μ-mesurable (et, par conséquent, puisqu'elle est bornée et à support compact, μ-intégrable Lebesgue) : elle vérifie trivialement le critère de Lusin (corollaire du théorème 33) : Si K est un compact quelconque, et si A est l'ensemble des points de discontinuité de $\vec{f}$, supposé de μ-mesure nulle, on peut trouver un ensemble ouvert $\mathcal{O}$ contenant A, de μ-mesure $\leq \delta$, et

* Cela coïncide avec la définition particulière donnée pour les fonctions réglées Page 184.

alors, si $K_\delta = K \cap \complement \mathcal{O}$, on a bien $\mu(K - K_\delta) \leq \delta$; en tout point de K , $\vec{f}$ est continue sur X , et par conséquent à fortiori la restriction de $\vec{f}$ à K est continue.

puisque $\vec{f}$ est μ - intégrable-Lebesgue, il existe, d'après le théorème 32, une fonction $\vec{g}$ continue décomposable à support compact, telle que, si $\vec{f} - \vec{g} = h$, on ait $\int h \, d\mu \leq \frac{\varepsilon}{2}$. Appelons h la "fonction supérieure" associée à h, c'est-à-dire définie de la façon suivante : pour tout voisinage $\mathcal{V}$ de a dans X , on posera $h^*(\mathcal{V}) = \operatorname{Sup}_{x \in \mathcal{V}} h(x)$, et alors on posera $h^*(a) = \operatorname{Inf}_{\mathcal{V}} h^*(\mathcal{V})$. On voit sans peine que h^* est semi-continue supérieurement : en effet, tout a possède un voisinage ouvert $\mathcal{V}$ tel que l'on ait : $h^*(\mathcal{V}) \leq h^*(a) + \varepsilon$ et alors, pour tout x de $\mathcal{V}$, on a $h^*(x) \leq h^*(\mathcal{V}) \leq h^*(a) + \varepsilon$. Par ailleurs h^* coïncide avec h en tout point où h est continue donc μ-presque partout. On a donc $\int h \, d\mu = \int h^* d\mu$.

Soit k une fonction continue quelconque majorant h^* , et intégrable (par exemple $k \in \mathcal{C}(X)$) . Alors $k - h^*$ est semi-continue inférieurement ≥ 0 , donc borélienne, et intégrable puisque majorée par k ; d'après le théorème 39 bis, $\int (k - h^*) d\mu$ est la borne supérieure des $\mu(\Psi), \Psi \in \mathcal{C}(X)$, $0 \leq \Psi \leq k - h^*$, donc $\int h^* d\mu$ est la borne inférieure des $\mu(k - \Psi) = \mu(\varphi)$, pour toutes les φ continues ≥ 0 intégrables qui majorent h^* (en l'espèce des $\varphi \in \mathcal{C}(X)$ si on a pris $k \in \mathcal{C}(X)$). On peut donc trouver $\varphi \in \mathcal{C}(X)$ $\varphi \geq h^* \geq h$, telle que $\mu(\varphi) \leq \int h^* d\mu + \frac{\varepsilon}{2} = \int h \, d\mu + \frac{\varepsilon}{2} \leq \varepsilon$. Mais alors on a $\| \vec{f} - \vec{g} \| \leq \varphi$, $\mu(\varphi) \leq \varepsilon$; ceci étant possible pour tout $\varepsilon > 0$, $\vec{f}$ est μ- intégrable-Riemann.

<u>Corollaire 1</u> - <u>Pour qu'une partie **A** de X, d'adhérence compacte, ait une fonction caractéristique μ-intégrable-Riemann il faut et il suffit que sa frontière soit de μ-mesure (de Lebesgue) nulle.</u>

<u>Démonstration</u> - L'ensemble des points de discontinuité de la fonction caractéristique de **A** n'est autre que l'ensemble des Points frontière de A , car au voisinage de tout point de l'intérieur de A cette fonction est constante et égale à 1, et au voisinage de tout point de l'extérieur elle est constante et égale à 0 .

<u>Corollaire 2</u> - Si les μ_n sont des mesures ≥ 0 sur X , convergeant vaguement vers μ Pour n tendant vers l'infini, et

si **A** est une partie de X, d'adhérence compacte, mesurable pour toutes les μ_n et dont la frontière $\dot{A}$ soit de mesure nulle pour la mesure limite μ, alors $\mu_n(A)$ converge vers $\mu(A)$.

Corollaire **3** - Si les μ_n sont des mesures de Radon sur $\mathbb{R}$, convergeant vaguement, pour n tendant vers l'infini, vers la mesure dx, et si φ est une fonction bornée, à support compact, mesurable pour toutes les μ_n, et intégrable-Riemann dx ; alors les $\int \varphi\, d\mu_n$ convergent vers $\int \varphi\, dx$.

Il existe, sur la droite $\mathbb{R}$, une réciproque des théorèmes précédents qui n'est pas sans intérêt :

Théorème **64** - Si les μ_n sont des mesures de Radon ≥ 0 sur la droite $\mathbb{R}$, et si, pour tout intervalle borné $[a,b]$ tel que $\mu(\{a\}) = \mu(\{b\}) = 0$, les $\mu_n([a,b])$ convergent, pour n infini, vers $\mu([a,b])$, alors les mesures μ_n convergent vaguement vers la mesure μ.

On pourra démontrer ce théorème à titre d'exercice.

Convergence vague et convergence uniforme

Nous allons maintenant chercher un autre genre de problèmes. Supposons que des mesures $\vec{\mu}_n$ convergent vaguement vers une limite μ ; peut-on trouver certains ensembles $\mathcal{A}$ de fonctions φ, tels que $\vec{\mu}_n(\varphi)$ converge vers $\vec{\mu}(\varphi)$, uniformément pour φ dans $\mathcal{A}$.

D'après ce que nous avons déjà vu, une telle convergence uniforme ne peut pas avoir lieu, si $\mathcal{A}$ est la boule unité de $\mathcal{C}_K(X)$. Il faudra donc des ensembles plus petits que la boule unité.

Le théorème **général** qui est alors utilisé est un théorème d'analyse fonctionnelle :

Théorème 65 (Banach-Steinhaus)

Soient u_n des applications linéaires continues d'un espace de Banach $\vec{E}$ dans un espace vectoriel normé $\vec{F}$. Si les fonctions u_n, définies **sur** $\vec{E}$ **à valeurs dans** $\vec{F}$, **convergent** simplement pour n infini vers la fonction **limite** u, alors les $\| u_n \|$ sont **bornées** par un nombre fixe M, u est linéaire et continue, et les u_n convergent vers u **uniformément** sur toute partie $\mathcal{A}$ compacte de $\vec{E}$.

Démonstration Il est essentiel **que** $\vec{E}$ soit un Banach, **c'est**-à-dire soit complet, faute de quoi tous les résultats énoncés deviennent inexacts. C'est un des rares théorèmes où l'on doit supposer l'espace objet complet. Le théorème de Banach-Steinhaus est un des théorèmes les plus profonds et les plus féconds de l'analyse moderne.

Démontrons d'abord que les normes des u_n sont majorées par un même nombre $M \geqslant 0$. Pour cela nous aurons d'abord besoin d'un lemme.

Lemme: théorème de Baire *

Soit E un espace métrique complet, et soit $\mathcal{O}_0, \mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_n, \ldots$, une suite d'ouverts denses de E. Alors l'intersection

$$\underline{A} = \bigcap_{n \geqslant 0} \mathcal{O}_n$$

(qui n'a plus aucune raison d'être ouverte).

est encore dense dans E.

Démonstration du lemme. Soit Ω un ouvert quelconque de $\mathbf{E}$; nous voulons démontrer que $\Omega \cap A$ n'est pas vide. Comme $\mathcal{O}_0$ est dense, il existe au moins un point de $\Omega \cap \mathcal{O}_0$, et comme c'est un ensemble ouvert il existe **même** une boule fermée $B_0 = B(a_0, \delta_0)$ entièrement contenue dans $\Omega \cap \mathcal{O}_0$. Répétons maintenant l'opération en remplaçant $\mathcal{O}_0$ par $\mathcal{O}_1$ et Ω par la boule ouverte $\overset{\circ}{B}_0$. Il existe alors une boule fermée $B_1 = B(a_1, \delta_1)$ entièrement contenue dans $\overset{\circ}{B}_0 \cap \mathcal{O}_1$; nous choisirons $\delta_1 \leqslant 1$. Et ainsi de suite; de proche en proche nous pouvons trouver une suite de points a_n et une suite de nombre $\delta_n \leqslant \frac{1}{n}$ telles que chaque boule fermée $B_n = B(a_n, \delta_n)$ soit contenue dans la boule ouverte $\overset{\circ}{B}_{n-1} = \overset{\circ}{B}_{n-1}(a_{n-1}, \delta_{n-1})$ et dans $\mathcal{O}_n$; toutes ces boules fermées sont donc elles-mêmes contenues dans Ω.

* Nous le baptisons lemme, mais c est un des plus importants théorèmes de l'analyse, dépassant de beaucoup l'usage qui en est fait ici !

La suite des centres a_n des boules est une suite de Cauchy. En effet tous les points a_{n+p} pour $p \geqslant 0$ sont contenus dans la boule fermée B_n; on a donc $d(a_{n+p}, a_n) \leqslant \frac{1}{n}$. Comme E est supposé complet, la suite des a_n converge vers un point limite a. Comme tous les a_{n+p} pour $p \geqslant 0$ sont contenus dans la boule B_n, a est nécessairement contenu dans B_n donc dans $\Omega \cap \mathcal{O}_n$; ceci étant vrai pour tout n , $a \in \Omega \cap A$, ce qui démontre le lemme.

Il est commode aussi d'utiliser ce lemme en utilisant le passage aux parties complémentaires; le complémentaire d'un ouvert est fermé, et le complémentaire d'un ensemble dense c'est-a-dire d'extérieur vide, est un ensemble d'intérieur vide; on alors :

<u>Soit E un espace métrique complet, et F_n une suite d'ensembles fermés d'intérieur vide; alors la réunion $F = \bigcup_{n \geq 0} F_n$ (qui n'est pas nécessairement fermé) a encore un intérieur vide (en particulier $F \neq E$).</u>

Il est bien évident qu'un espace métrique non complet n'a plus la même propriété; par exemple le corps $\mathbb{Q}$ des rationnels avec la métrique habituelle est réunion dénombrable d'ensembles réduits a des points, et un point est, dans $\mathbb{Q}$, fermé d'intérieur vide.

Démontrons maintenant la première partie du théorème, à savoir que les $\| u_n \|$ sont bornées. Appelons F_k l'ensemble des points x de E où tous les $\| \overrightarrow{u_n(x)} \|$ sont majorés par k. L'ensemble $F_{n,k} = \{ x \in E ; \| \overrightarrow{u_n(x)} \| \leqslant k \}$ est fermé puisque u_n est continue, donc $F_k = \bigcap_n F_{n,k}$ est aussi fermé. La réunion des F_k, pour $k = 0, 1, \ldots$ est l'espace entier E ; en effet, pour tout x de E, les $\overrightarrow{u_n(x)}$ sont supposés avoir une limite, donc leurs normes sont bornées, donc il existe un k (dépendant de x) tel que tous les $\| \overrightarrow{u_n(x)} \|$ soient $\leqslant k$, ou $x \in F_k$. comme E est métrique complet (Banach), le lemme de Baire nous dit qu'il est impossible que tous les F_k aient un intérieur vide. Donc il existe un k tel que F_k ait un intérieur non vide, donc une boule $B(a, \rho)$ contenue dans F_k. Alors, pour $\|x\| \leqslant \rho$, on a, pour tout n, $\| \overrightarrow{u_n(x)} \| \leqslant \| \overrightarrow{u_n(x+a)} \| + \| \overrightarrow{u_n(a)} \| \leqslant 2k$.

donc $\| u_n \| \leq \frac{2k}{\rho}$, ce qui démontre notre affirmation.
[Remarquons bien que ce résultat ne suppose nullement que les u_n convergent simplement, mais seulement que, pour tout x de E, les $\| \overrightarrow{u_n(x)} \|$ soient bornées; la borne dépend de x de manière a priori arbitraire, mais la démonstration précédente nous apprend que les $\| u_n \|$ sont bornées autrement dit qu'il existe M tel que la uorne des $\| \overrightarrow{u_n(x)} \|$ soit $\leq M \| x \|$]

Il est évident que u est linéaire; car, de $u_n(\vec{X} + \vec{Y}) = u_n(\vec{X}) + u_n(\vec{Y})$, on déduit Immédiatement, par passage à la limite, que $u(\vec{X} + \vec{Y}) = u(\vec{X}) + u(Y)$; de même $u(k\vec{X}) = k u(\vec{X})$. Il est aussi évident que u est continue et de norme $\leq M$; car, de $\| u_n(\vec{X}) \| \leq M \| \vec{X} \|$, on déduit, par passage à la limite, $\| u(\vec{X}) \| \leq M \| \vec{X} \|$.

Soit maintenant $\mathcal{C}$ un compact de $\vec{E}$.

Si $\varepsilon > 0$ est donné, il existe un nombre fini de boules ouvertes de rayon $\frac{\varepsilon}{3M}$, de centres sur $\mathcal{C}$, qui suffisent à recouvrir $\mathcal{C}$, d'après la définition même de la compacité. Soient $(\vec{a}_i)_{i \in I}$ les centres de ces boules. D'après l'hypothèse de convergence simple, pour chacun de ces centres $\vec{a}_i$, on peut trouver un entier p_i tel que $n \geq p_i$ entraîne

$$\| u_n(\vec{a}_i) - u(\vec{a}_i) \| \leq \frac{\varepsilon}{3}. \tag{IV,7;16}$$

Si alors on appelle? le plus grand des entiers p_i, on voit que, pour $n \geq p$ et $\vec{X}$ situé à une distance $< \frac{\varepsilon}{3M}$ de $\vec{a}_i$, on a nécessairement

(IV,7;17)
$$\|u_n(\vec{X}) - u(\vec{X})\| \leq \|u_n(\vec{X}) - u_n(\vec{a}_i)\|$$

$$+ \|u_n(\vec{a}_i) - u(\vec{a}_i)\| + \|u(\vec{a}_i) - u(\vec{X})\|$$

$$\leq M\frac{\varepsilon}{3M} + \frac{\varepsilon}{3} + M\frac{\varepsilon}{3M} = \varepsilon .$$

Mais, comme tout point de $\mathfrak{A}$ est à une distance $< \frac{\varepsilon}{3M}$ de l'un des $\vec{a}_i$, finalement on a, pour tout $\vec{X}$ de $\mathfrak{A}$ et $n \geq p$, l'inégalité $\|u_n(\vec{X}) - u(\vec{X})\| \leq \varepsilon$, donc les u_n convergent vers u uniformément sur $\mathfrak{A}$, et ceci démontre le théorème.

<u>Corollaire 1</u> - <u>Soient $\vec{\mu}_n$ des mesures sur X , à valeurs dans un espace vectoriel normé $\vec{E}$, convergeant vaguement pour n infini vers une mesure $\vec{\mu}$. Alors, pour tout compact K, les $\|\vec{\mu}_n\|_K$ sont bornées dans leur ensemble. Si $\mathfrak{A}$ est un compact de l'espace $\mathcal{C}_K(X)$, les $\vec{\mu}_n(\varphi)$ convergent vers $\vec{\mu}(\varphi)$, uniformément pour φ dans $\mathfrak{A}$</u> * .

Ce corollaire résulte immédiatement du théorème, parce que $\mathcal{C}_K(X)$ est un espace de Banach (voir page 432).

<u>Corollaire 2</u> - <u>Soient $\vec{\mu}_n$ des mesures convergeant vaguement vers une mesure $\vec{\mu}$, et φ_n des fonctions de $\mathcal{C}(X)$, convergeant uniformément vers une fonction φ, et gardant leurs supports dans un même compact K . Alors les $\vec{\mu}_n(\varphi_n)$ convergent vers $\vec{\mu}(\varphi)$</u> .

<u>Démonstration</u> D'après le corollaire 1, les $\|\vec{\mu}_n\|_K$ sont bornées par un même nombre $M \geq 0$. D'autre part, les φ_n convergent vers φ dans $\mathcal{C}_K(X)$. Alors :

(IV,7;17 bis)
$$\vec{\mu}_n(\varphi_n) - \vec{\mu}(\varphi) = \vec{\mu}_n(\varphi_n - \varphi) + \left(\vec{\mu}_n(\varphi) - \vec{\mu}(\varphi)\right) .$$

* Mais, $\mathcal{C}_K(X)$ étant un espace de dimension en général infinie, sa boule unité n'est pas compacte, on ne peut donc pas, pour $\mathfrak{A}$, prendre la boule unité, comme l'avions déjà signalé .

Le 2ème terme tend vers $\vec{0}$, d'après l'hypothèse de convergence vague; le premier aussi, puisque sa norme est majorée par $M \| \varphi_n - \varphi \|$, ce qui démontre le corollaire.

Les parties compactes de $\mathcal{C}_K(X)$

Pour pouvoir appliquer le corollaire 1, il nous faudra connaître les ensembles compacts de $\mathcal{C}_K(X)$.

L'étude générale de ces compacts sera faite ultérieurement, nous allons nous borner à des ensembles particuliers.

Théorème 66 - Soient A, B, C , trois espaces topologiques, A compact, C métrique. Soit f une application continue de $A \times B$ dans C . Alors, pour tout y de B , on peut définir une application partielle $f_y : x \to f(x,y)$, de A dans C : $f_y \in (C^A)_{cb}$.

Si y converge vers b dans B , la fonction partielle f_y converge vers la fonction partielle f_b , uniformément sur A ; ou encore, l'application $y \longrightarrow f_y$, qui. à tout élément y de B, fait correspondre la fonction partielle f_y , est une application continue de B dans $(C^A)_{cb}$.

Démonstration - Fixons-nous b dans B , et donnons-nous $\varepsilon > 0$. Quel que soit a , la fonction f est continue au point (a, b) ; donc il existe un voisinage ouvert U_a de a et un voisinage V_a de b , dans A et B respectivement, tels que, pour x dans U_a et y dans V_a , on ait l'inégalité

$$d\left(f(x,y),\, f(a,b)\right) \leq \frac{\varepsilon}{2} \,. \tag{IV,7;18}$$

On aura naturellement a fortiori l'inégalité

$$d\left(f(x,b),\, f(a,b)\right) \leq \frac{\varepsilon}{2} \,. \tag{IV,7;19}$$

On en déduit l'inégalité, valable pour $x \in U_a, y \in V_a$:

$$\begin{aligned} d\left(f(x,y),\, f(x,b)\right) &\leq d\left(f(x,y),\, f(a,b)\right) \\ &+ d\left(f(a,b),\, f(x,b)\right) \leq \varepsilon \end{aligned} \tag{IV,7;20}$$

Lorsque α varie, les U_α forment un recouvrement ouvert du compact A; donc il existe un nombre fini d'entre eux, $(U_{a_i})_{i\in I}$, qui forment encore un recouvrement de A. A ces U_{a_i} aorrespondent des V_{a_i}; soit V l'intersection de ces ensembles V_{a_i}, c'est encore un voisinage de b dans B.

Alors, pour x dans U_{a_i} et y dans V on a l'inégalité (IV, 7;20). Comme les U_{a_i} forment un recouvrement de A a l'inégalité précédente pour touts de A et tout y de"?, ce qui s'écrit encore :

$d(f_y, f_b) \leq \varepsilon$ pour y dans V; et ceci démontre le théorème.

Remarque - Si les espaces A et B sont tous les deux métriques et compacts, la démonstration est beaucoup plus élémentaire. En effet, dans ce cas, f est uniformément continue, d'après le théorème 31 du chapitre II. Etant donné $\varepsilon > 0$, il existe donc $\eta > 0$ tel que $d(x', x'') \leq \eta$, $d(y', y'') \leq \eta$, entraîne $d(f(x', y'), f(x'', y'')) \leq \varepsilon$. En particulier $d(y, b) \leq \eta$ entraîne, quel que soit x, l'inégalité $d(f(x,y), f(x,b)) \leq \varepsilon$, et par conséquent $d(f_y, f_b) \leq \varepsilon$, ce qui démontre le théorème.

Corollaire 1 - Soient A et B des espaces compacts, C un espace métrique, une application continue de $A \times B$ dans C. Alors, lorsque y 'parcourt B, l'ensemble des applications partielles f_y dé A dans C est une partie compacte de l'espace $(C^A)_{cb}$.

En effet c'est l'image directe d'un compact B par une application continue, $y \to f_y$.

Corollaire 2 - Soient X *un* espace localement compact, T un espace compact. Soit φ une fonction, à valeurs réelles ou complexes, continue sur $X \times T$. Alors, quel que soit t de T, on a une application partielle φ_t, qui est une fonction réelle ou complexe, continue sur X. Supposons que, lorsque t parcourt T, cette fonction φ_t ait toujours son support dans un même compact K de X. Alors, lorsque t parcourt T, l'ensemble des fonctions φ_t est une partie compacte de l'espace $\mathcal{C}_K(X)$.

Il suffit d'appliquer le corollaire 1 à $f = \varphi$, $A = K$, $B = T$, $C = \mathbb{R}$ ou $\mathbb{C}$

Corollaire 3 - Supposons vérifiées les conditions du corollaire 2. Soient $\vec{\mu}_n$ des mesures de Radon sur X, à valeurs dans un espace vectoriel normé E, qui, pour n, tendant vers l'infini, convergent vaguement vers une mesure limite $\vec{\mu}$. Alors les intégrales $\int \varphi(x,t)\, d\vec{\mu}_n(x)$ convergent, pour n tendant vers l'infini, vers $\int \varphi(x,t)\, d\vec{\mu}(x)$, uniformément pour t dans T.

Convergence vague d'une suite de mesures vers une mesure de Dirac

Théorème 67 - Soit μ_n une suite de mesures ≥ 0 sur un espace localement-compact X. Supposons que, pour tout voisinage compact $\mathcal{V}$ du point a de X, la suite des intégrales $\int_{\mathcal{V}} d\mu_n$ converge vers 1, pour n tendant vers l'infini.

Alors les mesures μ_n convergent vaguement vers la mesure de Dirac $\delta_{(a)}$.

Avant de commencer la démonstration, remarquons que notre hypothèse implique que, quel que soit le compact K de X ne contenant pas a, $\int_K d\mu_n$ converge vers 0 pour n tendant vers l'infini.

En effet, le complémentaire de K est un ouvert de X donc encore un espace localement compact * . Soit alors $\mathcal{V}$ un voisinage compact de a dans cet espace, autrement dit un voisinage compact de a dans X, sans point commun avec K.

Alors, d'une part, $\int_{\mathcal{V}} d\mu_n$ converge vers 1, d'autre part $\int_{\mathcal{V} \cup K} d\mu_n$ converge vers 1, donc $\int_K d\mu_n$ converge bien vers 0 .

Les conditions de l'énoncé expriment qu'a distance finie la masse totale tend vers 1 et tend à se concentrer au point a.

Démonstration - Soit φ une fonction de $\mathcal{C}(X)$, de support compact K . D'après la continuité de φ, $\varepsilon > 0$ étant donné, il existe un ouvert contenant a tel que, pour x dans cet ouvert, on ait l'inégalité $|\varphi(x) - \varphi(a)| \leq \frac{\varepsilon}{4}$. Si alors $\mathcal{V}$ est un voisinage compact de a dans cet ouvert, on a a fortiori cette inégalité sur $\mathcal{V}$. On peut alors décomposer la différence $\left(\int \varphi\, d\mu_n\right) - \varphi(a)$ en somme de trois termes, à savoir :

$$\left(\int \varphi\, d\mu_n\right) - \varphi(a) = \int_{\complement\mathcal{V}} \varphi\, d\mu_n + \int_{\mathcal{V}} (\varphi(x) - \varphi(a))\, d\mu_n(x) + \varphi(a)\left[\left(\int_{\mathcal{V}} d\mu_n\right) - 1\right]. \tag{IV,7;21}$$

* Voir note * page 437 et note * page 452.

Majorons successivement les modules de ces trois termes.

1°/ Le premier tend vers 0 lorsque n tend vers l'infini. En effet, comme φ a son support dans K , l'intégrale sur $\complement \mathcal{V}$ est en réalité une intégrale sur l'intersection $K \cap \complement \mathcal{V}$. Elle est alors majorée par $|||\varphi||| \int_A d\mu_n$,où A est l'adhérence de $K\cap\complement\mathcal{V}$. Or A est un compact, puisque c'est un ensemble fermé contenu dans le compact K ; d'autre part, il ne contient pas le point a , car A étant contenu dans $\complement \mathcal{V}$, a n'est pas adhérent à A .

Alors $\int_A d\mu_n$ converge bien vers 0 , d'après la remarque qui précède la démonstration.

2°/ Examinons ensuite la 3ème terme.

D'après l'hypothèse, la quantité entre crochets converge vers 0 , par conséquent le 3ème terme lui aussi converge vers 0 , pourn tendant vers l'infini. Il existe donc un entier p' tel que $n \geq p'$ entraîne que la somme des valeurs absolues du premier et du troisième terme soit $\leq \frac{\varepsilon}{2}$

3°/ Examinons maintenant le 2ème terme.

D'après le choix de $\mathcal{V}$, il est majoré par $\frac{\varepsilon}{4}\int_{\mathcal{V}} d\mu_n$. La dernière intégrale tend vers 1 pour n infini; donc il existe un entier p'' tel que cette intégrale soit $\leq$ 2 pour $n \geq p''$; alors, pour $n \geq p''$, le deuxième terme est majoré en module par $\frac{\varepsilon}{2}$. Alors, si p est le plus grand des entiers p', p'', on voit que l'on aura bien, pour $n \geq p$, l'inégalité $\left|\left(\int \varphi \, d\mu_n\right) - \varphi(a)\right| \leq \varepsilon$, et le théorème est démontré.

<u>Exemple 1</u> - Soit μ_n une suite de mesures ≥ 0 , dont la masse totale $\int d\mu_n$ converge vers 1, et dont le support converge uniformément vers le point a . Nous entendons par là que, quel que soit le voisinage $\mathcal{V}$ de a , le support de μ_n est contenu dans $\mathcal{V}$, pour n assez grand. Alors μ converge vaguement vers $\delta_{(a)}$ pour n infini. En effet les conditions du théorème sont trivialement réalisées.

<u>Exemple 1 bis</u> - Comme cas particulier, soit f_n une suite de fonctions ≥ 0 sur la droite réelle $\mathbb{R}$, intégrables pour la mesure dx , telles que $\int f_n(x)\,dx$ converge vers 1 pour n infini, et que le support de f_n converge uniformément vers le point a . Alors les mesures $f_n\,dx$ convergent vaguement vers $\delta_{(a)}$.

Cet exemple donne une interprétation nouvelle de certains cas où le théorème de convergence de Lebesgue n'est pas applicable.

Considérons par exemple la suite des fonctions f_n sur $\mathbb{R}$ définies comme suit :

(IV,7;22)
$$f_n(x) = \begin{cases} n \text{ pour } 0 < x \leq \frac{1}{n} \\ 0 \text{ ailleurs} \end{cases}$$

Ces fonctions convergent partout vers 0 pour n tendant vers l'infini. Mais elles ne sont pas majorées par une fonction ≥ 0 localement intégrable pour la mesure dx (voir formule (IV,4;11)). Le théorème de Lebesgue n'est donc pas applicable; et la meilleure preuve, c est que $\int f_n\, dx$ est constamment égal à 1, et par conséquent ne tend pas vers 0 .

Si alors φ appartient à $\mathscr{C}(\mathbb{R})$, $\int \varphi f_n\, dx$ ne converge pas vers 0 . Nous avions constaté cela à propos du théorème de Lebesgue comme un fait purement négatif. Nous l'inter- -maintenant d'une façon nouvelle : il résulte de ce qui est dit à l'exemple 1 bis que les masures $f_n\, dx$ convergent vaguement, non pas vers la mesure 0 , **mais vers la mesure δ** .

Alors $\int f_n \varphi\, dx$ converge vers $\varphi(0)$. C'est là un phénomène qui se produira souvent. Quand on verra des fonctions f_n convergeant simplement presque partout vers une fonction limite f , mais telles qu'il n'y ait pas convergence des intégrales, cela signifiera souvent * que les mesures $f_n\, dx$ convergent vers la somme de la mesure $f\, dx$, et d'une mesure présentant certaines masses ponctuelles.

Exemple 2 - soit f une fonction ≥ 0 sur l'espace $\mathbb{R}^n$, intégrable pour la mesure $dx = dx_1\, dx_2 \dots dx_n$ avec $\int f\, dx = 1$.

Si alors on pose, pour k réel $\neq 0$:

(IV,7;23)
$$\mu_k = |k|^n f(kx)\, dx ,$$

* Mais évidemment pas toujours ! Si, au lieu den , comme dans (IV,7;22), nous avions pris n^α , comme dans (IV,4;11), avec $\alpha > 1$, il n'y aurait rien eu de semblable.

μ_k est l'image de $f\,dx$ par l'homothétie $x \longrightarrow \frac{x}{k}$, (voir formule (IV,6;28), pour $n = 1$; et ultérieurement, théorème 102 pour n quelconque) . Les mesures μ_k convergent vaguement vers la mesure de Dirac δ pour k tendant vers l'infini. On peut le voir en appliquant le théorème 67. On peut le voir aussi directement, par changement de variables. Pour $\varphi \in \mathcal{C}(\mathbb{R}^n)$, on a :

$$\mu_k(\varphi) = \int \varphi(x)\, f(kx)\, |k|^n dx . \tag{IV,7;24}$$

En faisant le changement de variable $kx = \xi$, on a (voir ultérieurement, théorème 102) :

$$\mu_k(\varphi) = \int \varphi\left(\frac{\xi}{k}\right) f(\xi)\, d\xi . \tag{IV,7;25}$$

Lorsque k tend vers l'infini, $\frac{\xi}{k}$ tend vers 0 , alors, comme φ est continue, $\varphi\left(\frac{\xi}{k}\right)$ converge simplement vers $\varphi(0)$, donc la fonction $\xi \longrightarrow \varphi\left(\frac{\xi}{k}\right) f(\xi)$ converge simplement vers $\varphi(0)\, f(\xi)$. Comme elle reste bornée par $|||\varphi|||\, f(\xi)$, qui est ≥ 0 et intégrable pour $d\xi$, le théorème 35 de Lebesgue dit que $\mu_k(\varphi)$ converge pour k infini vers $\varphi(0)\int f(\xi)\, d\xi = \varphi(0)$, d'où la conclusion.

Mais il y a encore une démonstration plus directe. On voit que $f\,dx$ est une mesure ≥ 0 de masse totale 1; l'homothétie de centre 0 et de rapport $\frac{1}{k}$ draine toutes les masses vers l'origine pour k infini. L'exemple 3° page 553 nous donne alors la plus courte démonstration : l'homothétie H_k de rapport $\frac{1}{k}$ converge simplement, pour k infini, vers l'homothétie H de rapport nul, c'est-à-dire vers l'application constante $x \rightarrow 0$, donc $H_k\mu$ converge vers $H\mu$; mais $H\mu = \delta$ d'après l'exemple 1°/ page 546 , puisque la masse de μ est + 1.

Voici 2 exemples fréquemment utilisés :

a) La fonction f sur $\mathbb{R}$ est définie par :

$$f(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}} . \tag{IV,7;26}$$

Elle vérifie bien les conditions de l'énoncé, car, comme nous le verrons plus tard, l'intégrale de f est égale à 1.

Alors la suite des mesures $\sqrt{n}\ f(\sqrt{n}\,x)\,dx$:

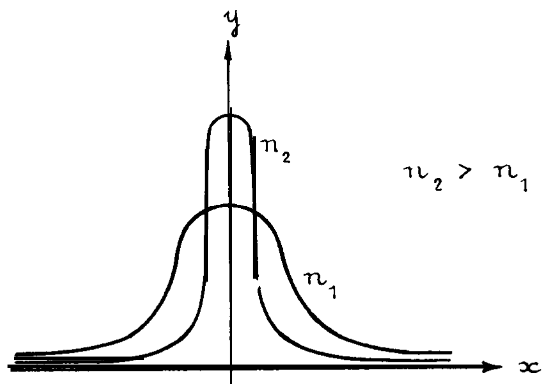
(IV,7 27) $$\mu_n = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{nx^2}{2}}\ \sqrt{n}\ dx$$

converge vaguement vers δ , pour n tendant vers l'infini.

On verra sur la figure la représentation graphique de 2 de ces courbes en cloche, qui donnent les densités des mesures μ_n par rapport à la mesure de Lebesgue, c'est-à-dire les fonctions $\sqrt{n}\ f(\sqrt{n}\ x)$.

On peut aussi prendre

(IV,7 ; 28) $$f(x) = e^{-\pi x^2} \quad ; \quad \sqrt{n}\ f(\sqrt{n}\ x) = \sqrt{n}\ e^{-n\pi x^2} .$$



b) La fonction f est donnée par la formule

(IV,7;29) $$f(x) = \frac{1}{\pi}\frac{1}{1+x^2} .$$

Son Intégrale est bien évidemment 1, alors les mesures

(IV,7;30) $$\mu_n = \frac{n\ dx}{\pi(1+n^2x^2)}$$

convergent vaguement vers δ , pour n tendant vers l'infini.

<u>Remarque</u> Ces propriétés auraient été Impossibles sans l'introduction de la convergence vague : si K est un intervalle compact $[-A,+A]$, $\|\mu_k - \delta\|_K = \|f(kx)|k|^n dx - \delta\|_K$, d'après ce que nous avons dit à propos de la formule (IV,2;16), est <u>égal</u> à $\int_K f(kx)\,|k|^n dx + 1$ vers 2 pour k Infini; il n'y a pas convergence locale en norme !

Théorème **68** - Il existe une suite de polynomes $P_n \geq 0$ d'une variable réelle, tels que les $P_n\,dx$ convergent vaguement vers δ pour n infini.

Ce théorème a une grande importance; on en déduira ultérieurement le théorème de Weierstrass (approximation des fonctions continues par des polynomes); théorème 80 bis.

Démonstration -- Soit e_n une suite de nombres > 0, tendant vers $+\infty$ pour n infini. Posons :

$$P_n(x) = \frac{e_n}{\sqrt{2\pi}}\left(1 - \frac{e_n^2 x^2}{2n}\right)^n, \tag{IV,7;31}$$

qui est ≥ 0 si n est pair.

Montrons que $P_n\,dx$ converge vaguement vers δ pour n infini, pourvu que $e_n/\sqrt{n}$ tende vers 0 . On pourrait utiliser le théorème 67; c'est aussi simple directement. **soit** $\varphi \in \mathscr{C}(\mathbb{R})$

$$(P_n\,dx)(\varphi) = \int \frac{e_n}{\sqrt{2\pi}}\left(1 - \frac{e_n^2 x^2}{2n}\right)^n \varphi(x)\,dx\,. \tag{IV,7;32}$$

soit $[-A, +A]$ un intervalle contenant le support de φ; cette intégrale est en réalité calculée sur $[-A, +A]$.

Faisons le changement de variables $e_n x = \xi$:

$$(P_n\,dx)(\varphi) = \frac{1}{\sqrt{2\pi}}\int\left(1 - \frac{\xi^2}{2n}\right)^n \varphi\left(\frac{\xi}{e_n}\right)d\xi\,, \tag{IV,7;33}$$

l'intégrale étant en réalité calculée sur $[-e_n A, +e_n A]$.

Lorsque n tend vers l'infini, pour ξ fixé, $\varphi\left(\frac{\xi}{e_n}\right)$ tend vers $\varphi(0)$, puisque φ est continue à l'origine; $\left(1 - \frac{\xi^2}{2n}\right)^n$ tend vers $e^{-\frac{\xi^2}{2}}$; donc la fonction à intégrer converge simplement vers la fonction $\varphi(0)e^{-\frac{\xi^2}{2}}$. Si le théorème **35** de Lebesgue est applicable, le second membre de (IV,7;33) converge vers $\frac{\varphi(0)}{\sqrt{2\pi}}\int e^{-\frac{\xi^2}{2}}d\xi = \varphi(0)$, et le théorème sera démontré.

Majorons donc la fonction à intégrer.

Pour $0 \leq u < 1$, le développement en série du logarithme montre que $\log(1-u) \leq -u$.

Donc $1 - u \leqslant e^{-u}$: et par suite $\left(1-\frac{\xi^2}{2n}\right)^n \leqslant e^{-\frac{\xi^2}{2}}$, si $\frac{\xi^2}{2n} < 1$.

Mais la fonction à intégrer est égale à $\varphi\left(\frac{\xi}{e_n}\right)\left(1-\frac{\xi^2}{2n}\right)^n$, donc majorée par $\|\varphi\| \left(1 - \frac{\xi^2}{2n}\right)^n$ pour $|\xi| \leqslant e_n A$, et nulle pour $|\xi| > e_n A$. Pour $|\xi| \leqslant e_n A$, on a $\frac{\xi^2}{2n} \leqslant \frac{A^2}{2}\frac{e_n^2}{n}$; puisque $\frac{e_n}{\sqrt{n}}$ tend vers 0 pour n infini, il existe un entier p tel que $n \geqslant p$ entraîne $\frac{A^2}{2}\frac{e_n^2}{n} < 1$.

Alors on a, pour $n \geqslant p$, la majoration

(IV,7;34) $$\left|\left(1-\frac{\xi^2}{2n}\right)^n \varphi\left(\frac{\xi}{e_n}\right)\right| \leqslant \|\varphi\| e^{-\frac{\xi^2}{2}},$$

et le second membre est intégrable, donc le théorème de Lebesgue est applicable, et les $P_n\,dx$ convergent bien vers δ.

Convergence étroite d'une suite de mesures de norme finie

Soit $\vec{\mu}$ une mesure de Radon sur X, à valeurs dans un espace vectoriel normé $\vec{E}$ de dimension finie, et de norme finie.

On sait alors, d'après le corollaire du théorème 54 bis, qu'on peut définir $\int \varphi\, d\vec{\mu}$, pour toute fonction φ scalaire, continue et bornée sur X.

On est donc amené à poser la définition suivante qui sera fondamentale en théorie des probabilités.

Définition On dit qu'une suite de mesures $\vec{\mu}_n$ sur X, à valeurs dans un esoace vectoriel normé $\vec{E}$ de dimension finie, de normes finies *, converge étroitement vers une mesure $\vec{\mu}$ de norme finie, si, quelle que soit la fonction φ scalaire, continue et bornée sur X, $\int \varphi\, d\vec{\mu}_n$ converge, pour n infini, vers $\int \varphi\, d\vec{\mu}$.

Naturellement la convergence étroite entraîne a fortiori la convergence vague; cependant elle est moins forte que la convergence en norme, qui exprimerait que $\|\vec{\mu}_n - \vec{\mu}\|$ converge vers 0, et que par conséquent $\int \varphi\, d\vec{\mu}_n$ converge vers $\int \varphi\, d\vec{\mu}$, uniformément lorsque φ reste bornée en norme sur X.

* Nous disons "finies" et non bornées; nous verrons précisément au théorème 69 que cela entraîne que les $\|\vec{\mu}_n\|$ soient bornées dans leur ensemble.

Si les $\vec{\mu}_n$ convergent-étroitement vers $\vec{\mu}$, on voit, en prenant $\varphi \equiv 1$, que $\int d\mu_n$ converge vers $\int d\vec{\mu}$; en particulier, si les μ_n et μ sont des mesures ≥ 0, $\|\mu_n\|$ converge vers $\|\mu\|$.

Théorème 69 - Si des mesures $\vec{\mu}_n$ sur X, à valeurs dans un espace vectoriel normé $\vec{E}$ de dimension finie, convergent étroitement vers $\vec{\mu}$ pour n infini, les $\|\vec{\mu}_n\|$ sont bornées dans leur ensemble.

Démonstration Chaque $\vec{\mu}_n$ définit une application linéaire continue u_n de l'espace $(\mathbb{K}^\times)_{cb}$, des fonctions scalaires continues bornées sur X, dans $\vec{E}$, à savoir : $\varphi \longrightarrow \int \varphi \, d\vec{\mu}_n$.

D'après la définition même de $\|\vec{\mu}_n\|$ comme borne supérieure de $\|\vec{\mu}_n(\varphi)\|$, pour $\varphi \in \mathcal{C}(X)$ et $|||\varphi||| \leq 1$, on a sûrement $\|u_n\| \geq \|\vec{\mu}_n\|$ *. Mais $(\mathbb{K}^\times)_{cb}$ est un espace de Banach (corollaire 3 du théorème 65 du chapitre II), donc le théorème 65 de Banach-Steinhaus dit que les u_n sont bornées, donc a fortiori les $\|\vec{\mu}_n\|$.

Soit μ une mesure ≥ 0, de norme finie. On appellera intégrale supérieure de Riemann par rapport à μ d'une fonction $f \geq 0$ bornée (à support quelconque) la borne inférieure des intégrales des fonctions continues ≥ 0 qui la majorent. Si f est à support compact, on retrouve la définition de la page 554; car, si g est continue bornée et majore f, et si α est une fonction continue à support compact, ≥ 0, égale à 1 sur un voisinage du support de f, αg majore encore f et $\int^{*R} f \, d\mu \leq \int \alpha g \, d\mu \leq \int g \, d\mu$. Nous appellerons fonction continue bornée décomposable sur X à valeurs dans $\vec{F}$ une somme finie $\sum_{i=1}^{N} \vec{g}_i \varphi_i$, où les $\vec{g}_i$ sont des vecteurs de $\vec{F}$, et les φ_i des fonctions scalaires continues bornées. On dira qu'une fonction $\vec{f}$ sur X, à valeurs dans $\vec{F}$ est μ-intégrable-Riemann, si elle admet une suite d'approximation, au sens de l'intégrale supérieure de Riemann, par des fonctions continues bornées décomposables. Alors :

* On démontre même que $\|u_n\| = \|\vec{\mu}_n\|$

Théorème 69 bis - Si les μ_n sont une suite de mesures ≥ 0 de normes finies, convergeant étroitement vers une mesure μ de norme finie pour n tendant vers l'infini, si φ est une fonction, réelle ou complexe, définie sur X, bornée, mesurable par rapport à toutes les μ_n, et μ-encadrable, alors $\int \varphi \, d\mu_n$ converge vers $\int \varphi \, d\mu$.

Démonstration analogue à celle du théorème 62.

Théorème 70 - Une fonction f sur X à valeurs dans F, bornée, est μ-intégrable-Riemann, si et seulement si l'ensemble de ses points de discontinuité est de μ-mesure (de Lebesgue) nulle.

Démonstration analogue à celle du théorème 63.

Corollaire 1 - Pour qu'une partie A de X ait une fonction caractéristique μ-intégrable-Riemann, il faut et il suffit que sa frontière $\dot{A}$ ait une μ-mesure (de Lebesgue) nulle.

Corollaire 2 - Si des mesures $\mu_n \geq 0$ de normes finies convergent étroitement vers une mesure μ de norme finie, et si A est une partie de X, mesurable pour toutes les μ_n et dont la frontière est de mesure nulle pour μ, alors les $\mu_n(A)$ convergent vers $\mu(A)$ pour n infini.

Convergence vague et convergence étroite.

Théorème 71 - Pour qu'une suite de mesures $\mu_n \geq 0$ de normes finies sur X convergent étroitement vers une mesure $\mu \geq 0$ de norme finie, il faut et il suffit qu'elles convergent vaguement vers μ, et que les $\|\mu_n\| = \int d\mu_n$ convergent vers $\|\mu\| = \int d\mu$.

Démonstration - La condition est évidemment nécessaire, montrons qu'elle est suffisante.

Soit φ une fonction continue bornée sur X, $M = |||\varphi|||$, et soit $\varepsilon > 0$ donné. D'après la définition même de la norme de μ, il existe une fonction α, continue, à support compact, $0 \leq \alpha \leq 1$, telle que

(IV,7;36) $$\int \alpha \, d\mu \geq \left(\int d\mu\right) - \frac{\varepsilon}{4M}.$$

Comme alors les mesures μ_n convergent vaguement vers μ pour n infini, les $\int \alpha \, d\mu_n$ convergent vers $\int \alpha \, d\mu$; comme par ailleurs $\int d\mu_n$ converge par hypothèse vers $\int d\mu$, on voit qu'il existe un entier p' tel que, pour $n \geq p'$, on ait

$$\int \alpha \, d\mu_n \geq \int d\mu_n - \frac{\varepsilon}{3M} \quad ; \text{ ou} \tag{IV,7;37}$$

$$\int (1-\alpha)\, d\mu_n \leqslant \frac{\varepsilon}{3M} \quad ; \text{ et } \int (1-\alpha)\, d\mu \leqslant \frac{\varepsilon}{4M} \tag{IV,7;38}$$

On a alors la décomposition suivante, pour $\varphi \in \mathscr{C}(X)$:

$$\mu_n(\varphi) - \mu(\varphi) = \Big(\mu_n(\alpha\varphi) - \mu(\alpha\varphi)\Big) + \mu_n\Big((1-\alpha)\varphi\Big) - \mu\Big((1-\alpha)\varphi\Big). \tag{IV,7;39}$$

Le premier terme converge vers 0 , en vertu de l'hypothèse de convergence vague; donc il existe un entier p'' tel que $n \geq p''$ entraîne

$$\Big|\mu_n(\alpha\varphi) - \mu(\alpha\varphi)\Big| \leqslant \frac{\varepsilon}{3} . \tag{IV,7;40}$$

D'après les majorations (IV,7;38) pour les 2 autres termes, on a donc, pour $n \geqslant p = \operatorname{Sup.}(p', p'')$:

$$\Big|\mu_n(\varphi) - \mu(\varphi)\Big| \leqslant \frac{\varepsilon}{3} + M\frac{\varepsilon}{3M} + M\frac{\varepsilon}{4M} < \varepsilon \quad , \tag{IV,7;41}$$

et ceci démontre le théorème.

Remarque Si nous reprenons les exemples 1, 1 bis, 2 donnes pour la convergence vague vers δ (pages565 A568), on voit qu'il y a convergence étroite. De méme dans les exemples 1°, 3°, 4°, pages 552-554.

Si maintenant nous reprenons le théorème 68, il n'y a sûrement pas convergence étroite, car, si P_n est un polynome $\geqslant 0$, $\int P_n(x)\,dx = +\infty$, il ne s'agit. pas de mesures de normes finies.

Si a_n est une suite de points de X s'éloignant indéfiniment (lorsque X est métrique; voir corollaire du théorème 55 pour la définition), $\delta_{(a_n)}$ converge vaguement vers 0 ; car, pour $\varphi \in \mathscr{C}(X)$, a_n est, pourn assez grand, en dehors du support (compact donc borné) de φ , donc $\varphi(a_n) = 0$ pour n assez grand. Mals $\delta_{(a_n)}$ ne converge pas étroitement vers0 , puis que $\|\delta_{(a_n)}\| = 1$. Ceci nous montre la significa - tion physique de la convergence étroite. La convergence vague ne fait jamais Intervenir que ce qui se passe a distance finie; une suite de mesures $\mu_n \geqslant 0$ peut parfaitement converger vaguement vers une mesure μ , sans que cela empêche une partie importante des masses de μ_n de 'se perdre" à l'infini; la convergence étroite au contraire, l'empêche, puisque $\|\mu_n\|$ tend vers $\|\mu\|$.

Théorème 72 - Soient X un espace localement compact, T un espace compact. Soit φ une fonction scalaire continue et bornée sur $X \times T$. Soit d'autre part μ_n une suite de mesures de Radon ≥ 0 sur X , de normes finies, convergeant étroitement, pour n tendant vers l'infini, vers une mesure limite μ .

Alors les intégrales $\int \varphi(x,t)\,d\mu_n$ convergent vers l'intégrale $\int \varphi(x,t)\,d\mu$, uniformément lorsque t parcourt T.

Démonstration Ceci naturellement généralise le corollaire 3 du théorème 66 au cas de la convergence étroite.

Posons $M = \operatorname{Sup.}_{x,t} |\varphi(x,t)|$.

En fonction de ce nombre M , on peut trouver, si nous reprenons les notations de la démonstration du théorème 71, une fonction α vérifiant (IV,7;36), donc (IV,7;38) , pour $n \geq p'$. Effectuons alors la décomposition (IV,7;39) relativement à la fonction partielle $\varphi_t : x \to \varphi(x,t)$. Le permier terme, $\mu_n(\alpha\varphi_t)$, converge vers $\mu(\alpha\varphi_t)$, uniformément pour t dans T . En effet, nous nous trouvons cette fois exactement dans les conditions du corollaire 3 du théorème 66, puisque nous avons une fonction continue scalaire $(x,t) \to \alpha(x)\varphi(x,t)$, et que , pour tout t , la fonction $x \longrightarrow \alpha(x)\varphi(x,t)$, définie sur X, a son support dans un compact fixe, à savoir le support de α . On peut donc déterminer p'' indépendamment de t , de façon qu'on ait (IV,7;40) pour φ_t , pour $n \geq p''$. Quant au deuxième et au troisième termes, ils se majorent à partir de (IV,7;38), indépendamment de t . On a alors (IV, 7;41) pour $n \geq p = \operatorname{Sup}(p',p'')$, et pour tout $t \in T$, ce qui démontre le théorème.

§ 8 PRODUITS TENSORIELS DE MESURES. INTÉGRALES MULTIPLES

Position du problème

Soient X et Y deux espaces localement compacts. Alors $X \times Y$ est aussi localement compact, comme il résulte immédiatement de la définition des voisinages. En outre, si X et Y sont dénombrables à l'infini, $X \times Y$ l'est aussi. Si en effet les H_n constituent une suite de compacts de réunion X, et les K_n une suite de compacts de réunion Y, les $H_n \times K_n$ forment une suite de compacts de réunion $X \times Y$.

Soient alors μ une mesure de Radon sur X ν une mesure de Radon sur Y ; peut-on construire à partir d'elles une nouvelle mesure de Radon sur $X \times Y$? Nous nous bornerons aux mesures de Radon scalaires * Quand on écrit, sur l'espace $\mathbb{R}^n$, la mesure qui sert à la définition des intégrales multiples usuelles, a savoir $dx_1\, dx_2 \dots dx_n$, elle est bien un "produit" des mesures dx_i définies sur les espaces facteurs $\mathbb{R}$ du produit $\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$.

Nous allons montrer que, si μ et ν sont des mesures scalaires, il existe toujours une mesure produit ou produit tensoriel, qui se note $\mu \otimes \nu$, ou $\mu_x \otimes \nu_y$, ou $d\mu(x) \otimes d\nu(y)$, ou $d\mu(x)\, d\nu(y)$; l'intégrale d'une fonction $\vec{f}$ sur $X \times Y$, à valeursdans un espace de Banach $\vec{F}$, par rapport à cette mesure produit s'appellera intégrale double, et se notera avec le symbole habituel des intégrales doubles

(IV,8;1) $$\int \vec{f}\, d(\mu \otimes \nu) = \iint \vec{f}(x,y)\, d\mu(x)\, d\nu(y).^{**}$$

Soient u et v des fonctions appartenant respectivement à $\mathcal{C}(X)$ et $\mathcal{C}(Y)$. Considérons alors la fonction. φ définie sur $X \times Y$ par la formule

(IV,8;2) $$\varphi(x,y) = u(x)\, v(y) ;$$

on notera aussi cette L'onction par $u \otimes v$.

* Pour des mesures à valeurs dans des espaces vectoriels de dimension finie, on a des résultats analogues, mais non pour des mesures à valeurs dans des espaces vectoriels normés de dimension infinie

** La notation $d\mu(x)\, d\nu(y)$ est évidemment incorrecte, puisque les variables x et y sont muettes ! par contre, dans (IV,8;1), c'est une notation correcte, x et y figurent 2 fois, et peuvent être remplacées par d'autres symboles arbitraires. De même, si $\lambda = \mu \otimes$ 3, il est correct d'écrire

$$d\lambda(x,y) = d\mu(x) \otimes d\nu(y) \text{ ou } d\mu(x)\, d\nu(y).$$

Elle est visiblement continue, car la fonction $(x,y) \to x \to u(x)$ est continue, comme composée de dew fonctions continues; de même la fonction $(x,y) \longrightarrow v(y)$ est continue, et la fonction ~~considérée est le~~ produit des deux précédentes. D'autre part elle est à support compact; si H et K sont les supports respectifs de u et de v dans X et dans Y, son support est exactement le produit $H \times K$. Appelons en effet A et B les ensembles de points où u et v sont $\neq 0$; alors φ est $\neq 0$, si et seulement si u et v sont $\neq 0$, c'est-à-dire sur $A \times B$. Le support de φ est par conséquent l'adhérence $\overline{A \times B}$, qui est le produit $\bar{A} \times \bar{B}$ des adhérences, c'est-à-dire $H \times K$.

Dans ce cas. on désire évidemment que l'intégrale double s'exprime sous la forme

$$\iint u(x)\,v(y)\,d\mu(x)\,d\nu(y) = \left(\int u(x)\,d\mu(x)\right)\left(\int v(y)\,d\nu(y)\right) \tag{IV,8;3}$$

c'est-à-dire que l'on ait

$$(\mu \otimes \nu)(u \otimes v) = \mu(u)\,\nu(v) \tag{IV,8;4}$$

Nous allons voir que ceci est suffisant pour déterminer d'une manière unique la mesure produit tensoriel $\mu \otimes$ 3.

Existence et unicité du produit tensoriel

<u>Théorème 73</u> - <u>Si X et Y sont deux espaces localement compacts, si μ et ν sont des mesures de Radon, réelles ou complexes, sur X et Y respectivement. il existe une mesure de Radon λ et une seule sur $X \times Y$, vérifiant l'égalité</u>

$$\lambda(u \otimes v) = \mu(u)\,\nu(v)\,. \tag{IV,8;5}$$

<u>Cette mesure s'appelle le produit tensoriel de μ et ν, et se note</u> $\mu \otimes$ 3 .

Nous démontrerons ce théorème en plusieurs etapes.

<u>Première étape : unicité du produit tensoriel.</u>

Soient H et K des compacts de X et Y respectivement. Considérons d'abord toutes les fonctions Ω "décomposables", de la forme $u \otimes v$, où u appartient à $\mathscr{C}_H(X)$ et v à $\mathscr{C}_K(Y)$. Considérons ensuite toutes les fonctions φ qui sont sommes d'un nombre fini (mais non limité à l'avance) de fonctions décomposables. Une telle fonction Q peut alors s'écrire

$$\varphi(x,y)=\sum_{i\in I} u_i(x)\,v_i(y), \quad u_i \in \mathscr{C}_H(X),\; v_i \in \mathscr{C}_K(Y). \tag{IV,8;6}$$

L'ensemble de toutes ces fonctions φ est un sous-espace vectoriel de $\mathscr{C}(X,Y)$, que nous appellerons $\Gamma_{H\times K}$.
On a alors le théorème suivant :

<u>Théorème **74**</u> (<u>théorème de densité) Soient</u> H_0 <u>et</u> K_0 <u>des compacts de</u> X <u>et</u> Y <u>respectivement,</u> H et K <u>des voisinages compacts de</u> H_0 <u>et</u> K_0 <u>respectivement</u> * . A<u>lors, dans l'espace</u> $\mathscr{C}_{H\times K}(X\times Y)$, <u>l'adhérence du sous-espace vectoriel</u> $\Gamma_{H\times K}$ <u>oontient le sous-espace</u> $\mathscr{C}_{H_0\times K_0}(X\times Y)$.

Ce théorème signifie que toute fonction φ, continue sur $X\times Y$, de support dans $H_0\times K_0$, est limite uniforme d'une suite de fonctions, qui peuvent chacune s'écrire sous la forme (IV,8:6), avec des fonctions u_i de support dans H et des fonctions v_i de support dans K .

<u>Démonstration</u>. Nous nous bornerons a donner la démonstration lorsque X et **Y** sont métriques; nous munirons alors $X\times Y$ de l'une des métriques naturelles du produit. Soit φ un élément de $\mathscr{C}_{H_0\times K_0}(X\times Y)$. Elle est alors uniformément continue (voir page **552**) : quel que soit $\varepsilon>0$, il existe $\eta>0$ tel que $d(x',x'')\leq\eta$, $d(y',y'')\leq\eta$, entraîne $|\varphi(x',y')-\varphi(x'',y'')|\leq\varepsilon$. Chaque point de H_0 **est** alors centre d'une boule ouverte contenue dans H , et dont on peut supposer le rayon $\leq\frac{\eta}{2}$. L'ensemble de ces boules ouvertes est un recouvrement de H_0, et on peut donc en extraire un sous-recouvrement fini, formé par des boules U_i, $i\in I$, de rayons $\leq\frac{\eta}{2}$, et toutes contenues dans **H** . De même on peut former un recouvrement fini de K_0 par des boules V_j, $j\in J$, toutes de rayon $\leq\frac{\eta}{2}$ et toutes contenues dans K . Alors les $U_i\times V_j$ forment un recouvrement ouvert fini de $H_0\times K_0$, et sont en outre toutes contenues dans $H\times K$.

* De tels voisinages existent; voir note * page **440**.

Si, par ailleurs, on a $x' \in U_i, x'' \in U_i, y' \in V_j, y'' \in V_j$, on a $d(x', x'') \leq \eta$ et $d(y', y'') \leq \eta$, d'où l'on déduit l'inégalité $|\varphi(x', y') - \varphi(x'', y'')| \leq \varepsilon$. Posons $I' = I \cup \{0\}$, $J' = J \cup \{0\}$, et considérons les recouvrements ouverts $(U_i)_{i \in I'}$ de X et $(V_j)_{j \in J'}$ de Y, avec $U_0 = \complement H_0$, $V_0 = \complement K_0$. Soient alors $(\alpha_i)_{i \in I'}$ et $(\beta_j)_{j \in J'}$ des partitions de l'unité associées à ces recouvrements (celà suppose X et Y dénombrables à l'infini; on peut s'affranchir de cette hypothèse, mais nous ne le ferons pas). Choisissons dans chaque U_i un point x_i dans chaque V_j un point y_j, et formons la fonction Φ sur $X \times Y$, définie par :

(IV, 8;7) $$\Phi(x, y) = \sum_{\substack{i \in I' \\ j \in J'}} \alpha_i(x) \beta_j(y) \varphi(x_i, y_j)$$

Comme $\varphi(x_i, y_j) = 0$ si i ou j est 0, elle peut s'écrire aussi avec $\sum_{\substack{i \in I \\ j \in J}}$ au lieu de $\sum_{\substack{i \in I' \\ j \in J'}}$.

Elle est de la forme (IV, 8,6), relativement à l'ensemble d'indices $I \times J$, avec, par exemple, $u_{ij}(x) = \alpha_i(x)$, $v_{ij}(y) = \varphi(x_i, y_j) \beta_j(y)$. Alors u_{ij} a son support dans H, et v_{ij} dans K, dès que i et j sont $\neq 0$. Remarquons que les $\alpha_i \otimes \beta_j$, $(i, j) \in I' \times J'$, forment une partition de l'unité de $X \times Y$, subordonnée au recouvrement par les $U_i \times V_j$, $(i, j) \in I' \times J'$. On a donc $\varphi(x, y) = \sum_{i,j} \alpha_i(x) \beta_j(y) \varphi(x, y)$, la sommation portant indifféremment sur $I \times J$ ou $I' \times J'$.

On a l'inégalité

(IV, 8;8) $$|\Phi(x, y) - \varphi(x, y)| \leq \sum_{i,j} \alpha_i(x) \beta_j(y) \, |\varphi(x_i, y_j) - \varphi(x, y)|$$

Le terme d'indice (i, j) du 2ème membre n'est $\neq 0$ que pour $i \neq 0$, $j \neq 0$, $(x, y) \in U_i \times V_j$; mais alors $|\varphi(x_i, y_j) - \varphi(x, y)| \leq \varepsilon$; le second membre est donc majoré par $\sum_{i,j} \alpha_i(x) \beta_j(y) = \varepsilon$, et le théorème est démontré.

<u>Corollaire - Si les mesures μ et ν sont données sur X et Y, il existe au plus une mesure λ satisfaisant. aux conditions du théorème 73.</u>

En effet, étant donné une fonction φ de $\mathscr{C}_{H \times K}(X \times Y)$, il existe une suite de fonctions ψ_n de $\mathscr{C}_{H \times K}(X \times Y)^{H_0 \times K_0}$, de la forme (IV,8;6), qui, pour n Infini, convergent uniformément vers φ. On a donc nécessairement, si λ existe, $\lambda(\varphi) = \lim_{n \to \infty} \lambda(\psi_n)$.

Mais, si λ existe, sa valeur est entièrement connue sur les fonctions décomposables $u \otimes v$, en vertu de l'égalité (IV,8;5), donc, par addition, sur toutes les fonctions ψ_n; et le passage à la limite précédent montre que $\lambda(\varphi)$ est entiérement connue pour toute φ de $\mathcal{C}(X \times Y)$, ce qui démontre bien l'unicité * .

Cette situation ressemble évidemment à celle du théorème 49 du chapitre II (prolongement des fonctions linéaires continues définies sur un sous-ensemble dense); mais, ici, il existe une petite difficulté, tenant à ce que $\mathcal{C}(X \times Y)$ n' a pas été considéré comme un espace topologique; c'est ce qui vient de nous obliger à faire une démonstration directe du corollaire.

Deuxième étape : existence du produit tensoriel,

Théorème 75 - Soit φ une fonction de $\mathcal{C}(X \times Y)$. Alors, pour tout x fixé dans X, on peut calculer l'intégrale

(IV,8;9) $$\psi(x) = \int \varphi(x,y)\, d\nu(y) .$$

La fonction $\psi : x \longrightarrow \psi(x)$, est continue et à support compact dans X ; il est donc possible de calculer

(IV,8;10) $$\lambda(\varphi) = \int \psi(x)\, d\mu(x) = \int\left[\int \varphi(x,y)\, d\nu(y)\right] d\mu(x) .$$

Alors la forme linéaire $\varphi \longrightarrow \lambda(\varphi)$ est une mesure de Radon sur $X \times Y$, la relation (IV,8;5).

* Inddpendamment de la recherche du produit tensoriel, nous démontrons plus : deux mesures λ_1, λ_2 sur $X \times Y$, qui prennent la même valeur pour toute fonction φ de $\mathcal{C}(X \times Y)$ de la forme $u \otimes v$, $u \in \mathcal{C}(X)$, $v \in \mathcal{C}(Y)$, prennent la même valeur pour toute fonction φ de $\mathcal{C}(X \times Y)$, donc sont égales.

<u>Démonstration</u> si nous appelons H et K les projections du support de φ sur X et Y respectivement, ce sont des compacts, comme images continues d'un compact. Alors la fonction partielle $\varphi_x : y \to \varphi(x,y)$, a son support dans K, quel que soit x. Par ailleurs $\varphi_x \equiv 0$, sauf pour $x \in H$.

Lorsque x tend vers a nous avons vu au théorème 66 que la fonction partielle φ_x tend vers la fonction partielle φ_a, uniformément sur le compact K (donc uniformément sur Y, puisqu'elle a toujours son support dans K), et par suite φ_x tend vers φ_a dans l'espace vectoriel normé $\mathcal{C}_K(Y)$. Il résulte alors de la définition de la mesure ν comme forme linéaire sur $\mathcal{C}(Y)$, dont la restriction à $\mathcal{C}_K(Y)$ est continue, que $\psi(x) = \int \varphi_x \, d\nu$ tend vers $\psi(a) = \int \varphi_a \, d\nu$. Autrement dit, ψ est une fonction continue sur X. Par ailleurs, elle a son support dans H, car pour $x \notin H$, $\varphi_x \equiv 0$ donc $\psi(x) = 0$. Donc elle appartient bien à $\mathcal{C}(X)$, et il est bien possible de calculer $\mu(\psi)$, donc l'expression (IV,8;10) a bien un sens. Il est évident qu'elle dépend linéairement de 'y.

Par ailleurs, si φ garde son support dans un compact fixe H Y K alors ψ garde son support dans H, et vérifie la majoration

(IV,8;11) $$|\psi(x)| \leq |||\varphi||| \quad \|\nu\|_K \; .$$

Il en résulte que l'on a la majoration

(IV,8;12) $$|\lambda(\varphi)| \leq \|\mu\|_H \, |||\psi||| \leq \|\mu\|_H \, \|\nu\|_K \, |||\varphi||| .$$

Ceci prouve bien que λ est une mesure sur $X \times Y$, et en outre on a la majoration

(IV,8;13) $$\text{II } \lambda\|_{H \times K} \leq \|\mu\|_H \|\nu\|_K \; .$$

Il reste donc seulement à vérifier que λ satisfait à la relation (IV,8;5). Si donc nous posons $\varphi(x,y) = u(x)v(y)$, on a $\psi(x) = u(x)\int v(y)d\nu(y) = u(x)\nu(v)$, donc $\lambda(\varphi) = \nu(v)\int u(x)d\mu(x) = \mu(u)\nu(v)$, et le théorème est démontré.

<u>Corollaire - Si μ et ν sont données, il existe au moins une mesure λ satisfaisant aux conditions du théorème 73.</u>

Ainsi le théorème 73 est démontré, par les démonstrations, sans rapport l'une avec l'autre, de l'existence et de l'unicité.

<u>Remarque</u> Nous aurinns pu au contraire commencer par fixer y, et calculer l'intégrale

(IV,8;14) $$\theta(y) = \int \varphi(x, y)\, d\mu(x)$$

La fonction obtenue θ aurait encore été une fonction continue à support compact sur Y , et on aurait pu calculer

(IV,8;15) $$\lambda(\varphi) = \int \theta(y)\, d\nu(y) = \int\left[\int \varphi(x,y)\, d\mu(x)\right] d\nu(y).$$

On aurait ainsi défini une deuxième mesure de Radon λ, satisfaisant à la relation (IV,8;5). L'unicité démontrée au corollaire du théorème 74 nous affirme que c'est la même.

On a donc la relation

(IV,8;16) $$\int\left[\int \varphi(x,y)\, d\nu(y)\right] d\mu(x) = \int\left[\int \varphi(x,y)\, d\mu(x)\right] d\nu(y);$$

l'égalité entre ces deux derniers termes constitue ce qu'on appelle le théorème d'interversion des signes d'intégration.

Exemples de produits tensoriels

Si a est un point de X et b un point de Y , on a entre les mesures de Dirac la relation évidente suivante

(IV,8;17) $$\delta_{(a)} \otimes \delta_{(b)} = \delta_{(a,b)} \quad \text{ou} \quad d\delta_{(a)}(x)\, d\delta_{(b)}(y) = d\delta_{(a,b)}(x,y).$$

En effet le second membre vérifie (IV,8;5).
Si maintenant μ est la mesure $\sum\limits_i c_i\, \delta_{(a_i)}$ et 3 la mesure $\sum\limits_j d_j\, \delta_{(b_j)}$, on a la formule

(IV,8;18) $$\sum_i c_i\, \delta_{(a_i)} \otimes \sum_j d_j\, \delta_{(b_j)} = \sum_{i,j} c_i\, d_j\, \delta_{(a_i, b_j)}$$

Propriétés élémentaires

Il est bien évident que l'application $(\mu,\nu) \longrightarrow \mu\otimes 3$ de $\mathscr{C}'(X) \times \mathscr{C}'(Y)$ dans $\mathscr{C}'(X \times Y)$ est une application bilinéaire.*

Autrement dit on a

$$(\text{IV}, 8\,; 19)\qquad \begin{cases} (\mu_1+\mu_2)\otimes(\nu_1+\nu_2) = \mu_1\otimes\nu_1 + \mu_1\otimes\nu_2 + \mu_2\otimes\nu_1 + \mu_2\otimes\nu_2\,. \\ h\mu\otimes k\nu = hk(\mu\otimes\nu)\,, \quad h \text{ et } k \quad \text{scalaires.} \end{cases}$$

Il est bien évident aussi que, si μ et ν sont réelles, $\mu\otimes\nu$ est réelle; et que, si μ et 3 sont $\geqslant 0$, $\mu\otimes 3$ est 30 , car il suffit, pour $\varphi \geqslant 0$, de faire le calcul du théorème 75, pour vérifier que $(\mu\otimes\nu)(\varphi)$ est $\geqslant 0$.

Enfin (IV,8;13) montre que, si μ et ν sont de normes finies, il en est de même de $\mu\otimes\nu$, et

$$(\text{IV}, 8\,; 19^{bis})\qquad \|\mu\otimes\nu\| \leqslant \|\mu\|\ \|\nu\|\,.$$

Support de $\mu\otimes\nu$

Théorème 76 - Si ~~A est le support de μ, et B le support de ν, le support~~ de $\mu\otimes\nu$ est l'ensemble produit A x B .

Démonstration 1°/ Dans l'ensemble ouvert $\complement A \times Y$. le produit $\mu\otimes\nu$ est nul, puisque μ est nulle dans $\complement A$. De même $\mu\otimes 3$ est nulle dans $X \times \complement B$, puisque 3 est nulle dans $\complement B$. Donc $\mu\otimes 3$ est nulle dans la réunion de ces deux ouverts (théorème 13). Or cette réunion est exactement le complémentaire de A × B ; cela prouve que $\mu\otimes\nu$ a nécessairement son support dans 4 × B .

2°/ Soit maintenant (a,b) un point quelconque de A × B , et soit $\mathscr{A}$ (resp. $\mathscr{B}$) un voisinage ouvert de a dans X (resp. de b dans Y). D'après la caractérisation du support (théorème 14) il existe une fonction u de $\mathscr{C}(X)$, de support dans $\mathscr{A}$, telle que $\mu(u) \neq 0$. De même, il existe une fonction v de $\mathscr{C}(Y)$, de support dans $\mathscr{B}$, telle que $\nu(v) \neq 0$. si alors nous prenons $\varphi(x,y) = u(x)\,v(y)$, on a nécessairement $(\mu\otimes V)((p) = \mu(u)\,\nu(v) \neq 0$. Cela prouve qu'il

* d'où le nom de produit tensoriel

n'existe aucun voisinage ouvert de (a, b) dans lequel $\mu \otimes \nu$ soit nulle, et que, par conséquent, (a, b) appartient au support de $\mu \otimes \nu$. Ce support est donc bien $A \times B$.

Calcul d'une intégrale double par 2 intégrations simples successives

Dans le théorème 75, nous avons donné un mode de calcul des intégrales doubles par deux intégrations simples successives; mais ce théorème n'était a priori valable que pour l'intégrale d'une fonction φ appartenant à $\mathcal{C}(X \times Y)$. Il existe un théorème bien plus général lorsque μ et ν sont ≥ 0.

Théorème 77 (Théorème de Fubini-Lebesgue) - Soient X et Y des espaces localement compacts dénombrables à l'infini. Soient μ et ν des mesures ≥ 0 sur X et Y respectivement. Soit $\vec{F}$ un espace de Banach. Soit $\vec{f}$ une fonction, définie sur $X \times Y$, à valeurs dans $\vec{F}$, et intégrable (resp. mesurable) pour $\mu \otimes \nu$. Alors, si l'on fixe x dans X , la fonction partielle $\vec{f}_x : y \longrightarrow \vec{f}(x,y)$, n'est pas toujours intégrable (resp. mesurable) par rapport à ν, mais elle l'est au moins pour μ-presque toutes les valeurs de x dans X . On peut ainsi définir *

$$\vec{g}(x) = \int \vec{f}(x,y)\, d\nu(y) \in \vec{F} , \tag{IV,8;20}$$

pour μ-presque toutes les valeurs de x .

La fonction $\vec{g} : x \longrightarrow \vec{g}(x)$, définie μ-presque partout sur X , à valeurs dans $\vec{F}$, est alors μ-intégrable, et l'on a la formule

$$\iint \vec{f}(x,y)\, d\mu(x)\, d\nu(y) = \int \vec{g}(x)\, d\mu(x) = \int \left[\int \vec{f}(x,y)\, d\nu(y) \right] d\mu(x) \tag{IV,8;21}$$

On l'écrit généralement sous la forme

$$\iint \vec{f}(x,y)\, d\mu(x)\, d\nu(y) = \int d\mu(x) \int \vec{f}(x,y)\, d\nu(y). \tag{IV,8;22}$$

* Nous supposons maintenant $\vec{f}$ intégrable, la suite ne se rapporte plus au cas écrit "resp. mesurable".

Cette notation est évidemment un peu étrange a priori. Pour éviter une parenthèse ou un crochet, on fait comme si $\int \vec{g}(x)\, d\mu(x)$ s'écrivait $\int d\mu(x)\, \vec{g}(x)$, ce qui alors justifie l'écriture (IV,8;22) par suppression des crochets. Il est bon de se familiariser avec cette notation très courante.

Démonstration

Lemme 1 - Soit f une fonction $\geqslant 0$ (à valeurs finies ou infinies), semi-continue inférieurement sur $X \times Y$; on a la formule

(IV,8;22 bis) $$\int^* d\mu(x) \int^* f(x,y)\, d\nu(y) = \iint^* f(x,y)\, d\mu(x)\, d\nu(y)$$

Considérons l'ensemble $\mathcal{C}'_+(X \times Y ; f)$ de toutes les fonctions $\geqslant 0$ de $\mathcal{C}(X \times Y)$ qui minorent f. Il vérifie la propriété de croissance donnée au théorème 39 quarto; d'autre part l'enveloppe supérieure de ces fonctions est exactement f [en effet, elle est trivialement $\leqslant f$; mais, si $a \in X$, il existe un voisinage ouvert $\mathcal{V}$ de a tel que $f(\mathcal{V}) \geqslant (f(a)-\varepsilon)^+ = \mathrm{Max}(f(a)-\varepsilon, 0)$. Il existe ensuite un voisinage compact $\mathcal{W}$ de a dans $\mathcal{V}$. Il existe alors, d'après le corollaire 1 du théorème 11, une fonction continue Ψ, $0 \leqslant \Psi \leqslant 1$, à support dans $\mathcal{V}$, égale à 1 dans $\mathcal{W}$; on a donc $0 \leqslant (f(a)-\varepsilon)^+ \Psi \leqslant f$ (car c'est vrai dans $\mathcal{V}$, mais les deux sont nulles en dehors de $\mathcal{V}$); mais cette fonction prend en a la valeur $(f(a)-\varepsilon)^+$. Donc la borne supérieure des $\varphi(a)$, pour toutes les φ de $\mathcal{C}_+(X \times Y)$ majorées par f, est bien $f(a)$]. On peut appliquer à $\mathcal{C}_+(X \times Y ; f)$ le théorème 39 quarto; cela redonne d'ailleurs le théorème 39 bis, à savoir

$$\iint^* f(x,y)\, d\mu(x)\, d\nu(y) = \sup_{\varphi \in \mathcal{C}_+(X \times Y; f)} (\mu \otimes \nu)(\varphi) .$$

Mais en outre, considérons l'ensemble des fonctions; partielles correspondantes $\varphi_x : y \longrightarrow \varphi(x,y)$; c'est un ensemble de fonctions continues $\geqslant 0$ vérifiant encore la condition de croissance du théorème 50 quarto, et son enveloppe supérieure est la fonction $f_x : y \longrightarrow f(x,y)$. Ce théorème donne donc

$$\int^* f(x,y)\, d\nu(y) = \sup_{\varphi \in \mathcal{C}_+(X \times Y; f)} \int \varphi(x,y)\, d\nu(y)$$

Mais, pour toute φ, la fonction $J_\varphi : x \longrightarrow \int \varphi(x,y)\, d\nu(y)$ est continue $\geqslant 0$ à support compact sur X (théorème 75); l'enveloppe supérieure de ces fonctions J_φ c'est-a-dire

$F : x \longrightarrow \int^* f(x,y)\, d\nu(y)$ est donc une fonction $\geqslant 0$ semi-continue inférieurement. Mais l'ensemble des J_φ vérifie encore la condition de croissance du théorème 39 quarto, car J_φ et J_Ψ sont majorées par $J_{\varphi+\Psi}$, donc $\mathrm{Sup}\,(J_\varphi, J_\Psi) \leqslant J_{\varphi+\Psi}$. On peut donc appliquer une nouvelle fois ce théorème :

$$\int^* d\mu(x) \int^* f(x,y)\, d\nu(y) = \operatorname*{Sup}_{\varphi \in \mathcal{C}_+(X\times Y; f)} \int d\mu(x) \int \varphi(x,y)\, d\nu(y) = \operatorname*{Sup}_{\varphi \in \mathcal{C}_f(X\times Y; f)} (\mu\otimes\nu)(\varphi)$$

$$(\text{théorème } 75) = \iint^* f(x,y)\, d\mu(x)\, d\nu(y)$$

ce qui démontre le lemme 1.

Lemme 2 - Soit $f \geqslant 0$ quelconque sur $X \times Y$. On a

$$\text{(IV,8;23)} \qquad \int^* d\mu(x) \int^* f(x,y)\, d\nu(y) \leqslant \iint^* f(x,y)\, d\mu(x)\, d\nu(y)^{*}$$

En effet, soit $g \geqslant f$ semi-continue inferieurement. On a, d'après le lemme 1 :

$$\int^* d\mu(x) \int^* f(x,y)\, dy \leqslant \int^* d\mu(x) \int^* g(x,y)\, d\nu(y) = \iint^* g(x,y)\, d\mu(x)\, d\nu(y)$$

Mais, lorsque g varie, la borne inférieure du dernier membre est, d'après le théorème **50** ter, $\iint^* f(x,y)\, d\mu(x)\, d\nu(y)$, d'où (IV,8;23).

Démontrons maintenant le théorème. Soit $\vec{f}$ une fonction sur $X \times Y$, à valeurs dans $\vec{F}$, $(\mu\otimes\nu)$ - intégrable. Elle a une suite d'approximation formée de fonctions $\vec{f}_n$ continues décomposables à support compact, de la forme $\sum_i g_{i,n}\, \varphi_{i,n}$, $\varphi_{i,n} \in \mathcal{C}(X\times Y)$. Pour une telle fonction, le théorème de Fubini est évident, il se ramène à celui d'une fonction de $\mathcal{C}(X\times Y)$, et c'est alors la définition même du produit tensoriel par le théorème **75.**

* Le lemme 1 pourrait laisser croire que les fonctions semi-continues inférieurement sont très privilégiées parmi les fonctions $\geqslant 0$, en ce sens qu'on a égalité au lieu d'inégalité. Le théorème 78 montrera qu'on a égalité pour toutes les fonctions $(\mu\otimes\nu)$ - mesurables $\geqslant 0$ sur $X \times Y$, en particulier pour toutes les fonctions boréliennes; et toute fonction semi-continue est borélienne.

Les quantités $\int^* d\mu(x) \int^* \| \vec{f} - \vec{f_n} \| d\nu(y) \leq \iint^* \| \vec{f} - \vec{f_n} \| d\mu(x)\, d\nu(y)$ (lemme 2) convergent vers 0 . Il résulte alors du théorème 38 qu'on peut extraire une suite partielle que nous continuerons a appeler $\vec{f_n}$, pour laquelle $\int^* \| \vec{f} - \vec{f_n} \| \, d\nu(y)$ converge vers 0 pour $d\mu$- presque tout x de X . Il existe donc un ensemble A de X , de μ-mesure nulle, tel que, pour $x \notin A$, cette convergence ait lieu. Mais la fonction $(\vec{f_n})_x : y \longrightarrow \vec{f_n}(x, y)$, continue à support compact décomposable, est ν- intégrable; donc, pour tout $x \notin A$, $\vec{f}_x$ est ν- intégrable, et les $(\vec{f_n})_x$ en forment une suite d'approximation : $\vec{g}(x) = \int \vec{f}(x,y)\, d\nu(y) = \lim_{n \to \infty} \int \vec{f_n}(x,y)\, d\nu(y) = \lim_{n \to \infty} \vec{g}_n(x)$.

Chaque fonction $\vec{g_n}$ est continue décomposable à support compact (théorème 75). donc μ- intégrable, et $\int \vec{g}_n \, d\mu = \iint \vec{f_n} \, d(\mu \otimes \nu)$; $\vec{g}$ est définie μ- presque partout, et l'on a

$$\int^* \| \vec{g} - \vec{g}_n \| \, d\mu \leq \int^* d\mu(x) \int^* \| \vec{f} - \vec{f_n} \| d\nu(y) \leq \iint^* \| \vec{f} - \vec{f_n} \| d(\mu \otimes \nu)$$

tend vers 0 .

Donc $\vec{g}$ est μ-intégrable, et les $\vec{g_n}$ en forment une suite d'approximation :

$$\int d\mu(x) \int \vec{f}(x,y)\, d\nu(y) = \int \vec{g}\, d\mu = \lim_{n \to \infty} \int \vec{g}_n \, d\mu = \lim_{n \to \infty} \iint \vec{f_n}\, d(\mu \otimes \nu) = \iint \vec{f}\, d(\mu \otimes \nu),$$

Ce qui démontre le théorème.

Remarques 1°/ Naturellement, on pourrait procéder en sens inverse et écrire la formule

(IV,8;24) $$\iint \vec{f}(x,y)\, d\mu(x)\, d\nu(y) = \int d\nu(y) \int \vec{f}(x,y)\, d\mu(x).$$

Il en résulte en particulier que, si l'on sait à l'avance que $\vec{f}$ est intégrable pour $\mu \otimes \nu$, on a la formule d'interversion des intégrations

(IV,8;25) $$\int d\mu(x) \int \vec{f}(x,y)\, d\nu(y) = \int d\nu(y) \int \vec{f}(x,y)\, d\mu(x)$$

2°/ Si $\vec{f}$ n'est pas intégrable pour $\mu \otimes \nu$, il peut arriver que l'un des deux membres de (IV,8;25) ait un sens, sans que l'autre ait un sens. Il peut même arriver que chacun ait un sens et que leurs valeurs soient distinctes.

Considérons par exemple, sur le carré $[0,1] \times [0,1]$ de $\mathbb{R}^2$, la mesure produit $dx \otimes dy$, et la fonction réelle f définie comme suit :

(IV,8;26) $$f(x,y) = \frac{x^2 - y^2}{(x^2+y^2)^2}, \text{ pour } (x,y) \neq (0,0).$$

Nous ne prenons pas la peine de définir $f(0,0)$, car cela reviendra, pour x fixé, à ne pas la définir éventuellement pour une valeur de y ou au contraire, pour y fixé, à ne pas la définir éventuellement pour une valeur de x, et de toute façon, un point est un ensemb de mesure nulle pour dx ou dy.

Alors nous allons voir que l'expression

$$(IV,8;27)\qquad \int_{[0,1]} dx \int_{[0,1]} \frac{x^2 - y^2}{(x^2 + y^2)^2}\, dy$$

a un sens, et qu'elle est égale à $\frac{\pi}{4}$.

En effet, pour x fixé $\neq 0$, une primitive en y de la fonction continue $y \longrightarrow \frac{x^2 - y^2}{(x^2 + y^2)^2}$ est la fonction $y \longrightarrow \frac{y}{x^2 + y^2}$. On a donc, pour $x \neq 0$:

$$(IV,8;28)\qquad \int_{[0,1]} \frac{x^2 - y^2}{(x^2 + y^2)^2}\, dy = \frac{1}{x^2 + 1}$$

On n'a pas besoin, pour appliquer le processus (IV,8;27), de calculer l'intégrale pour $x = 0$, car $\{0\}$ est de mesure nulle pour dx.

Alors

$$(IV,8;29)\qquad \int_{[0,1]} dx \int_{[0,1]} \frac{x^2 - y^2}{(x^2 + y^2)^2}\, dy = \int_0^1 \frac{dx}{1 + x^2} = \frac{\pi}{4} .$$

Si maintenant nous essayons de calculer en sens inverse l'intégrale

$$(IV,8;30)\qquad \int_{[0,1]} dy \int_{[0,1]} \frac{x^2 - y^2}{(x^2 + y^2)^2}\, dx \ ,$$

nous voyons qu'elle peut s'écrire, en permutant les variables muettes x et y ,

$$(IV,8;31)\qquad \int_{[0,1]} dx \int_{[0,1]} \frac{y^2 - x^2}{(y^2 + x^2)^2}\, dy \ ;$$

elle est donc opposée à (IV,8;27) et vaut $-\frac{\pi}{4}$. Ainsi (IV,8;27) et (IV,8;30) ont toutes les deux un sens mais n'ont pas la même valeur. Cela prouve que le théorème de Fubini n'était pas applicable, donc nécessairement la fonction f n'était pas intégrable pour la mesure produit $dx \otimes dy$; ce résultat négatif en constitue la preuve. A l'aide de méthodes qui seront développées plus loin, on peut le voir directement; on pourra le montrer en faisant d'abord un changement de variables, passant aux coordonnées polaires, $x = r\cos\varphi$, $y = r\sin\varphi$ (formule (IV,9;92)), puis en appliquant le corollaire 4 du théorème 108, avec $\alpha = 2$.

Z

3°/ <u>On ne pouvait pas espérer que $\vec{f}_x$ serait intégrable en y, pour toutes les valeurs de x.</u>

Si en effet nous reprenons l'exemple des mesures dx et dy sur $\mathbb{R}$, on voit qu'on peut modifier $\vec{f}$ sur une parallèle $x = a$ à l'axe des y, sans modifier son intégrabilité par rapport à $dx \otimes dy$ parce que la droite $x = a$ est un ensemble de mesure nulle dans $\mathbb{R}^2$.

Cela permet de prendre, pour fonction $y \to f(a, y)$, <u>n'importe quelle fonction de y</u>, donc n'ayant aucune raison d'être intégrable en y pour dy. <u>Mais cette circonstance n'a évidemment aucune importance pour ce qui nous occupe;</u> du moment que nous savons que la fonction% est définie pour μ-presque toutes les valeurs de x, et intégrable par rapport à μ, c'est entièrement suffisant pour calculer l'intégrale double par deux intégrations simples successives. Il y a même là une commodité dont il faut savoir profiter : si, pour certaines valeurs exceptionnelles de x formant un ensemble de p-mesure nulle, l'intégrale $\int \vec{f}(x,y)\, d\nu(y)$, ou bien n'existe pas, ou bien est difficile à calculer, on n'a aucun besoin de s'en occuper.

4°/ Le Théorème de Fubini est à comparer au théorème de sommation par paquets dans les séries commutativement convergentes (théorème 59 du chapitre 2). Dans ce théorème également, <u>nous avons dû supposer d'abord que la série était commutativement convergente,</u> pour ensuite pouvoir la sommer par paquets. Or nous avons vu à ce moment là qu'il existait cependant un cas où l'on n'avait pas besoin de savoir à l'avance la série commutativement convergente : à savoir le cas où les termes de la série étaient réels ≥ 0. Il en est de même ici, comme le montreront le théorème suivant et son corollaire 1.

<u>Théorème 78 (Fubini-Fatou).</u> Soient X, Y, des espaces localement compacts, dénombrables à l'infini. <u>Soit f une fonction ≥ 0 (à valeurs finies ou non) sur $X \times Y$, $(\mu \otimes \nu)$-mesurable. Alors on a l'égalité</u>

$$(IV,8;33\,bis)\qquad 0 \leq \int^* d\mu(x) \int^* f(x,y)\, d\nu(y) = \iint^* f(x,y)\, d\mu(x)\, d\nu(y) \leq +\infty .$$

<u>Démonstration</u> - Puisque f est $(\mu \otimes \nu)$-mesurable, elle est limite d'une suite croissante de fonctions $f_n \geq 0$ mesurables, bornées, à supports compacts (Par exemple des fonctions $f_{M,K}$ comme page 474). Pour chaque f_n on a, d'après Fubini-Lebesgue (théorème 77) :

$$\int d\mu(x)\int f_n(x,y)\,d\nu(y)=\iint f_n(x,y)\,d\mu(x)\,d\nu(y)\ ,$$

la deuxième intégrale écrite au premier membre, $\int f_n(x,y)d\nu(y)$, ayant un sens pour p-presque toutes les valeurs de x. On a **donc la même égalité avec** partout des $\int^*$ au lieu de $\int$, toutes les $\int^*$ ayant alors un sens (a valeurs finies ou non). Il suffit alors de faire tendre n vers l'infini, et d'appliquer le théorème 36 de Fatou, deux fois à gauche et une fois a droite.

Remarque - On peut dire un peu plus : Soit A_n l'ensemble. de μ- mesure nulle, des x pour lesquels $(f_n)_x$ n'est pas ν- intégrable. Soit $A = \bigcup_n A_n$, encore de μ- mesure nulle. Pour $x \notin A$, toutes les $(f_n)_x$ sont ν- mesurables, donc aussi leur limite f_x ; ainsi la fonction partielle f_x est ν- mesurable pour μ- presque tous les x (voir généralisation au corollaire 4).

En outre, si $g_n(x)=\int f_n(x,y)\,d\nu(y)$, $g(x)=\int^* f(x,y)\,d\nu(y)$, g_n est définie sur $\complement A_n$ et μ- mesurable; donc g est définie sur $\complement A$ et μ- mesurable.

Mais on a vu que, pour des fonctions mesurables $\geqslant 0$, on pouvait écrire $\int$ au lieu d' $\int^*$, **même** si elle était infinie (remarque 2 après corollaire 3 du théorème 39). Dans ce cadre-la, on peut donc écrire, si $f \geqslant 0$ est $(\mu \otimes \nu)$- mesurable :

$$0 \leqslant \int d\mu(x)\int f(x,y)\,d\nu(y)=\iint f(x,y)\,d\mu(x)\,d\nu(y) \leqslant +\infty\ ; \tag{IV,8;33}$$

la 2ème intégrale du 1er membre a un sens, pour μ- presque toutes **les valeurs** de x, comme intégrale (finie ou infinie) de la fonction f_x, ν- mesurable $\geqslant 0$; la 1ère intégrale du 1er membre (finie ou infinie) est celle de la fonction g définie k-presque partout mais k-mesurable $\geqslant 0$; l'intégrale double du 2ème membre est celle de la fonction f, $(\mu\otimes\nu)$- mesurable $\geqslant 0$.

<u>Corollaire 1 - si X et Y sont des espaces localement compacts dénombrables à l'infini. si μ et ν sont des mesures de Radon $\geqslant 0$ sur X et Y respectivement. si f est une fonction réelle $\geqslant 0$, mesurable pour $\mu \otimes \nu$, si, pour μ- presque toutes les valeurs de x, la fonction partielle $f_x : y \longmapsto f(x,y)$, est intégrable pour ν, si la fonction g, définie pour μ- presque toutes les valeurs de x par</u>

(IV,8;32) $$g(x) = \int f(x,y)\,d\nu(y)\ ,$$

est intégrable pour μ , alors la fonction f est intégrable pour $\mu \otimes \nu$, et l'on a l'égalité (IV,8;22).

La démonstration est immédiate. Les hypothèses signifient que le 1er membre de (IV,8;33 bis) est fini, donc aussi le 2ème; alors f, étant $(\mu \otimes \nu)$- mesurable et d'intégrale superieure finie, est intégrable d'après le théorème 39.

Corollaire 2-Si $\vec{f}$ est une fonction définie sur $X \times Y$, à valeurs dans un espace de Banach $\vec{F}$ et $(\mu \otimes \nu)$ mesurable, et si l'une quelconque des trois intégrale

(IV,8;34) $$\iint^* \|\vec{f}(x,y)\|\,d\mu(x)\,d\nu(y)\ ,\quad \int^* d\mu(x)\int^* \|\vec{f}(x,y)\|\,d\nu(y)\ ,$$

$$\int^* d\nu(y)\int^* \|\vec{f}(x,y)\|\,d\mu(x)\ ,$$

a une valeur finie, alors nécessairement la fonction $\vec{f}$ est intégrable pour $\mu \otimes \nu$, et l'on a les formules (IV,8;22) et (IV,8;24).

En effet, la fonction $\|\vec{f}\|$ étant mesurable, on peut appliquer le théorème de Fubini-Fatou sous la forme précédente, et le fait que l'une des trois intégrales considérées soit finie signifie que les trois intégrales considérées sont finies, donc en particulier que $\iint^* \|\vec{f}(x,y)\|\,d(\mu \otimes \nu)$ est finie; il résulte alors du théorème 39, $\vec{f}$ étant supposée $(\mu \otimes \nu)$-mesurable, que $\vec{f}$ est bien $(\mu \otimes \nu)$-intégrable. On se trouve alors dans les conditions d'application du théorème de Fubini-Lebesgue.

Remarque - Naturellement il est possible d'appliquer les théorèmes de Fubini pour une intégrale du type $\iint_A \vec{f}(x,y)\,d\mu(x)\,d\nu(y)$, où A est un ensemble $(\mu \otimes \nu)$ - mesurable de $X \times Y$, pourvu que $\vec{f}$ soit intégrable sur A , ou soit $(\mu \otimes \nu)$ - mesurable et $\geqq 0$. Comme cela revient à calculer $\iint \tilde{\vec{f}}(x,y)\,d\mu(x)\,d\nu(y)$, où $\tilde{f}$ est la fonction égale à $\vec{f}$ sur A , prolongée par $\vec{0}$ en dehors de A , il n'y a pas là un théorème nouveau, Pour $a \in X$ appelons $A(a)$ la "section de A par la verticale d'abscisse a", c'est-à-dire l'ensemble des points y de Y tels que $(a,y) \in A$. Alors on aura

(IV,8;35) $$\iint_A \vec{f}(x,y)\,d\mu(x)\,d\nu(y) = \int_X d\mu(x) \int_{A(x)} \vec{f}(x,y)\,d\nu(y).$$

En particulier, si $f = 1$ sur A :

(IV,8;35 bis) $$(\mu \otimes \nu)(A) = \int \nu(A(x))\,d\mu(x),$$

où $\nu(A(x))$ est la ν-mesure de la section $A(x)$; elle est définie pour μ-presque tout x (a valeur fini ou infinie).

Corollaire 3 - Soit A une partie $(\mu \otimes \nu)$-mesurable de $X \times Y$. Pour qu'elle soit de $(\mu \otimes \nu)$-mesure nulle, il faut et il suffit que, pour μ-presque tout x, sa section $A(x)$ soit de ν-mesure nulle.

En effet, $\int \nu(A(x))\,d\mu(x)$ est nulle, d'après le théorème 26, si et seulement si la fonction qu'on intègre est μ-presque partout nulle.

Remarque - Il est essentiel dans cet énoncé de supposer A $(\mu \otimes \nu)$-mesurable. On peut (avec l'axiome de Zermelo) prouver l'existence d'une partie A de $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ qui coupe toute droite $x =$ constante suivant un seul point, mais qui n'est pas de $(dx \otimes dy)$-mesure nulle (parce que non $(dx \otimes dy)$-mesurable).

Corollaire 4 - Soit f une fonction à valeurs dans un espace métrisable séparable F. Si f est $(\mu \otimes \nu)$-mesurable, alors, pour μ-presque toutes les valeurs de $x \in X$, la fonction partielle $f_x : y \longrightarrow f(x,y)$, est ν-mesurable.

Démonstration - Soit **A** une partie borélienne de $X \times Y$. Alors, pour tout x de **X**, la section **A** (x) est borélienne. En effet les parties de $X \times Y$ dont la section par x donné est borélienne forment manifestement une tribu, qui contient tous les ouverts de $X \times Y$, donc tous les boréliens.

Alors, si f est une fonction borélienne sur $X \times Y$, toutes ses fonctions partielles f_x sont boréliennes : car, si $\mathcal{O}$ est un ouvert de **F**, $f_x^{-1}(\mathcal{O}) = (f^{-1}(\mathcal{O}))(x)$ est borélien comme section par x de $f^{-1}(\mathcal{O})$ borélien.

Si maintenant f est $(\mu \otimes \nu)$-mesurable, il existe une fonction g borélienne, égale à f sauf aux points d'un ensemble A de $(\mu \otimes \nu)$-mesure nulle (remarque après le théorème 23 bis). Pour μ-presque tous les x, la section $A(x)$ est de ν-mesure nulle (corollaire 3) et f_x coïncide avec g_x borélienne sauf sur $A(x)$ donc est ν-mesurable.

Cas où la fonction à intégrer est le produit d'une fonction de x par une fonction de y

Théorème 79 - Soient μ et ν des mesures ≥ 0 sur X et Y, espaces localement compacts dénombrables à l'infini. Soit f une fonction définie sur X à valeurs dans un espace de Banach $\vec{F}$, $\vec{g}$ une fonction définie sur Y, à valeur dans un espace de Banach $\vec{G}$, et soit B une application bilinéaire **continue** de $\vec{F} \times \vec{G}$ dans un espace de Banach $\vec{H}$.

Alors, si $\vec{f}$ est μ-intégrable (resp. μ-mesurable), et si $\vec{g}$ est ν-intégrable (resp. ν-mesurable), la' fonction $B(\vec{f},\vec{g})$: $(x,y) \longrightarrow B(\vec{f}(x),\vec{g}(y))$ est $(\mu \otimes \nu)$-intégrable (resp. $(\mu \otimes \nu)$-mesurable *), et l'on a la formule

(IV,8;36)
$$\iint B(\vec{f},\vec{g})\, d(\mu \otimes \nu) = B\left(\int \vec{f}\, d\mu \,,\, \int \vec{g}\, d\nu\right)$$

Inversement si $\vec{F}$, $\vec{G}$, $\vec{H}$, sont le corps des scalaires, si B est le produit ordinaire, si $f\,g$ est $\mu \otimes \nu$-intégrable, et si aucune des deux fonctions f, g, n'est presque partout nulle, **alors** f est μ-intégrable et g est ν-intégrable.

Avant de donner la démonstration, remarquons que, si $\vec{F} = \vec{G} = \vec{H}$ est le corps des scalaires, et si B est le produit ordinaire, (IV,8;36) s'écrit

(IV,8;37)
$$\iint f(x)\, g(y)\, d\mu(x)\, d\nu(y) = \left(\int f(x)\, d\mu(x)\right)\left(\int g(y)\, d\nu(y)\right)$$

Cela ne résulte pas immédiatement de la définition (IV,8;4) de $\mu \otimes \nu$, car et g ne sont pas continues a support compact.

Démonstration. Démontrons d'abord la réciproque.
Supposons donc $f\,g$ intégrable pour $\mu \otimes$ 3 . Alors le théorème de Fublni-Lebssgue nous dit que, pour k-presque toutes les valeurs de x, la fonction Y $\longrightarrow f(x)\, g(y)$ est ν-intégrable. Comme par ailleurs nous avons supposé que f n'était pas presque partout nulle, il existe au moins un x pour lequel $y \longrightarrow f(x)\, g(y)$ est 3 -intégrable et $f(x) \neq 0$;on en déduit bien que la fonction g est ν-intégrable. Le même raisonnement montre que f est μ-intégrable ** .

* Contrairement à ce qui a lieu dans les **théorèmes** de Fubini, on part de mesurabilités par rapport à μ et à 3 pour en déduire une mesurabilité par rapport à $\mu \otimes \nu$.

** En dehors du cas particulier étudié, ce résultat ne subsiste en général pas. Il est, par exemple, évidemment faux pour $B \equiv 0$. Demême, si $f \equiv$ o , $f(x)g(y)$ est $\equiv 0$, donc $(\mu \otimes \nu)$-intégrable, sans qu'on puisse en déduire que g est ν-intégrable.

Pour démontrer le théorème direct, nous aurons besoin d'un lemme :

Lemme - Si A est une partie de X , de mesure nulle pour μ , $A \times Y$ est de mesure nulle pour $(\mu \otimes \nu)$.

Ce lemme est assez remarquable, car, s'il est vrai que $\mu(A) = 0$, $\nu(Y)$ est en général infini. Nous avions déjà indiqué que, dans le plan $\mathbb{R}^2$, pour la mesure $dx\,dy$, une parallèle à l'un des axes était de mesure nulle.

Démonstration du lemme soit $\varepsilon > 0$. Soit K un compact de Y . Sa mesure est finie, et il existe un ouvert $\mathcal{O}' \supset K$ de Y , tel que $\nu(\mathcal{O}') \leqslant \nu(K) + 1$. Il existe alors un ouvert $\mathcal{O} \supset A$ de X tel que $\mu(\mathcal{O}) \leqslant \frac{\varepsilon}{\nu(K)+1}$. Alors $\mathcal{O} \times \mathcal{O}'$ est ouvert dans $X \times Y$, donc certainement $(\mu \otimes \nu)$ -mesurable, et nous pouvons appliquer à sa fonction caractéristique la formule (IV,8;33); elle donne

$$(\mu \otimes \nu)\,(\mathcal{O} \times \mathcal{O}') = \mu(\mathcal{O})\,\nu(\mathcal{O}') \leqslant \varepsilon \qquad \text{(IV,8;38)}$$

Donc $(\mu \otimes \nu)^*(A \times K) \leqslant \varepsilon$; comme ε est arbitrairement petit, on a $(\mu \otimes \nu)(A \times K) = 0$. Mais Y est la réunion d'une suite croissante de compacts K_n , donc $A \times Y$, réunion de la suite des $A \times K_n$, de $(\mu \otimes \nu)$ mesure nulle, est lui aussi de $(\mu \otimes \nu)$-mesure nulle.

Fin de la démonstration du théorème direct

Nous allons d abord démontrer que la fonction $B(\vec{f}, \vec{g})$ est $(\mu \otimes \nu)$-mesurable.

Comme $\vec{f}$ (resp. $\vec{g}$) est mesurable, il existe, d'après le théorème 32, une suite d'approximation de $\vec{f}$ (resp. $\vec{g}$) formée de fonctions continues dont on peut extraire (par le théorème 38) une suite partielle $\vec{f}_n$ (resp. $\vec{g}_n$), μ (resp. ν) - presque partout convergente, c'est-à-dire convergeant vers $\vec{f}$ (resp. $\vec{g}$) sur un ensemble A (resp. B) dont le complémentaire est de μ-mesure (resp. ν-mesure) nulle. Alors $B(\vec{f}_n, \vec{g}_n)$ est continue; et elle converge vers $B(\vec{f}, \vec{g})$ sur $A \times B$, dont le complémentaire $\complement(A \times B) = (\complement A \times Y) \cup (X \times \complement B)$ est de mesure nulle d'après le lemme. Donc $B(\vec{f}, \vec{g})$ est bien mesurable, d'après le théorème 23.

Alors pour démontrer qu'elle est intégrable, il suffit de montrer que **l'intégrale supérieure** de sa norme est finie., Or celle-ci admet la majoration

$$\iint^{*} \|B(\vec{f},\vec{g})\| \, d(\mu \otimes \nu) \leq \|B\| \iint^{*} \|\vec{f}(x)\| \; \|\vec{g}(y)\| \, d\mu(x) \, d\nu(y) \tag{IV,8;39}$$

Mais, de la même manière, **la fonction** $(x,y) \longrightarrow \|\vec{f}(x)\| \, \|\vec{g}(y)\|$ **est** $(\mu \otimes \nu)$ -mesurable, et, comme elle est ≥ 0 , **on** peut certainement lui **appliquer** le théorème de Fubini - Fatou. Celui-ci donne (avec $'^{*}\!\int = '\!\int$) :

$$\iint \|\vec{f}(x)\| \; \|\vec{g}(y)\| \, d\mu(x) \, d\nu(y) \tag{IV,8;40}$$

$$= \int d\mu(x) \int \|\vec{f}(x)\| \; \|\vec{g}(y)\| \, d\nu(y)$$

Comme $\vec{g}$ est supposée **ν-intégrable,** la dernière Intégrale est finie, et vaut $\|\vec{f}(x)\| \int \|\vec{g}(y)\| \, d\nu(y)$, donc l'intégrale supérieure cherchée vaut $\left(\int \|\vec{f}\| d\mu\right)\left(\int \|\vec{g}\| d\nu\right) < +\infty$, et $B(\vec{f},\vec{g})$ est bien $(\mu \otimes \nu)$ **-intégrable.**

On peut donc lui appliquer le théorème de Fubini-Lebesgue, et écrire

$$\iint B(\vec{f},\vec{g}) \; d(\mu \otimes \nu) \tag{IV,8;41}$$

$$= \int d\mu(x) \int B(\vec{f}(x), \vec{g}(y)) \, d\nu(y) \, .$$

Mais la fonction $\vec{g}$ est 3 - **intégrable,** et, **pour** x fixé, l'application $\vec{q} \longrightarrow B(\vec{f}(x), \vec{q})$ de $\vec{G}$ dans $\vec{H}$ **est** linéaire continue; donc le théorème 45 (**permutabilité** de l'intégrale et d'une application linéaire continue) donne

$$\int B(\vec{f}(x), \vec{g}(y)) \, d\nu(y) = B\left(\vec{f}(x), \int \vec{g} \, d\nu\right) \tag{IV,8;42}$$

, d'où

$$\iint B(\vec{f},\vec{g}) \, d(\mu \otimes \nu) = \int B\left(\vec{f}(x), \int \vec{g} \, d\nu\right) d\mu(x) \tag{IV,8;43}$$

Mais $\vec{f}$ est μ-tégrable. ,et $\vec{f} \longrightarrow B(\vec{f}, \int \vec{g}\, d\nu)$ est uns application linéaire continue de $\vec{F}$ dans $\vec{H}$, donc le même théorème 45 donne

(IV,8;44) $$\int B\left(\vec{f}(x), \int \vec{g}\, d\nu\right) d\mu(x)$$

$$= B\left(\int \vec{f}\, d\mu, \int \vec{g}\, d\nu\right)$$

et le théorème est démontré.

Remarque - Si $F = \vec{G} = \vec{H}$ sont le corps des scalaires. et si B est le produit ordinaire! la fin de la démonstra - tion est plus simple, une fois démontré que $(x,y) \to f(x)g(y)$ est intégrable, puisqu'on a immédiatement :

(IV,8;44[bis]) $$\iint f(x)g(y)\,d\mu(x)\,d\nu(y) = \int d\mu(x) \int f(x)\,g(y)\,d\nu(y)$$

$$= \int f(x)\,d\mu(x) . \int g(y)\,d\nu(y) .$$

Corollaire 1 - Soient A et B des parties de X et Y respectivement. Si elles sont μ -mesurable et ν -mesurable respectivement, de mesures finies ou infinies, alors $A \times B$ est $(\mu \otimes \nu)$- mesurable, et sa mesure est le produit des mesures.

(IV,8;45) $$(\mu \otimes \nu)(A \times B) = \mu(A)\,\nu(B) \leqslant +\infty ,$$

le produit étant considéré comme nul si l'un de ses fac - teurs est nul, même si l'autre vaut $+\infty$.

En particulier $(\mu \otimes \nu)(X \times Y) = \mu(X)\,v\,(Y)$.

Réciproquement, si $A \times B$ est mesurable par rapport à $(\mu \otimes \nu)$, et si aucun des deux ensembles n'est de mesure nulle, alors A est μ -mesurable, et B est 3 -mesurable.

Corollaire 2 - Soient μ et ν des mesures ≥ 0 sur X et Y, et soit p et 4 des fonctions complexes, respectivement sur X et Y, p localement μ-intégrable, q localement ν - intégrable. Alors la fonction $p \otimes q : (x,y) \to p(x)\,q(y)$ est localement $(\mu \otimes \nu)$ -intégrable, et la mesure $(p \otimes q)(\mu \otimes \nu)$ coïncide avec la mesure $p\mu \otimes q\nu$.

La locale $(\mu \otimes 3)$ intégrabilité résulte du théorème. Ensuite l'égalité du produit multiplicatif des produits tensoriels, $(p \otimes q)(\mu \otimes \nu)$, et du produit tensoriel des produits multiplicatifs, $p\mu \otimes q\nu$ résulte de ce que ces 2 mesures prennent la même valeur sur toute fonction φ de $\mathscr{C}(X \times Y)$ de la forme $u \otimes v$, $u \in \mathscr{C}(X)$, $v \in \mathscr{C}(Y)$, à savoir $\left(\int p u \, d\mu\right)\left(\int q v \, d\nu\right)$ (voir note (*) page 579).

Extension aux intégrales multiples quelconques

Soient par exemple, X, Y, Z trois espaces localement compacts, munis de trois mesures scalaires λ, μ, 3 .

Alors on peut, d'un seul coup, définir une mesure produit tensoriel $\lambda \otimes \mu \otimes \nu$ sur $X \times Y \times Z$, grâce a un théorème analogue à 73. On a cette fois, pour $u \in \mathscr{C}(X)$, $v \in \mathscr{C}(Y)$, $w \in \mathscr{C}(Z)$:

$$(\lambda \otimes \mu \otimes \nu)(u \otimes v \otimes w) = \lambda(u)\mu(v)\nu(w) \tag{IV,8;46}$$

Les intégrales relatives à $\lambda \otimes \mu \otimes 3$ sont des intégrales triples; on définit de même des intégrales multiples d'ordre fini quelconque.

D'autre part, il est possible de définir d'abord la mesure $\lambda \otimes \mu$ sur $X \times Y$, puis la mesure $(\lambda \otimes \mu) \otimes \nu$ sur $(X \times Y) \times Z$, qu'on peut identifier à $X \times Y \times Z$. Il est facile de voir, grâce aux formules (IV,8;46) et (IV,8; 4) que ces deux mesures sur $X \times Y \times Z$, $\lambda \otimes \mu \otimes \nu$ et $(\lambda \otimes \mu) \otimes 3$, coïncident : c'est <u>l'associativité du produit tensoriel.</u> Naturellement la méthode de calcul développée au théorème 75 est valable, et, lorsque λ, μ, ν, sont $\geqslant 0$, il existe des théorèmes de Fubini, généralisant les théorèmes 77 et 78. Boit à calculer, par exemple, l'intégrale

$$\iiint \vec{f}(x,y,z)\, d\lambda(x)\, d\mu(y)\, d\nu(z) \; , \tag{IV,8;47}$$

lorsque l'on sait d'avance que la fonction $\vec{f}$ est intégrable pour $\lambda \otimes \mu \otimes 3$, ou qu'elle est mesurable et $\geqslant 0$.

On pourra, par exemple, calculer d'abord une **intégrale** double, puis une intégrale simple suivant la formule :

(IV, 8;48) $$\int d\mu(y) \iint \vec{f}(x,y,z)\, d\lambda(x)\, d\nu(z)\ ,$$

l'intégrale double ayant un sens pour p-presque toutes les valeurs de y .

On pourra aussi calculer d'abord une **intégrale** simple puis une intégrale double suivant la formule

(IV, 8;49) $$\iint d\mu(y)\, d\nu(z) \int \vec{f}(x,y,z)\, d\lambda(x)\ ,$$

l'intégrale simple ayant un sens pour $(\mu \otimes \nu)$**-presque** toutes les valeurs de (y,z) .

On pourra encore calculer trois intégrales simples successives, par exemple suivant la formule

(IV, 8;50) $$\int d\mu(y) \int d\lambda(x) \int \vec{f}(x,y,z)\, d\nu(z)\ ,$$

l'intégrale en $d\nu(z)$ ayant un sens **pour**$(\lambda \otimes \mu)$ presque toutes les valeurs de (x,y) , **l'intégrale** suivante en $d\lambda(x)$ **pour** k-presque **toutes** les valeurs de y .

Tous les résultats obtenus seront les mêmes

Cette méthode a été couramment utilisée en mathématiques spéciales, pour calculer des volumes.

Supposons par exemple que X , Y , Z soient tous les trois la droite **réelle** $\mathbb{R}$**,** et que l'on ait

$$d\lambda = d\mu = du = \text{mesure de Lebesgue}$$

Alors la mesure produit $dx\ dy\ dz$ est la mesure des **volumes dans l'espace** $\mathbb{R}^3$ **.**

Soit alors A une partie mesurable **de,**$\mathbb{R}^3$. et **supposons que nous** cherchions **à** calculer son volume, **c'est-à-dire** l'intégrale $\int_A dx \otimes dy \otimes dz$. Alors, d'après la formule **(IV,8;35),** dans le cadre du **théorème** de **Fubini-Fatou,** cette intégrale pourra par **exemple** se calculer par

(IV, 8;51) $$\int S(z)\, dz \leq +\infty\ ,$$

où l'on désigne par $S(z)$ l'aire de la section $A(z)$ de cote z de l'ensemble A , c'est-à-dire précisément l'intégrale double

$$\iint_{A(z)} dx\, dy \leqslant +\infty \tag{IV,8;52}$$

Elle pourra aussi se calculer par le procédé

$$\iint \ell(x,y)\, dx\, dy \leqslant +\infty\ , \tag{IV,8;53}$$

où l'on désigne par $\ell(x,y)$ la "longueur" de la section de l'ensemble A par la droite parallèle à l'axe des z de coordonnées x, y , c'est-à-dire l'intégrale

$$\int_{A(x,y)} dz \leqslant +\infty\ . \tag{IV,8;54}$$

Si on considère l'espace $\mathbb{R}^n$ c'est le produit de-n espaces identiques à $\mathbb{R}$. Si on convient de noter $dx_1, dx_2, \ldots, dx_n$, les mesures canoniques (ou de Lebesgue) sur les espaces facteurs * , alors $dx_1 \otimes dx_2 \otimes \ldots \otimes dx_n$ est une mesure bien déterminée sur $\mathbb{R}^n$; on l'appelle mesure canonique ou mesure de Lebesgue sur $\mathbb{R}^n$, et on la note aussi dx ou $dx_1\, dx_2\, \ldots\, dx_n$.

Convergences vagues de produits tensoriels

Théorème 80 - Si une suite de mesures μ_n sur X converge vaguement vers μ pourn tendant vers l'infini, et si une suite de mesures 3, sur Y converge vaguement vers ν pour n tendant vers l'infini, alors les mesures $\mu_n \otimes \nu_n$ convergent vaguement vers $\mu \otimes \nu$.

Démonstration Soit φ un élément de $\mathcal{C}(X \times Y)$. Utilisons la méthode de calcul du théorème 75. Les quantités $\int \varphi(x,y)\, d\nu_n(y)$ convergent vers $\int \varphi(x,y)\, d\nu(y)$, et ceci pour tout x, d'après l'hypothèse de convergence vague.

* Notation incorrecte, voit note * page 429 . Tous les espaces facteurs sont les mêmes, c'est donc toujours la même mesure ! On aurait le droit de dire :

$1(x_1)dx_1,\ 1(x_2)dx_2, \ldots ,\ 1(x_n)$ dz, , ce qui est en effet toujours la même mesure !

En outre, d'après le corollaire 3 du théorème 66, cette convergence est uniforme par rapport à x, lorsque x décrit un compact. Mais, si nous appelons $H \times K$ un produit de compacts contenant le support de φ, toutes ces intégrales sont nulles pour $x \notin H$.

Autrement dit les fonctions ψ_n définies dans la démonstration du théorème 75 convergent, pour n infini, vers la fonction ψ, uniformément sur H, tout en gardant leur support dans H.

Alors, d'après le corollaire 2 du théorème 65, les $\mu_n(\psi_n)$ convergent vers $\mu(\psi)$, et ceci démontre le théorème.

Corollaire - Il existe une suite P_ℓ de polynômes ≥ 0 sur $\mathbb{R}^n$ tels que les mesures $P_\ell\, dx$ (où $dx = dx_1 \otimes dx_2 \otimes \ldots \otimes dx_n$) convergent vaguement vers δ pour ℓ tendant vers $+\infty$.

Démonstration Nous avons vu au théorème 68 qu'il existe une suite Q_ℓ de polynômes ≥ 0 sur $\mathbb{R}$ tels que les mesures $Q_\ell(x)\, dx$ convergent vaguement vers $\delta = d\delta(x)$ pour ℓ infini. Alors, les polynômes $P_\ell(x_1, x_2, \ldots, x_n) = Q_\ell(x_1) Q_\ell(x_2) \ldots Q_\ell(x_n)$ sont bien ≥ 0, et $Q_\ell(x_1)\, dx_1 \otimes Q_\ell(x_2)\, dx_2 \ldots \otimes Q_\ell(x_n)\, dx_n$ converge bien, pour ℓ infini, d'après le théorème 80, vers $d\delta(x_1) \otimes d\delta(x_2) \ldots \otimes d\delta(x_n) = d\delta(x_1, x_2, \ldots, x_n)$.

Théorème 80 bis (théorème d'approximation polynomiale de Weierstrass).

Soit K un compact de $\mathbb{R}^n$, $\vec{f}$ une fonction continue sur K à valeurs dans un Banach $\vec{F}$. Quel que soit $\varepsilon > 0$, il existe un polynome $\vec{Q}$ sur $\mathbb{R}^n$ à valeurs dans F, tel que $\underset{x \in K}{\mathrm{Max}} \| \vec{f}(x) - \vec{Q}(x) \| \leq \varepsilon$. Ou encore il existe une suite de polynomes $\vec{Q}_k$ qui convergent vers $\vec{f}$ uniformément sur K pour k tendant vers l'infini.

Démonstration - La formule $\vec{g} = \sum_{i=1}^{N} \vec{f}(a_i)\alpha_i = \sum_{i=1}^{N} \vec{g}_i \alpha_i$, $\alpha_i \in C(\mathbb{R}^n)$, utilisée au corollaire 8 du théorème 11 (partition de l'unité) donne d'abord une fonction continue $\vec{g}$ sur $\mathbb{R}^n$ tout entier décomposable, à support compact, telle que

$$\| \vec{f} - \vec{g} \| \leq \frac{\varepsilon}{2} \text{ sur } K \;^{*}$$

soit alors P_ℓ une suite de polynomes tels que les mesures $P_\ell\, dx$ convergent vaguement vers δ, suivant le corollaire précédent. Posons

$$(IV,8;55)\quad \begin{cases} \vec{Q}_\ell(x) = \int_{\mathbb{R}^n} \vec{g}(x-t) P_\ell(t)\,dt = \int_{\mathbb{R}^n} P_\ell(x-t)\vec{g}(t)\,dt \\ = \iint\cdots\int_{\mathbb{R}^n} \vec{g}(x_1-t_1, x_2-t_2, \ldots, x_n-t_n) P_\ell(t_1, t_2, \ldots, t_n)\,dt_1\,dt_2 \ldots dt_n . \end{cases}$$

Ces quantités ont un sens, comme intégrales de fonctions continues à support compact.

Montrons que la suite des $\vec{Q}_\ell$ converge uniformément vers $\vec{g}$ sur K, et que ce sont des polynomes; alors il en existera un, $\vec{Q}_\ell$ tel que $\| \vec{Q}_\ell - \vec{g} \| \leq \frac{\varepsilon}{2}$ sur k, et le théorème sera démontré.

Or d'abord on a $P_\ell = \sum_p c_p x^p$ (notations de la page 249; $p = (p_1, p_2, \ldots, p_n) \in \mathbb{N}^n$, et $x^p = x_1^{p_1} x_2^{p_2}, \ldots, x_n^{p_n}$; nous avons supprimé des flèches pour alléger) d'où l'on déduit

$$(IV,8;56)\quad \begin{cases} (x-t)^p = \sum_{q \leq p} \binom{p}{q} x^q (-t)^{p-q} \\ P_\ell(x-t) = \sum_{\substack{p,q \\ q \leq p}} c_p \binom{p}{q} x^q (-t)^{p-q} \\ \vec{Q}_\ell(x) = \sum_{\substack{p,q \\ q \leq p}} c_p \binom{p}{q} x^q \int_{\mathbb{R}^n} (-t)^{p-q} \vec{g}(t)\,dt \\ \quad = \sum_q d_q x^q , \end{cases}$$

donc $\vec{Q}_\ell$ est bien un polynome.

Montrons la convergence uniforme, Comme $\vec{g}$ est décomposable, $\vec{g} = \sum_{i=1}^{N} \vec{g}_i \alpha_i$, on se ramène immédiatement à montrer que chaque intégrale $\int_{\mathbb{R}^n} \alpha_i(x-t) P_\ell(t)\,dt$ converge vers $\alpha_i(x) = \int_{\mathbb{R}^n} \alpha_i(x-t)\, d\delta(t)$, uniformément pour x dans K.

* Dans l'énoncé de ce corollaire 8, $\vec{f}$ était supposée définie sur X (ici $\mathbb{R}^n$) tout entier, et à support compact. Mais c'était inutile, seules les valeurs de $\vec{f}$ sur K interviennent,

or $(x,t) \longrightarrow \alpha_i(x-t)$ est une fonction continue sur $K \times \mathbb{R}^n$; lorsque x parcourt K, la fonction partielle $t \longrightarrow \alpha_i(x-t)$ a son support dans un compact fixe de $\mathbb{R}^n$, à savoir la somme vectorielle $K - T_i$ où T_i est le support de α_i ; on est donc exactement dans les conditions d'application du corollaire 3 du théorème 66, avec interversion des rôles des variables x et t, ce qui démontre le théorème.

Remarque. Il y a d'autres méthodes pour démontrer le théorème d'approximation de Weierstrass, mais celle-ci a l'avantage de donner certains résultats beaucoup plus forts. Supposons $\vec{f}$ définie et continue, non seulement sur K, mais sur $\mathbb{R}^n$.

Alors, pour construire une fonction $\vec{g}$ à support compact proche de $\vec{f}$ sur K, il y a un procédé beaucoup plus simple : on prend $\vec{g} = \alpha \vec{f}$, où α est une fonction continue, ou même C^∞, à support compact, égale à 1 sur un voisinage de K (corollaire 1 du théorème 11). Alors on a bien plus : $\vec{g}$ est égale à $\vec{f}$ sur un voisinage de K. Les $\vec{Q}_\ell$ convergeront vers $\vec{g}$ donc vers $\vec{f}$ uniformément sur K. Il est vrai qu'alors $\vec{g}$ n'est plus décomposable, et qu'on ne peut plus appliquer directement le corollaire 3 du théorème 66; mais on se convaincra facilement que, si ce dernier n'a été démontré que pour des fonctions scalaires, c'est à cause de la présence de mesures vectorielles, mais qu'il était valable pour des fonctions vectorielles en cas de mesures scalaires.

Mais, dans ce cadre, les $\vec{Q}_\ell$ ont des propriétés de convergence bien plus fortes. Supposons que $\vec{f}$ soit de classe C^m sur un voisinage de K ; alors, si α est C^m et de support dans ce voisinage, $\vec{g}$ est C^m dans $\mathbb{R}^n$ tout entier. Alors, pour tout p d'ordre $\leq m$, les dérivées $D^p \vec{Q}_\ell$ convergent vers $D^p \vec{f}$, uniformément sur K. Autrement dit on peut réaliser l'approximation polynomiale, à la fois pour $\vec{f}$ et ses dérivées. En effet, on se trouve dans un cas trivial où la dérivation sous le signe d'intégration est légitime (voir plus loin, corollaire du théorème 115) de sorte que

$$D^{p} \vec{Q}_{\ell}(x) = \int_{\mathbb{R}^n} D^{p} \vec{g}(x-t)\, P_{\ell}(t)\, dt \; ; \tag{IV,8;57}$$

et on est ramené au problème initial, avec $D^{p}\vec{g}$ au lieu de $\vec{g}$.

§ 9 PROPRIÉTÉS PARTICULIÈRES AUX MESURES DE RADON SUR LA DROITE RÉELLE R

Les propriétés que nous allons voir dans ce § font intervenir le fait que $\mathbb{R}$ est non seulement un espace localement compact, mais un ensemble totalement ordonné. Si a et b sont 2 points de $\mathbb{R}$, nous appellerons $[a,b[$ le segment $a \leq x < b$, si $a < b$, et le segment $b < x \leq a$ si $b < a$; de même pour $]a,b]$, $[a,b]$, $]a,b[$
Ce sera donc plutôt le segment au sens de la page 174 que l'intervalle au sens de la page 21.

Introduction du symbole $\int d\vec{\mu}$

Définition Soit $\vec{\mu}$ une mesure de Radon sur un intervalle $\mathbb{R}_1$ (ouvert, semi-ouvert ou fermé) de la droite réelle $\mathbb{R}$, à valeurs dans un espace de Banach $\vec{E}$. Supposons que cette mesure admette une base μ_0, mesure réelle $\geqslant 0$. On peut alors la prolonger en particulier aux fonctions caractéristiques des intervalles compacts. Alors, pour a et b dans $\mathbb{R}_1$, on appelle $\int_a^b d\vec{\mu}$ la quantité définie comme suit :

$$\int_a^b d\vec{\mu} = \begin{cases} \int_{[a,b[} d\vec{\mu}, & \text{si } a < b, \\ \vec{0} & \text{si } a = b, \\ -\int_{[b,a[} d\vec{\mu}, & \text{si } b < a^{*}. \end{cases} \tag{IV,9;1}$$

Si, en particulier, la mesure $d\vec{\mu}$ s'écrit sous la forme $\vec{f}(x)\,dx$, où $\vec{f}$ est localement intégrable par rapport à la mesure dx, alors on a, par définition

$$\int_a^b d\vec{\mu} = \begin{cases} \int_{[a,b]} \vec{f}(x)\,dx, & \text{si } a \leqslant b \\ -\int_{[b,a]} \vec{f}(x)\,dx, & \text{si } a \geqslant b. \end{cases} \tag{IV,9;2}$$

Théorème 81 - On a la relation de Chasles :

$$\int_a^b d\vec{\mu} + \int_b^c d\vec{\mu} + \int_c^a d\vec{\mu} = \vec{0}. \tag{IV,9;3}$$

Démonstration évidente, par examen des différents cas de figure.

Intégrales indéfinies

Définition On dit qu'une fonction M définie sur $\mathbb{R}_1$, à valeurs dans un espace affine normé complet E, est une Intégrale indéfinie de la mesure de Radon $\vec{\mu}$, définie sur $\mathbb{R}_1$ à valeurs dans $\vec{E}$, ayant une base réelle $\geqslant 0$, si, quels que soient a et b dans $\mathbb{R}_1$, on a :

* Il faut faire très attention : $\vec{\mu}$ a peut-être des masses ponctuelles, et il n'est donc pas indifférent qu'on ait pris des intervalles semi-ouverts, contenant l'extrémité gauche, et non l'extrémité droite.

(IV,9;4) $$\overline{M(b) - M(a)} = \int_a^b d\vec{\mu} \,.$$

Si $\vec{\mu}$ est de la forme $d\vec{\mu} = \vec{f}\, d\mu_0$ où μ_0 est une mesure réelle ≥ 0, et $\vec{f}$ une fonction localement μ_0-intégrable, on dit aussi que M est une <u>intégrale indéfinie de $\vec{f}$ par rapport à la mesure</u> μ_0.

<u>Thecrème 82 - Soit E une ace affine normé complet, $\vec{\mu}$ mesure de Radon sur $\mathbb{R}_1$, à valeurs dans $\vec{E}$, de base ≥ 0. Alors $\vec{\mu}$ a une infinité d'intégrales indéfinies à valeurs dans E. Si c est un point de $\mathbb{R}_1$, la fonction $\vec{M}: x \to \int_c^x d\vec{\mu}$ est une intégrale indéfinie à valeurs dans $\vec{E}$, et on obtient toutes les intégrales indéfinies a valeurs dans E en lui ajoutant une constante arbitraire de E.</u>

<u>Démonstration</u> Que M réponde à la question résulte de la relation de Chasles. D'autre part, si 2 fonctions, a valeurs respectivement dans $\vec{E}$ et E_1 sont des intégrales indéfinies de $\vec{\mu}$, la formule (IV,9;4) montre que leur différence prend la même valeur en deux points a et b quelconques de $\mathbb{R}_1$, donc est constante sur $\mathbb{R}_1$; inverse=ment la somme d'une intégrale indéfinie à valeurs dans E et d'une constante de E est une intégrale indéfinie à valeurs dans E.

<u>Exemple</u> Considérons <u>la fonction d'Heaviside</u>, c'est-à-dire la fonction réelle Y sur $\mathbb{R}$ définie par :

(IV,9;5) $$Y(x) = \begin{cases} 0 \text{ pour } x \leqslant 0 \\ 1 \text{ pour } x > 0 \,; \end{cases}$$

c'est une intégrale indéfinie de δ, avec $Y(x) = \int_{-\infty}^x d\delta$ ou $\int_a^x d\delta$, $a \leqslant 0$ quelconque.

<u>Remarque</u> Cn en déduit aussi que, la différence entre 2 intégrales indéfinies à valeurs dans E est une constante de $\vec{E}$ et que la somme d'une intégrale indéfinie à valeurs dans E et d'une constante de $\vec{E}$ est une intégrale indéfinie à valeurs dans E.

Théorème 83 - Toute fonction M à valeurs dans E, qui est une intégrale indéfinie de la mesure de Radon $\vec{\mu}$ sur $\mathbb{R}_1$ à valeurs dans E est continue à gauche en tout point de $\mathbb{R}_1$. Elle est continue à droite en un point c de $\mathbb{R}_1$, si ~~et~~ seulement si $\vec{\mu}(\{c\}) = \vec{0}$; de toute manière, elle a une limite droite, et un saut égal à $\vec{\mu}(c)$.

Démonstration soit $h \geq 0$. Alors $M(c) - M(c-h)$ est la mesure $\vec{\mu}([c-h, c[)$, ou intégrale, par rapport à $\vec{\mu}$, de la fonction caractéristique φ_h de cet intervalle. Soit $d\vec{\mu} = \vec{f}\, d\mu_0$, $\mu_0 \geq 0$. Alors $\vec{\mu}([c-h, c[) = \int \varphi_h \vec{f}\, d\mu_0$. Si $h > 0$ tend vers 0, la fonction caractéristique φ_h converge simplement vers la fonction 0, fonction caractéristique de l'ensemble vide; donc la fonction $\varphi_h \vec{f}$ converge simplement vers $\vec{0}$. Par ailleurs, elle reste bornée en norme par $\|\vec{f}\|$ (PR, , dès que $h \leq h_0$; or cette fonction est ≥ 0 et μ_0-intégrable, puisque $\vec{f}$ est localement μ_0-intégrable. D'après le théorème de Lebesgue, étendu aux limites généralisées (remarque page 491), on voit donc bien que les différences précédentes convergent vers 0 pour h tendant vers 0, ce qui démontre la continuité à gauche. Si par contre nous raisonnons sur les différences $M(c+h) - M(c) = \int \psi_h \vec{f}\, d\mu_0$, où ψ_h est la fonction caractéristique de $[c, c+h[$, nous voyons que, cette fois, ces fonctions ψ_h convergent vers la fonction caractéristique ψ_0 de l'ensemble $\{c\}$ réduit au point c, les $\vec{f}\psi_h$ sont encore majorées par une fonction ≥ 0 intégrable fixe, et le théorème de Lebesgue nous montre cette fois que ces différences convergent. pour h tendant vers 0, vers $\int \psi_0 \vec{f}\, d\mu_0 = \vec{\mu}(\{c\})$. On en déduit bien la continuité à droite, si et seulement si $\{c\}$ est de mesure nulle pour $\vec{\mu}$. En outre, M a toujours une limite à droite, et $M(c+0) = M(c) + \vec{\mu}(\{c\})$

Remarque Soient a, c, des points de $\mathbb{R}_1$, $a < c$. Le raisonnement précédent montre que $\vec{\mu}([a, c'[)$ tend vers $\vec{\mu}([a, c[)$ quand c' tend vers c par valeurs $< c$, et vers $\vec{\mu}([a, c]) = \vec{\mu}([a, c[) + \vec{\mu}(\{c\})$ quand c' tend vers c par valeurs $> c$. Bien entendu $\vec{\mu}([a, c'])$ a les mêmes limites, $\vec{\mu}([a, c[)$ et $\vec{\mu}([a, c])$.

<u>Corollaire 1 - Si $\vec{f}$ est une fonction définie sur $\mathbb{R}_1$, à valeurs dans un espace de Banach E, intégrable par rapport à la mesure de Lebesgue dx, toute intégrale indéfinie $x \longrightarrow \int_c^x \vec{f}(t)\,dt$ est une fonction continue de x.</u>

En effet, par rapport à la mesure $\vec{f}\,dx$, n'importe quel point est de mesure nulle.

<u>Corollaire 2.- Une intégrale indéfinie d'une mesure sur $\mathbb{R}_1$ est une fonction réglée, donc borélienne et par suite mesurable par rapport à toute mesure de Radon sur $\mathbb{R}_1$.</u>

Le fait que la fonction soit réglée résulte immédiatement du théorème. Comme la fonction caractéristique d'un intervalle est borélienne, il en est de même d'une fonction en escalier, donc d'une fonction réglée, par passage à la limite (théorèmes 8 et 23).

Fonctions à variation bornée sur la droite

Soit M une application d'une partie $\mathbb{R}_1$ de la droite réelle $\mathbb{R}$ dans un espace métrique E. On appelle <u>variation totale de M</u> la borne supérieure $V(M)=V(\mathbb{R}_1;M)$ des sommes

(IV,9;6)
$$\sum_A = \sum_{i=0}^{n-1} d\left(M(c_i),\, M(c_{i+1})\right)$$

pour toutes les décompositions A de $\mathbb{R}_1$, c'est-à-dire toutes les suites finies croissantes de points de $\mathbb{R}_1$, $c_0, c_1, \ldots, c_n$. C'est un nombre ≥ 0, fini ou égal à $+\infty$; si $\mathbb{R}_2 \subset \mathbb{R}_1$, on a bien entendu $V(\mathbb{R}_2;M) \leq V(\mathbb{R}_1;M)$. Si $V(\mathbb{R}_1;M)$ est fini, on dit que <u>M est à variation bornée sur $\mathbb{R}_1$</u> (on devrait dire : à variation totale finie). Les cas les plus fréquents sont ceux où $\mathbb{R}_1$ est un intervalle (borné ou non, ouvert, semi-ouvert, ou fermé), ou l'ensemble $\mathbb{N}$ des entiers ≥ 0 (voir page 133), ou l'ensemble $\mathbb{Z}$ de tous les entiers. <u>Désormais nous supposerons que $\mathbb{R}_1$ est un intervalle.</u>

Une fonction lipschitzienne sur un intervalle $\mathbb{R}_1$, borné de $\mathbb{R}$ est à variation bornée

En effet, de $d(M(x''),M(x')) \leq k|x''-x'|$, on déduit que la variation totale de M est majorée par k fois la longueur de $\mathbb{R}_1$. En particulier, une fonction M sur $\mathbb{R}_1$, à valeurs dans un espace affine normé E, dérivable et à dérivée bornée, est à variation bornée si $\mathbb{R}_1$ est de longueur finie; car, d'après le corollaire 2 du théorème 13 du chapitre III, elle est lipschitzienne.

Bien entendu si N est à valeurs dans un espace affine normé E, $\vec{M}$ à valeurs dans l'espace vectoriel associé $\vec{E}$, et si k est un scalaire, on a :

$$(IV,9;6^{bis}) \qquad \begin{cases} V(N+\vec{M}) \leq V(N) + V(\vec{M}) \\ V(k\vec{M}) = |k|\, V(\vec{M}) \end{cases}$$

Cela prouve que les fonctions à variation bornée sur $\mathbb{R}_1$, à valeurs dans $\vec{E}$, forment un espace vectoriel.

Mais $\vec{M} \longrightarrow V(\vec{M})$ n'est pas une norme sur cet espace vectoriel; en effet $V(\vec{M}) = 0$ est équivalent à $\vec{M}$ = constante, et no* nécessairement $\vec{M} \equiv \vec{0}$.

Remarquons que la somme (IV,9;6) augmente, quand on remplace la suite finie des c_i par une suite finie qui contient. On peut donc, dans la recherche de la borne supérieure, se borner à considérer celles des décompositions Δ pour lesquelles figurent, parmi les c_i, des points en nombre fini donnés à l'avance.

Théorème 83 bis - Si a, b, c, sont 3 points de $\mathbb{R}_1$, $a < b < c$, la variation totale de M dans $[a, c]$ est la somme des variations totales dans $[a, b]$ et dans $[b, c]$:

$$(IV,9;6^{ter}) \qquad V([a,c]\ ;M) = V([a,b];M) + V([b,c];M)^{*}.$$

* Ce renvoi se trouve à la page 602 _

En effet, on peut calculer le 1er membre comme borne supérieure de sommes $\sum$, correspondant à des suites finies de points $c_0 = a, c_1, c_2, \ldots, c_n = c$, parmi lesquels figure le point b.

Théorème 84 - Toute fonction à variation bornée sur un intervalle $\mathbb{R}_1$ de $\mathbb{R}$, à valeurs dans un espace métrique E est bornée; si E est complet, elle est reglée, donc borélienne et mesurable pour toute mesure de Radon sur $\mathbb{R}_1$.

Si M est à variation totale finie, celle-ci majore $d(M(x), M(y))$, pour x et y dans $\mathbb{R}_1$; donc M est bien bornée. Montrons qu'elle est réglée, si E est complet. soit $x_0, x_1, x_2, \ldots, x_n, \ldots$, une suite de points > C tendant vers c pour n infini.

Pour montrer que $M(x_0), M(x_1), \ldots, M(x_n), \ldots$ a une limite pour n infini, on peut changer l'ordre des termes de cette suite, ce qui permet de supposer la suite des x_n décroissante.

Alors la somme $\sum_{i=0}^{n-1} d(M(x_i), M(x_{i+1}))$ est bornée, indépendamment de n, par la variation totale de M. Donc $\sum_{i=0}^{\infty} d(M(x_i), M(x_{i+1})) < +\infty$. Alors, pour $m \geq n+1$, m et n tendant vers $+\infty$, la somme $\sum_{i=n}^{m-1}$ tend vers 0, et a fortiori la distance $d(M(x_n), M(x_m))$ tend vers 0 : la suite des $M(x_n)$ est une suite de Cauchy; comme E est complet, elle est bien convergente. Sa limite est indépendante de la suite x_n considérée; car, avec 2 suites, $(x'_n)_{n \in N}$, $(x''_n)_{n \in N}$, on peut former la suite mélangée $x'_0, x''_0, x'_1, x''_1, x'_2, x''_2, \ldots x'_n, x''_n, \ldots$, pour laquelle ia suite des valeurs de M doit avoir une limite, ce qui oblige bien les

* Renvoi de la page précédente.

On pourrait plutôt s'attendre à avoir, par exemple :

$$V([a,c[;M) = V([a,b[;M) + V([b,c[;M).$$

c'est inexact en général. C'est vrai si M est continue à gauche au point b, car alors (lemme 2, page 604) :

$$V([a,b];M) = V([a,b[;M).$$

2 suites $M(x_n)$, $M(x'_n)$, à avoir la même limite; soit $M(c+0)$ la limite commune à toutes ces suites.

Alors $M(x)$ tend vers $M(c+0)$, quand x tend vers c par valeurs $> c$; sans quoi, il existerait un $\varepsilon > 0$ tel que, pour tout $n \geq 1$, on puisse trouver un x_n,

$c < x_n \leq c + \frac{1}{n}$, avec $d(M(x_n), M(c+0)) > \varepsilon$; alors x_n tendrait vers c pour n infini, par valeurs $> c$ et

$M(x_n)$ ne tendrait pas vers $M(c+0)$, ce qui serait contraire aux résultats antérieurs.

On voit de même que $M(x)$ a une limite $M(c-0)$, quand x tend vers c par valeurs $< c$; donc M est bien réglée.

Remarque - La réciproque est inexacte : une fonction réglée, et même une fonction continue sur un intervalle $[a, b]$ de $\mathbb{R}$, même à valeurs réelles, n'est pas nécessairement à variation bornée.

Si par exemple on considère la fonction réelle suivante sur $[0, 1]$: au point d'abscisse $\frac{1}{n}$, $n \geq 1$, $M(\frac{1}{n}) = \frac{(-1)^n}{n}$; dans l'intervalle $\left[\frac{1}{n+1}, \frac{1}{n}\right]$, elle est linéaire affine; $M(0) = 0$; M est manifestement continue; sa variation totale est $(1 + \frac{1}{2}) + (\frac{1}{2} + \frac{1}{3}) + (\frac{1}{3} + \frac{1}{4}) + \cdots = +\infty$.

Z

Théorème 84 bis - Soit M une application de $[a, b]$ $(a < b)$ dans E, à variation bornée. Alors, si on appelle $V([a,x]; M)$ la variation totale de la restriction de la fonction M à l'intervalle $[a, x]$ la fonction $x \longrightarrow V([a,x]; M)$ est croissante, et elle est continue à gauche (resp. à droite) au point c, si et seulement si M est continue à gauche (resp. à droite) au point c.

Nous montrerons deux lemmes, d'où résultera le théorème.

Lemme 1 - Si M est une application de $[a, c[$ dans un espace métrique E, alors, lorsque x tend vers c par valeurs $< c$, $V([a,x]; M) \leq +\infty$ tend vers $V([a, c[; M) \leq +\infty$.

Démonstration du lemme 1 - La fonction considérée est croissante, et a donc une limite $\leqslant V([a,c[;M)$ pour x tendant vers c par valeurs $< c$. Mais, si V est un nombre quelconque $< V([a,c[;M)$, il existe, par définition, une décomposition A de $[a,c[$, soit $c_0 = a \leqslant c_1 \leqslant c_2 \ldots \leqslant c_n < c$, telle que

(IV,9;7) $$\sum_{\Delta} \geqslant V$$

Alors a fortiori, pour $c_n \leqslant x < c$:

(IV,9;8) $$V([a,x];M) \geqslant V([a,c_n];M) \geqslant \sum_{\Delta} \geqslant V ,$$

ce qui démontre le lemme 1.

Lemme 2 - Si M est une application de $[a,c]$ $(a<c)$ dans un espace métrique E, à variation bornée, on a $V([a,c];M) = V([a,c[;M)$ si et seulement si M est continue à gauche au point c *.

Démonstration du lemme 2 - On a évidemment $V([a,c[;M) \leqslant V([a,c];M)$.

Supposons d'abord M continue à gauche au point c. Alors, $\varepsilon > 0$ étant donné, il existe $\eta > 0$ tel que $c - x \leqslant \eta$, $x < c$, entraîne $d(M(c), M(x)) \leqslant \frac{\varepsilon}{2}$. Il existe d'autre part une décomposition A: $c_0 = a, c_1, c_2, \ldots, c_n = c$ de $[a,c]$, telle que

(IV,9;9) $$\sum_{\Delta} \geqslant V([a,c];M) - \frac{\varepsilon}{2} .$$

De plus, on peut supposer que $c_{n-1} \geqslant c - \eta$ (sans quoi on rajouterait à A le point $c - \eta$, ce qui ne pourrait qu'augmenter $\sum_A$).

* Si M n'est pas à variation bornée sur $[a,c]$ elle ne l'est pas non plus sur $[a,c[$, et on a $V([a,c];M) = V([a,c[;M) = +\infty$, que M soit ou non continue à gauche au point c.

(IV,9;10)
$$V([a,c[;M) \geq V([a,c_{n-1}];M) \geq \sum_{\Delta} - d(M(c),M(c_{n-1})) \geq \left(V([a,c];M) - \frac{\varepsilon}{2}\right) - \frac{\varepsilon}{2} = V([a,c];M) - \varepsilon .$$

comme ε est quelconque, on a bien $V([a,c];M) = V([a,c[;M)$.

Inversement, supposons cette égalité réalisée. Il existe alors une décomposition Δ de $[a,c[$ telle que l'on ait (IV,9;7), avec $V = V([a,c];M) - \varepsilon$. Alors, pour $c_n \leq x < c$, on aura

(IV,9;11)
$$V([a,c];M) - \varepsilon \leq \sum_{\Delta} \leq \sum_{\Delta} + d(M(x),M(c)) \leq V([a,c];M) ,$$ d'où

(IV,9;12)
$$d(M(x),M(c)) \leq \varepsilon \quad \text{pour} \quad c_n \leq x < c ,$$

et M est bien continue à gauche au point c .

<u>Démonstration du théorème.</u> Nous avons donc bien montré que $x \longrightarrow V([a,x];M)$ est continue à gauche au point c , si et seulement si M est continue à gauche au point c . Nous avons même trouvé. dans tous les cas, la limite de $V([a,x];M)$ quand x tend vers c par valeurs $< c$.

Si maintenant x tend vers C par valeurs $> c$, $V([x,b];M)$ tend toujours. comme le montre un raisonnement symétrique de celui du lemme 1, vers $V(]c,b];M)$; ceci est égal à $V([c,b];M)$, comme le montre un raisonnement symétrique de celui du lemme 2, si et seulement si M est continue à droite au point c . Alors $V([a,x];M) = V([a,b];M) - V([x,b];M)$ tend vers $V([a,c];M) = V([a,b];M) - V([c,b];M)$ si et seulement si M est continue à droite au point c .

Si $\mathbb{R}_1$ est un intervalle fermé $[a, b]$, et si M est continue à gauche *, au lieu d'exprimer la variation totale de M comme borne supérieure, on peut l'exprimer comme la limite des sommes (IV,9;6), correspondant à n'importe quelle suite $A_1, A, A_2, \ldots, A_n, \ldots$ de décompositions de $[a, b]$, de finesse tendant vers **0** pour n infini :

Théorème **85** - Soit **M** une application d'un intervalle fermé $[a, b]$ $(a < b)$ de $\mathbb{R}$ dans un espace métrique E, partout continue à gauche. Quel que soit le nombre V_1, strictement inférieur à la variation totale **V(M)** de **M** dans $[a, b]$, il existe un nombre $\eta > 0$ ~~, ayant~~ la propriété suivante : quelle que soit la suite finie croissante

$c_0 = a \leq c_1 \leq c_2 \leq \cdots \leq c_n = b$ de points de $[a, b]$, telle que tous les intervalles $[c_i, c_{i+1}]$ soient de longueur $\leq \eta$, on a l'inégalité :

$$\sum_{i=0}^{n-i} d(M(c_i), M(c_{i+1})) \geq V_1 . \tag{IV,9;13}$$

Démonstration. Soit V_2 tel que $V_1 < V_2 < V(M)$. D'après la définition de la variation totale comme borne supérieure, il existe une décompcsition $\Delta_0 : d_0 = a, d_1, d_2, \ldots, d_N = b$, telle que

$$\sum_{\Delta_0} = \sum_{j=0}^{'N-I} d(M(d_j), M(d_{j+1})) \geq V_2 . \tag{IV,9 14}$$

Choisissons alors η assez petit pour que :

1°/ $\eta < d_{i+j} - d_j$, pour tout j ; alors, si $c_{i+1} - c_i \leq \eta$, on sera sûr que $[c_i, c_{i+1}]$ contient au plus un des points d_j ;

* ou, naturellement, continue à droite.

2°/ $d(x, d_j) \leq \eta$, $x < d_j$. entraîne
$d(M(x), M(d_j)) \leq \frac{V_2 - V_1}{2(N-1)}$; et ceci pour tout $j = 1, 2, \ldots, N-1$
Un tel choix est possible, parce que M est continue à gauche en chaque point-d_6 .

Nous allons montrer que ce nombre η répond à la question. Soit donc A: $c_0 = a, c_1, \ldots, c_n = b$, une décomposition de $[a, b]$, de finesse $\leq \eta$

Si on ajoute aux c_i les points d_j , on a une décomposition A' qui donne une somme $\sum_{\Delta'}$ certainement au moins égale à $\sum_{\Delta_0}$, c'est-à-dire à V_2 .

Mais alors

$$\sum_{\Delta} = \sum_{\Delta'} - \sum_{i,j} \Big(d(M(c_i), M(d_j)) + d(M(d_j), M(c_{i+1})) - d(M(c_i), M(c_{i+1})) \Big), \tag{IV,9;15}$$

la dernière somme étant étendue à tous les systèmes (c_i, d_j, c_{i+1}) tels que $c_i < d_j < c_{i+1}$.

On a la majoration (inégalité triangulaire);

$$d(M(d_j), M(c_{i+1})) - d(M(c_i), M(c_{i+1})) \leq d(M(c_i), M(d_j)) \tag{IV,9;16}$$, d'où

$$\begin{aligned} & d(M(c_i), M(d_j)) + d(M(d_j), M(c_{i+1})) \\ & - d(M(c_i), M(c_{i+1})) \\ & \leq 2\, d(M(c_i), M(d_j)) \leq 2\, \frac{V_2 - V_1}{2(N-1)} . \end{aligned} \tag{IV,9;17}$$

Alors

(IV,9;18) $$\sum_{\Delta} \geq \sum_{\Delta'} - 2(N-1)\frac{V_2 - V_1}{2(N-1)}$$

$$= \sum_{\Delta'} - (V_2 - V_1),$$

et comme $\sum_{A'} \geq V_2$, on a bien $\sum_{A} \geq V_1$.

Le théorème suivant permet de caractériser les fonctions à variation bornée à valeurs dans un espace affine normé de dimensinn finie :

Théorème 86 - Soit E un espace affine normé de dimension finie-n, sur le corps des réels, muni d'un référentiel $0, \vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$. Pour qu'une fonction M sur un intervalle $\mathbb{R}_1$ de $\mathbb{R}$, à valeurs dans E, soit à variation bornée, il faut et il suffit que chacune de ses composantes soit à variation bornée. Pour qu'une fonction réelle sur $\mathbb{R}_1$ soit à variation bornée, il faut et il suffit qu'elle soit rence de 2 fonctions croissantes et bornées.

Démonstration . Si on remplace la norme de $\vec{E}$ par une norme equivalente, il résulte du théorème 12 du chapitre II qu'une fonction à variation bornée reste à variation bornée * . Le référentiel de E l'identifie à $\mathbb{R}^n$, et nous pouvons remplacer la norme de $\vec{E}$ par la norme équivalente (théoréme 13 du chapitre II) : $\|(x_1, x_2, \dots, x_n)\| = \sum_{j=1}^{n} |x_j|$ sur $\mathbb{R}^n$.

Si alors on représente la fonction M sur $\mathbb{R}_1$ par

$M(x) = 0 + \sum_{j=1}^{n} X_j(x)\vec{e}_j$, or a, pour toute décomposition Δ de $\mathbb{R}_1$:

(IV,9;19) $$\|\overrightarrow{M(c_{i+1}) - M(c_i)}\| = \sum_{j=1}^{n} \left| X_j(c_{i+1}) - X_j(c_i) \right|$$

* Naturellement la valeur de la variation totale change avec la norme; mais si elle est finie pour une norme, elle est finie pour toute autre norme équivalente.

Donc la variation totale de M est au moins égale à celle de chaque fonction X_j, et au plus égale à la somme des variations totales des X_j ; cela prouve bien que M est à variation bornée si et seulement si les X_j le sont.

Soit maintenant M une fonction réelle sur $\mathbb{R}_1$. Si elle est monotone, tous les $M(c_{i+1}) - M(c_i)$, correspondant à une décomposition quelconque Δ, sont de même signe, et la somme (IV,9;6) est exactement égale, quelle que soit Δ, à $|M(c_n) - M(c_0)|$; donc, si M est monotone et bornée, elle est à variation bornée. Alors la différence de 2 fonctions croissantes et bornées est sûrement à variation bornée. Inversement, soit M une fonction réelle à variation bornée sur $\mathbb{R}_1$. Choisissons un point c de $\mathbb{R}_1$. Pour x dans $\mathbb{R}_1$ appelons $V(x)$ la quantité :

$$(IV,9;20)\quad V(x) = \begin{cases} \text{variation de } M \text{ sur } [c,x] \text{, si } x > c \text{ ;} \\ 0 \quad \text{si } x = c. \\ -(\text{variation de } M \text{ sur } [x,c]) \text{, si } x < c \text{ ;} \end{cases}$$

$V(x)$ est évidemment une fonction croissante et bornée de $x \in \mathbb{R}_1$. Mais, si $y > x$, la différence $V(y) - V(x)$ est la variation totale de M sur $[x, y]$, donc elle est au moins égale à $M(y) - M(x)$. Donc

$$(IV,9;21)\quad \begin{cases} V(y) - V(x) \geq M(y) - M(x) \quad \text{ou} \\ V(y) - M(y) \geq V(x) - M(x) \text{ ,} \end{cases}$$

• En utilisant encore le fait que la somme (IV,9;6) ne peut qu'augmenter quand on remplace la suite des c_i par une autre qui la contient, on voit même que la variation de M est exactement la somme des variations des X_j. .

donc la différence $V - M$ est une fonction croissante et bornée W, et $M = V-W$ est bien différence de 2 fonctions croissantes et bornées.

Corollaire - soit $|a, b|$ un intervalle de $\mathbb{R}$, M une fonction réelle sur $|a,b|$ bornee et monotone par morceaux. Alors M est a variation bornee.

On dit que M est monotone par morceaux, si $|a,b|$ est réunion d'un nombre fini d'intervalles, $|a_0, a_1]$, $[a_1, a_2], \ldots, [a_{n-1}, a_n|$, $a_0 = a$, $a_n = b$, tels que dans chaque intervalle $]a_i, a_{i+1}[$ elle soit monotone (Toutes les fonctions usuelles sont monotones par morceaux, si $|a,b|$ est borné). Alors on a, en supposant par exemple $|a,b| =]a,b[$:

$$(\text{IV},9;22) \quad V(|a,b|;M) = V(|a_0, a_1];M) + V([a_1, a_2];M) + \cdots + V([a_{n-1}, a_n|;M)$$

$$\sum_{i=0}^{n-1} |M(a_{i+1}) - M(a_i)| + \sum_{i=1}^{n-1} \left(|M(a_i) - M(a_i - 0)| + |M(a_i+0) - M(a_i)| \right)$$

Fonctions à variation bornée et intégrales indéfinies

Si nous avons introduit les fonctions à variation bornée, c'est qu'elles ont une relation simple avec les intégrales indéfinies de mesures de Radon. Disons d'abord qu'une fonction sur $\mathbb{R}_1$, à valeurs dans, E, est dite localement à variation. bornée, si sa restriction à tout intervalle compact $[a,b]$ de $\mathbb{R}_1$ est à variation bornée.

Théorème 87 - Soit $\vec{\mu}$ une mesure de Radon sur un intervalle $\mathbb{R}$ de $\mathbb{R}$, à valeurs dans un espace de Banach E, de base $\lambda \geq 0$: $\vec{\mu} = \vec{\mu}\lambda$.

Alors toute intégrale indéfinie M de $\vec{\mu}$ est localement à variation bornée, et elle est à variation bornée sur tout intervalle (ouvert, semi-ouvert, ou fermé) $|a,b|$ de $\mathbb{R}_1$ sur lequel $\vec{\mu}$ est λ-intégrable. Pour tout intervalle $|a,b|$ de $\mathbb{R}_1$, $a < b$, on a :

$$(\text{IV},9;23) \quad V(|a,b|;M) = \int_{|a,b[} \|\vec{\mu}\| \, d\lambda \leq +\infty .$$

Tout le théorème résulte de (IV, 9; 23). Or, si $c_0, c_1, c_2, \dots, c_n$, est une décomposition Δ de $]a,b[$, on a

$$\text{(IV,9;24)} \quad \sum_{\Delta} = \sum_{i=0}^{n-1} \left\| \overrightarrow{M(c_{i+1}) - M(c_i)} \right\| = \sum_{i=0}^{n-1} \left\| \int_{c_i}^{c_{i+1}} \vec{f}\, d\lambda \right\| \leq \int_{]a,b[} \|\vec{f}\|\, d\lambda .$$

En prenant la borne supérieure des $\sum_{\Delta}$, on a l'inégalité

$$\text{(IV,9;24bis)} \quad V(]a,b[\,; M) \leq \int_{]a,b[} \|\vec{f}\|\, d\lambda .$$

Nous avons démontré (IV,9;23) avec $\leq$ au lieu de $=$. Montrons d'abord que l'on a bien l'égalité si $\vec{f}$ est continue sur $\mathbb{R}_1$, et si $[a,b]$ est un intervalle fermé. donc compact, de $\mathbb{R}_1$. Posons en effet $\Lambda = \int_a^b d\lambda$. La fonction $\vec{f}$ est uniformément continue sur $[a,b]$; donc,

$\varepsilon > 0$ étant donné, il existe $\eta > 0$ tel que $|x' - x''| \leq \eta$ entraîne $\|\vec{f}(x') - \vec{f}(x'')\| \leq \frac{\varepsilon}{2\Lambda}$.

Soit alors Δ une décomposition de $[a,b]$, $c_0 = a$, $c_n = b$, de finesse $\leq \eta$. On a, pour $c_i \leq x \leq c_{i+1}$

$$\text{(IV,9;25)} \quad \Big| \|\vec{f}(x)\| - \|\vec{f}(c_i)\| \Big| \leq \|\vec{f}(x) - \vec{f}(c_i)\| \leq \frac{\varepsilon}{2\Lambda} .$$

Alors on a les minorations suivantes :

$$\text{(IV,9;26)} \quad \overrightarrow{M(c_{i+1}) - M(c_i)} = \int_{c_i}^{c_{i+1}} \vec{f}(x)\, d\lambda(x)$$

$$= \vec{f}(c_i) \int_{c_i}^{c_{i+1}} d\lambda + \int_{c_i}^{c_{i+1}} \left(\vec{f}(x) - \vec{f}(c_i)\right) d\lambda(x) ;$$

$$\left\| \overrightarrow{M(c_{i+1}) - M(c_i)} \right\| \geq \|\vec{f}(c_i)\| \int_{c_i}^{c_{i+1}} d\lambda - \frac{\varepsilon}{2\Lambda} \int_{c_i}^{c_{i+1}} d\lambda .$$

Mais

$$(\text{IV},9;27)\qquad \|\vec{f}(c_i)\| \int_{c_i}^{c_{i+1}} d\lambda \geq \int_{c_i}^{c_{i+1}} \|\vec{f}(x)\| d\lambda(x) - \int_{c_i}^{c_{i+1}} \|\vec{f}(x) - \vec{f}(c_i)\| d\lambda(x)$$

$$\geq \int_{c_i}^{c_{i+1}} \|\vec{f}\| d\lambda - \frac{\varepsilon}{2\Lambda}\int_{c_i}^{c_{i+1}} d\lambda .$$

Finalement

$$(\text{IV},9;28)\qquad \|\overline{M(c_{i+1}) - M(c_i)}\| \geq \int_{c_i}^{c_{i+1}} \|\vec{f}\| d\lambda - \frac{\varepsilon}{\Lambda}\int_{c_i}^{c_{i+1}} d\lambda .$$

On en déduit

$$(\text{IV},9;29)\qquad V([a,b];M) \geq \sum_{\Delta} \geq \int_a^b \|\vec{f}\| d\lambda - \frac{\varepsilon}{\Lambda}\int_a^b d\lambda$$

$$\geq \int_a^b \|\vec{f}\| d\lambda - \varepsilon ;$$

comme ε est arbitraire, on a $V([a,b];M) \geq \int_a^b \|\vec{f}\| d\lambda$,

ce qui, avec (IV,9;24 bis), démontre bien (IV,9;23). Si maintenant $\vec{f}$ est quelconque et $[a,b]$ toujours compact, on sait tout de même que $\vec{f}$ est λ-intégrable sur $[a,b]$; donc, d'après le théorème 32, $\varepsilon > 0$ étant donné, il existe une fonction $\vec{q}$, continue sur $[a,b]$, telle que

$$(\text{IV},9;30)\qquad \int_a^b \|\vec{f} - \vec{q}\| d\lambda \leq \frac{\varepsilon}{2} .$$

Soit N une intégrale indéfinie de la mesure $\vec{q}\,\lambda$, de base λ . Alors on a :

$$(\text{IV},9;31)\qquad V([a,b];M) \geq V([a,b];N) - V([a,b];\overrightarrow{M-N}) .$$

Mais on a l'égalité

$$\text{(IV,9;32)} \qquad V([a,b];N) = \int_a^b \|\vec{q}\|\,d\lambda\,,$$

puisque $\vec{q}$ est continue sur $[a,b]$. Et on a, d'après (IV,9;24 bis), l'inégalité :

$$\text{(IV,9;33)} \qquad V([a,b];\overrightarrow{M-N}) \leqslant \int_a^b \|\vec{p}-\vec{q}\|\,d\lambda \leqslant \frac{\varepsilon}{2}$$

Donc

$$\text{(IV,9;34)} \qquad \begin{aligned} V([a,b];M) &\geqslant \int_a^b \|\vec{q}\|\,d\lambda - \frac{\varepsilon}{2} \\ &\geqslant \int_a^b \|\vec{p}\|\,d\lambda - \int_a^b \|\vec{p}-\vec{q}\|\,d\lambda - \frac{\varepsilon}{2} \\ &\geqslant \int_a^b \|\vec{p}\|\,d\lambda - \varepsilon. \end{aligned}$$

Comme ε est arbitrairement petit, on en déduit encore (IV,9;23).

Si maintenant $\vec{p}$ est quelconque, et $|a,b|$ non nécessairement fermé, la formule subsiste par passage à la limite : par exemple, si $]a,b[$ est ouvert, $V(]a,b[;M)$ est la limite de $V([a',b'];M)$ lorsque a' tend vers a par valeurs $> a$ et b' vers b par valeurs $< b$, d'après le lemme 1 du théorème 84 bis, et $\int_{]a,b[} \|\vec{p}\|\,d\lambda$ est la limite de $\int_{[a',b'[} \|\vec{p}\|\,d\lambda$, d'après le théorème 36 de Fatou.

<u>Corollaire</u> - <u>Si $\vec{p}$ est une fonction sur $\mathbb{R}_1$, à valeurs dans un espace de Banach $\vec{E}$, localement dx-intégrable, l'intégrale indéfinie $M : x \longrightarrow \int_c^x \vec{p}(t)\,dt$ est une fonction continue, à variation bornée sur tout intervalle fermé $[a,b]$ de $\mathbb{R}_1$; sa variation totale sur un intervalle (ouvert, semi-ouvert, ou fermé) $|a,b|$ $(a<b)$ de $\mathbb{R}_1$ est donnée</u> par :

$$\text{(IV,9;35)} \qquad V(|a,b|;M) = \int_{|a,b|} \|\vec{p}(t)\|\,dt$$

On a la réciproque suivante, si $\vec{E}$ est de dimension finie :

Théorème 88 - Soit M une fonction, localement à variation bornée et continue à gauche, sur un intervalle $\mathbb{R}_1$ de $\mathbb{R}$, à valeurs dans un espace affine normé E de dimension finie. Alors M est l'intégrale indéfinie d'une mesure $\vec{\mu}$ sur $\mathbb{R}_1$, à valeurs dans $\vec{F}$, de base $\geqslant 0$; cette mesure est déterminée d'une manière unique, si $\mathbb{R}_1$ n'a pas de maximum, ou si l'on impose à $\vec{\mu}$ de ne pas porter de masse en ce maximum. En outre, pour tout intervalle $|a,b|$ de $\mathbb{R}_1$, $a<b$, on a :

$$(IV,9;36) \qquad \int_a^b d\,|\vec{\mu}| = |\vec{\mu}|([a,b[) = V([a,b[\,;M)$$

$$= V([a,b]\,;M)$$

$|\vec{\mu}|$ étant la plus petite majorante absolue de $\vec{\mu}$. La mesure $\vec{\mu}$ est de norme finie sur $\mathbb{R}_1$, si et seulement si M est à variation bornée sur $\mathbb{R}_1$.

Si L est une fonction réelle, croissante, continue à gauche sur $\mathbb{R}_1$, elle est l'intégrale indéfinie d'une mesure $\lambda \geqslant 0$; si, pour tout couple de points x', x'', de $\mathbb{R}_1$, $x' \leqslant x''$, on a,

$$(IV,9;37) \qquad \left\| \overrightarrow{M(x'')-M(x')} \right\| \leqslant L(x'') - L(x') ,$$

alors, si $\mathbb{R}_1$ n'a pas de maximum, ou si $\vec{\mu}$ n'a pas de masse en ce maximum, la mesure $\vec{\mu}$ est absolument majorée par la mesure $\lambda \geqslant 0$.

On voit pourquoi il faut supposer M continue à gauche : le théorème 83 nous y oblige. Alors le lemme 2 du théorème 84 bis démontre l'égalité des 2 derniers membres de (IV,9;36).

Assez généralement, on représente par $\overrightarrow{dM}$, au lieu de $\vec{\mu}$ ou $d\vec{\mu}$, la mesure dont M est l'intégrale indéfinie; on dit aussi que cette mesure est la mesure dérivée de M.

Nous verrons en effet, au corollaire 1 du théorème 89, que si M est de classe C^1, cette mesure est $\vec{M}'(x)\,dx$

Démonstration -

Montrons d'abord que, si M est intégrale indéfinie d'une mesure $\vec{\mu}$ sur $\mathbb{R}_1$, à valeurs dans E, $\vec{\mu}$ est nécessairement unique, si $\mathbb{R}_1$ n'a pas de maximum.

Nous connaissons en effet $\vec{\mu}([a,b[)$, pour les intervalles semi-ouverts $[a,b[$ de $\mathbb{R}_1$. Mais, si $[a,b]$ est un intervalle fermé de $\mathbb{R}_1$, b n'étant pas un maximum de $\mathbb{R}_1$, $\vec{\mu}([a,b])$ est la limite des $\vec{\mu}([a,b'[)$, quand b' tend vers b par valeurs $> b$ (remarque suivant le théorème 83), donc $\vec{\mu}([a,b])$ est également connu. De même $\vec{\mu}(]a,b[)$ est connu comme limite de $\vec{\mu}([a',b[)$ quand a' tend vers a par valeurs $> a$; enfin $\vec{\mu}(]a,b])$ est limite de $\vec{\mu}([a',b])$ quand a' tend vers a par valeurs $> a$. Alors $\vec{\mu}(\varphi)$ est connue pour toute fonction en escalier à support compact φ, si $\mathbb{R}_1$ n'a pas de maximum. Comme alors toute fonction φ de $\mathcal{C}(\mathbb{R}_1)$ est limite uniforme de telles fonctions, gardant leur support dans un compact fixe de $\mathbb{R}_1$, $\vec{\mu}(\varphi)$ est connue pour toute fonction φ de $\mathcal{C}(\mathbb{R}_1)$, donc $\vec{\mu}$ est bien déterminée d'une manière unique. Ce raisonnement est évidemment en défaut si $\mathbb{R}_1$ a un maximum b_0. Alors si l'on ajoute à $\vec{\mu}$ n'importe quelle masse portée par b_0 on ne change évidemment pas ses intégrales indéfinies, dont la définition ne fait intervenir que les intervalles $[a,b[$ de $\mathbb{R}_1$, ouverts à droite. Mais au contraire, si $\vec{\mu}$ présente une masse $\vec{\mu}(\{b_0\}) \neq \vec{0}$ en b_0, on peut la retirer, c'est-à-dire retrancher de $\vec{\mu}$ la mesure $\vec{\mu}(\{b_0\})\,\delta_{(b_0)}$, et on obtient alors une nouvelle mesure, sans masse en b_0, dont M est intégrale indéfinie; une telle mesure est alors unique.

Montrons maintenant que toute fonction M à valeurs dans E de dimension finie, localement à variation bornée, est intégrale indéfinie d'une mesure.

Soit en effet $0, (\vec{e}_i)_{i\in I}$ un référentiel de E par rapport au corps $\mathbb{R}$. Alors $M = 0 + \sum_{i\in I} M_i \vec{e}_i$, et chacune des M_i est une fonction réelle continue à gauche et localement à variation bornée; si nous démontrons qu'elle est intégrale indéfinie d'une mesure réelle μ_i, alors M sera intégrale indéfinie de $\vec{\mu} = \sum_{i=1}^{n} \mu_i \vec{e}_i$. Donc on peut se borner aux fonctions réelles.

Mais, si M est une f'onction réelle, continue a gauche et localement à variation bornée, elle peut, d'après la méthode du théorème 86, s'exprimer comme différence de 2 fonctions croissantes, $M = V - (V - M)$. ta fonction V est continue à gauche puisque M est continue à gauche (théorème 84 bis), donc aussi $V - M$. Si chacune des fonctions V, $V - M$, est intégrale indéfinie d'une mesure $\geqslant 0$, M sera bien intégrale indéfinie d'une mesure réelle. Ainsi nous sommes ramenés à démontrer que toute fonction sur $\mathbb{R}_1$, réelle, croissante, continue à gauche, est intégrale indéfinie d'une mesure $\geqslant 0$.

Soit donc M une telle fonction. Si M est continue et strictement croissante, c'est évident, car alors M est un homéomorphisme $x \longrightarrow y = M(x)$ de $\mathbb{R}_1$ sur un intervalle $\mathbb{R}_2$ de $\mathbb{R}$ (théorème 37 du chapitre II). Soit M^{-1} son homéomorphisme réciproque, et appelons μ l'image par M^{-1} de la mesure de Lebesgue dy de $\mathbb{R}_2$. Si alors $[a,b[$ est un intervalle de $\mathbb{R}_1$, $[\alpha,\beta[= [M(a), M(b)[$ son image par M dans $\mathbb{R}_2$, le corollaire 1 du théorème 60 montre que

$$\mu([a,b[) = (M^{-1}(dy))([a,b[) = dy([M(a), M(b)[) = M(b) - M(a),$$

donc M est bien intégrale indéfinie de μ.

Soit maintenant M croissante (non nécessairement strictement) et non nécessairement continue. Soient a_i ses points de discontinuité, en infinité dénombrable, et soient c_i ses sauts; pour tout intervalle $[a,b]$ de $\mathbb{R}_1$, on a $\sum_{a_i \in [a,b[} c_i \leq M(b) - M(a)$. Considérons la mesure de Radon $\lambda = \sum_i c_i \delta_{(a_i)}$. Elle a une intégrale indéfinie L (théorème 82); L a exactement les a_i comme points de discontinuité, et les sauts c_i (théorème 83). Par ailleurs on a, d'après ce qui précède, $M(b) - M(a) \geqslant \sum_{a_i \in [a,b[} c_i = L(b) - L(a)$ donc $M-L$ est encore croissante, mais maintenant continue. Alors $M-L+\varepsilon x$ est strictement croissante et continue; elle est donc l'intégrale indéfinie d'une mesure de Radon μ_ε, et alors M est l'intégrale indéfinie de $\mu = \mu_\varepsilon + \lambda - \varepsilon\, dx$; μ n'est peut-être pas $\geqslant 0$, mais $\mu + \varepsilon\, dx \geqslant 0$; comme μ est indépendante du procédé utilisé (unicité démontrée antérieurement) donc de ε, on a aussi $\mu \geqslant 0$.

Supposons qu'il existe une fonction réelle croissante L sur $\mathbb{R}_1$ telle que l'on ait (IV,9;37). Soient $\vec{\mu}$ et λ des mesures, dont M et L sont des intégrales indéfinies. Pour tout intervalle semi-ouvert $[a,b[$, on a donc $\|\vec{\mu}([a,b[)\| \leqslant \lambda([a,b[)$. Par le même passage à la limite que précédemment, on voit que cette majoration reste vraie pour tout intervalle d'un autre type, si $\mathbb{R}_1$ n'a pas de maximum, ou si $\vec{\mu}$ n'a pas de masse en ce maximum. Alors, si φ est $\geqslant 0$, en escalier, à support compact sur $\mathbb{R}_1$, on a $\|\vec{\mu}(\varphi)\| \leqslant \lambda(\varphi)$; par passage à la limite, on a la même inégalité pour toute fonction $\varphi \geqslant 0$ de $\mathcal{C}(\mathbb{R}_1)$, donc $\vec{\mu}$ est absolument majorée par λ. Remarquons enfin que, si $\vec{\mu}$ ~~est~~ de base $\lambda \geqslant 0$, $\vec{\mu} = \vec{f}\lambda$, on a vu au théorème 54 que $|\vec{\mu}|$ n'est autre que la mesure $\|\vec{f}\| \, d\lambda$. Alors (IV,9;36) est équivalent à (IV,9;23).

La mesure $\vec{\mu}$ est de norme finie sur $\mathbb{R}_1$, si et seulement si $|\vec{\mu}|$ est de norme finie [On a toujours $\|\vec{\mu}\| \leqslant \| |\vec{\mu}| \|$; nous avons vu qu'on n'a pas en général égalité si E est de dimension $\geqslant 2$ (voir théorème 15). Il n'est donc pas immédiat que $\|\vec{\mu}\| < +\infty$ implique $\| |\vec{\mu}| \| < +\infty$ Mais cela résulte de ce que E est de dimension finie.

Si $(\vec{e}_i)_{i\in I}$ est une base de $\vec{E}$ sur le corps $\mathbb{R}$, $\vec{\mu} = \sum_{i\in I} \vec{e}_i \mu_i$, on sait que $\sum_{i\in I} \|\vec{e}_i\| \, |\mu_i|$ est une majorante absolue de $\vec{\mu}$ (page 456); si $\vec{\mu}$ est de norme finie, il en est de même des μ_i, donc des $|\mu_i|$ qui ont la même mesure (théorème 18); et si une majorante absolue de μ a une norme finie, il en est de même a fortiori de sa plus petite majorante absolue], c'est-à-dire si et seulement si $\int_{\mathbb{R}_1} d|\vec{\mu}| < +\infty$, ou encore si et seulement si M est à variation bornée sur $\mathbb{R}_1$ tout entier.

<u>Corollaire - Pour qu'une fonction M sur un intervalle $\mathbb{R}_1$ de $\mathbb{R}$, à valeurs dans un espace affine normé F de dimension finie, soit lipschitzienne de coefficient de Lipschitz k il faut et il suffit qu'elle soit intégrale indéfinie, par rapport à la mesure de Lebesgue, d'une fonction $\vec{\mu}$ sur $\mathbb{R}_1$, à valeurs dans $\vec{F}$, de norme partout majorée par k.</u>

Démonstration

Si, pour $x' \leq x''$, $\overrightarrow{M(x'') - M(x')} = \int_{x'}^{x''} \vec{f}(x)\,dx$, $\|\vec{f}(x)\| \leq k$,

on a bien $\|\overrightarrow{M(x'') - M(x')}\| \leq k(x'' - x')$.

Inversement, si on a cette inégalité, on a (IV,9;37), avec $L(x) = k\,x$; donc M est intégrale indéfinie d'une mesure $\vec{\mu}$, qu'on peut supposer sans masse au maximum de $\mathbb{R}_1$ si ce maximum existe; et alors $\vec{\mu}$ est absolument majorée par $k\,dx$. Alors le théorème 53 de Dunford Pettis montre que $d\vec{\mu} = \vec{f}\,dx$, avec $\|\vec{f}\| \leq k$, sur $\mathbb{R}_1$.

Longueur d'un chemin dans un espace métrique

Considérons un chemin dans un espace métrique E ; d'après la définition de la page 90, c'est une application continue M d'un segment $[a,b]$ de $\mathbb{R}$ dans l'espace métrique E.

On appelle longueur de ce chemin la variation totale $V(M)$ de M, $0 \leq V(M) \leq +\infty$. Le chemin est de longueur finie, si et seulement si M est à variation bornée; si E est un espace affine de dimension finie, muni d'un référentiel, il en est ainsi, si et seulement si les composantes de M sont à variation bornée.

Supposons en particulier que E soit un espace affine normé. Soit d'autre part * $t \longmapsto A + t\,\overrightarrow{B - A}$ un chemin "rectiligne", d'extrêmités A et B correspondant à $t = 0$ et $t = 1$ respectivement. Pour toute décomposition Δ de $[0,1]$, on a : $\|\overrightarrow{M(c_{i+1}) - M(c_i)}\| = (c_{i+1} - c_i)\,\|\overrightarrow{B - A}\|$, d'où $\sum_\Delta = \|\overrightarrow{B - A}\|$, et par suite la longueur du chemin rectiligne est $\|\overrightarrow{B - A}\|$, comme on pouvait s'y attendre. Si maintenant on considère un chemin quelconque M, si $c_0 = a, c_1, c_2, \ldots, c_n = b$ est une décomposition Δ de $[a,b]$, et si nous considérons la quantité $\sum_\Delta$ définie à (IV,9;6), correspondant à ce chemin, c'est tout simplement

* Dans l'étude des chemins et de leurs longueurs, nous emploierons plutôt t ou u au lieu de x, pour la variable; c est ce qu'on fait ordinairement en géométrie pour la représentation paramétrique d une courbe .

la longueur de la ligne polygonale inscrite dans le chemin, et dont les sommets successifs sont $M(c_0), M(c_1), \dots, M(c_n)$. La variation totale ou longueur du chemin apparaît alors simplement comme borne supérieure des longueurs des lignes polygonales inscrites dans le chemin.

(Si E est un espace métrique quelconque, nous continuerons parfois à employer cette image de borne supérieure des longueurs des lignes polygonales inscrites, alors qu'il n'y a pas à proprement parler de lignes polygonales, et qu'on fait intervenir la distance $d(M(c_i), M(c_{i+1}))$ sans tracer un segment de droite qui les joigne, sans même qu'il soit nécessairement possible de le faire).

Considérons deux chemins équivalents * de E ; autrement dit, nous avons deux applications M et N de deux segments $[a, b]$ et $[\alpha, \beta]$ de $\mathbb{R}$, dans E, et nous supposons qu'il existe un homéomorphisme ξ de $[\alpha, \beta]$ sur $[a, b]$, tel que l'on ait $N = M \circ \xi$. Dans ce cas, il est bien évident que ces deux chemins ont la même longueur.

La fonction longueur, qui, à chaque chemin, associe sa longueur, est alors une fonction définie sur tous les chemins de E, à valeurs dans $[0, +\infty]$, et possedant les deux propriétés évidentes suivantes :

1°/ Quel que soit le chemin d'origine A et d'extrémité B dans E, sa longueur est au moins égale à $d(A, B)$, correspondant à la décomposition triviale $c_0 = a$, $c_1 = b$; si E est un espace affine normé, ceci montre que la ligne droite est le plus court chemin d'un point à un autre ** ;

2°/ Si un chemin est la Juxtaposition de deux chemins, sa longueur est égale à la somme de leurs longueurs *** .

Qu'entendons-nous exactement par là ?

* Dans le même sens que les variétés paramétriques équivalentes, voir page 335.

** Mais elle n'est pas forcément seule à réaliser le plus court chemin; on peut donner, dans certains espaces affines normés, un exemple d'un système de 3 points A, B, C, non alignés, tels que $\|\overrightarrow{C-A}\| = \|\overrightarrow{B-A}\| + \|\overrightarrow{C-B}\|$.

• +* Bien faire attention à ce caractère additif. Si, par exemple, quand t décrit $[a, b]$, M(t) décrit d'abord le segment $[A, C]$, puis le segment $[C, A]$, de sorte que $M(a) = M(b) = A$, le longueur est, bien entendu, $2\|\overrightarrow{C-A}\|$, et non pas 0 .

Soit M_1 un chemin, application continue d'un segment $[a,b]$, $a < b$, dans E ; soit M_2 un deuxième chemin, application d'un segment $[b,c]$, $b < c$, dans E , et supposons que $M_1(b) = M_2(b) = B$.

Alors M_1 et M_2 définissent évidemment une application continue M de $[a,c]$ dans E, en posant

$$M(t) = \begin{cases} M_1(t) \text{ pour } a \leqslant t < b ; \\ M_2(t) \text{ pour } b \leqslant t \leqslant c ; \end{cases} \tag{IV,9;37 bis}$$

et celle-ci définit un chemin qu'on appelle juxtaposition des deux premiers chemins. L'affirmation 2°/ résulte alors de (IV,9;6 ter).

<u>La fonction longueur est la plus petite fonction définie sur les chemins de E , à valeurs $\geqslant 0$, et possédant les 2 propriétés précédentes.</u>

Si en effet une fonction possède les propriétés précédentes, alors la longueur d'un chemin est certainement $\geqslant$ la longueur de toutes les lignes polygonales inscrites, donc à leur borne supérieure, c'est-à-dire précisément à la longueur du chemin, telle que nous venons de la définir.

Le théorème 85 montre que la longueur d'un chemin peut s'obtenir comme limite des longueurs d'une suite de lignes polygonales inscrites, et non nécessairement comme borne supérieure.

Soient $\mathbb{R}_1$ un intervalle non nécessairement fermé de $\mathbb{R}$, M une fonction continue sur $\mathbb{R}_1$ à valeurs dans l'espace métrique E . On dit qu'elle définit un chemin impropre. On dit que ce chemin impropre est <u>rectifiable</u>, si, pour tout intervalle compact $[a,b]$ de $\mathbb{R}_1$, la restriction $M_{[a,b]}$ de M à $[a,b]$ * définit un chemin de longueur finie (le chemin impropre lui-même pouvant être de longueur infinie; cela signifie simplement que M est localement à variation bornée).

* A la page 10 , nous avons introduit f_A pour la restriction d'une fonction f sur E à une partie A de E. On emploie aussi la notation $f \mid A$. Alors l'arc ci-dessus est le chemin défini par l'application restreinte $M_{[a,b]}$ ou $M \mid [a,b]$.

soit alors c un point de $\mathbb{R}_1$. A tout point t de $\mathbb{R}_1$, on pourra faire correspondre le nombre réel $s(t)$ défini comme suit : si $t \geq c$, $d(t)$ est la longueur de l'arc $M_{[c,t]}$; si $t < c$, $s(t)$ est l'opposé de la longueur de l'arc $M_{[t,c]}$. On définit ainsi une fonction $s : t \longrightarrow s(t)$ sur $\mathbb{R}_1$, à valeurs réelles (déjà définie à la formule (IV,9;20)), continue (théorème 84 bis), croissante et en général strictement croissante; $s(t)$ est appelée l'abscisse curviligne relative à $t \in \mathbb{R}_1$, lorsqu'on prend C comme origine des abscisses *. Si la fonction s est strictement croissante, elle définit un homéomorphisme de $\mathbb{R}_1$ sur un intervalle $\mathbb{R}_2$ de la droite réelle (théorème 37 du chapitre II); on peut alors l'inverser par une fonction $s \longrightarrow t(s)$, on peut donc prendre s comme paramètre, et définir ainsi une nouvelle paramétrisation du chemin impropre, équivalente à la paramétrisation initiale: les différentes paramétrisations obtenues en faisant varier C, ou en changeant les signes + et - dans la définition de s, sont toutes équivalentes.

On les appelle paramétrisations intrinsèques (elles dépendent évidemment de la métrique de E, et non seulement de sa topologie). Si s n'est pas strictement croissante, elle ne définit pas une paramétrisation équivalente du chemin. Toutefois, si& est donné, et s'il existe une infinité de valeurs de t pour lesquelles $s(t) = s_0$, elles forment un Intervalle de $\mathbb{R}_1$ sur lequel M est constante, de sorte que les M(t) correspondants sont tous confondus avec un même point P_0 de E ; si donc on pose $P_0 = M_0(s_0)$ on voit qu'il existe une fonction M_0 sur l'intervalle $\mathbb{R}_2$ de $\mathbb{R}$, telle que $M(t)= M_0(s(t))$; M_0 est continue et même lipschitzienne, car l'image par M_0 de l'intervalle $[s',s'']$ a la longueur $|s''-s'|$, donc on a, a fortiori :

(IV,9;37^ter) $$\left\| \overrightarrow{M_0(s'') - M_0(s')} \right\| \leq |s'' - s'|$$

* On ne peut pas, en général, l'appeler **abscisse** curviligne du point $P = M(t)$, car un même point P de E peut être l'image par la fonction M de plusieurs points t de $[a,b]$ ("points multiples'< du chemin).

Elle ne définit pas un chemin équivalent, puisque s n'est pas un homéomorphisme de $\mathbb{R}_1$ sur $\mathbb{R}_2$, n'étant pas injective. On voit bien que la considération de la paramétrisation M_0 reste "intéressante" (et même plus intéressante en pratique que la paramétrisation M), et on continue à l'appeler paramétrisation intrinsèque du chemin impropre . (Avec l'abus de langage, déjà signalé page 11, qui consiste à donner à la fonction le même nom, même après chargement de variables, on écrit souvent $M(s)$ au lieu de $M_0(s)$, en disant que c'est "la même fonction exprimée a l'aide de la variable s au lieu de la variable t" (!)).

La fonction croissante s sur $\mathbb{R}_1$ est l'intégrale indéfinie d'une mesure de Radon $\geqslant 0$ sur $\mathbb{R}_1$ (théorème 88), par rapport à laquelle tout intervalle $[t', t'']$ de $[a, b]$ a pour mesure la longueur $|s''- s'|$ du chemin $M_{[t',t'']}$. On appelle généralement ds cette mesure sur $\mathbb{R}_1$ (quand on passe à la représentation paramétrique intrinsèque, elle devient bien la mesure de Lebesgue ds). Noter que ds est une mesure sur $\mathbb{R}_1$ et non sur l'image du chemin. On l'appelle <u>mesure des lon-Lueurs</u> sur le chemin.

Il resulte de la définition même de s (Formule (17, 9;20)) et du théorème 88 (formule (IV,9;36)) que ds est la plus petite majorante absolue de $\overrightarrow{dM}$, **si E** est **un** espace affine normé de dimension finie. Le théorème 54 de Dunford-Pettis montre alors qu'on peut écrire

$\overrightarrow{dM} = \vec{\mu}\, ds$, $\|\vec{\mu}\| = 1$ ds -presque partout.

'il se trouve que E soit un espace localement compact dénombrable à l'infini et si l'application M est& propre, il existe aussi une image de la mesure précédente, mesure de Radon $\geqslant 0$ sur E , dont le support est contenu dans l'adhérence de l'image du chemin. Souvent malgré la confusion que cela peut apporter dans les notations, on désigne aussi Par ds cette mesure de Radon image, au lieu de $M(ds)$.

Si $\vec{f}$ est une 'onction définie sur E, à valeurs dans un espace de Banach $\vec{F}$, on pourra parler de <u>l'intégrale de courbe</u> $\int \vec{f}\, ds$; par définition, c'est $\int_{\mathbb{R}_1} \vec{f}(M(t))\, d(s(t))$; c'est aussi, d'après le théorème OC, l'intégrale de $\vec{f}$ par rapport à la mesure image $M(ds)$, quand celle-ci existe * .

* La notation $\int \vec{f}\, ds$ est insuffisante, car le chemin M n'y apparait pas. On pourra écrire $\int_{(M)} \vec{f}\, ds$, intégrale relativement au chemin défini par l'application M.

Théorème 88 bis - L'image d'un chemin impropre rectifiable dans un espace $\mathbb{R}^n$ a nécessairement une mesure nulle par rapport à $dx = dx_1 \otimes dx_2 \otimes \ldots \otimes dx_n$

Démonstration - Le chemin étant rectifiable, on peut, pour démontrer ce théorème, le remplacer par le chemin $M_0 : s \longrightarrow M_0(s)$, correspondant à une paramétrisation intrinsèque puisque les image s de M et de M_0 sont identiques. Soit alors $[a,b]$ un intervalle de $\mathbb{R}_2$.

Soit Δ une décomposition de $[a,b]$, $c_0 = a, c_1, c_2, \ldots, c_\ell = b$, de finesse $\leq \eta$. soit c'_i le milieu de $[c_i, c_{i+1}]$. Comme tout arc a une longueur au moins égale à la corde qui joint ses extrémités, les distances des points $M(t), t \in [c_i, c_{i+1}]$, à $M(c'_i)$, sont majorées par les longueurs des arcs $M_{[c'_i, t]}$, donc par $\frac{1}{2}(c_{i+1} - c_i)$; alors l'image du chemin $M_{[c_i, c_{i+1}]}$ est contenue dans un cube de centre $M(c'_i)$, de côtés parallèles aux axes, égaux à $(c_{i+1} - c_i)$. La mesure volumétrique de cette image est donc majorée par $(c_{i+1} - c_i)^n \leq \eta^{n-1}(c_{i+1} - c_i)$. La mesure volumétrique de l'image de $[a,b]$ est donc majorée par $\eta^{n-1}(b-a)$. Comme η est arbitraire, le volume de cette image est nul. Comme $\mathbb{R}_2$ est réunion dénombrable d'intervalles tels que $[a,b]$, son image est bien encore de volume nul.

Remarque. Il ne faudrait pas croire que cette conclusion subsiste pour un chemin -rectifiable. voir à ce sujet la remarque ultérieure qui suit le corollaire 1 du théorème 102 bis

Intégrale indéfinie et primitive

Théorème 89 - Soit $\vec{f}$ une fonction définie sur un intervalle $\mathbb{R}_1$ de la droite réelle $\mathbb{R}$, à valeurs dans un espace de Banach $\vec{F}$, et localement int rable pour la mesure de Lebesgue. Alors une intégra% indéfinie F de $\vec{f}$ par rapport à dx, à valeurs dans un espace affine normé F, d'espace vectoriel associé $\vec{F}$, admet $\vec{f}(c)$ comme dérivée (resp. comme dérivée a gauche, resp. comme dérivée à droite), en tout point c où est continue (resp. continue a gauche, resp. continue à d oite).

Démonstration - Montrons par exemple la dérivabilité de F si $\vec{f}$ est continue.

On a

$$\overline{\frac{F(c+h)-F(c)}{h}} - \vec{f}(c) = \frac{1}{h}\int_c^{c+h} (\vec{f}(x)-\vec{f}(c))\, dx , \tag{IV,9;38}$$

d'où la majoration

$$\left\| \overline{\frac{F(c+h)-F(c)}{h}} - \vec{f}(c) \right\| \leqslant \sup_{|x-c|\leqslant h} \left\| \vec{f}(x)-\vec{f}(c) \right\| , \tag{IV,9;38bis}$$

et en vertu de l'hypothèse de continuité le deuxième membre tend vers 0 lorsque h tend vers 0 , ce qui démontre le théorème.

<u>Corollaire 1 - Si $\vec{f}$ est une fonction continue sur $\mathbb{R}_1$, à valeur dans $\vec{E}$, il y a identité entre les intégrales indéfinies de $\vec{f}$ par rapport à dx, à valeurs dans E, et les primitives de $\vec{f}$, à valeurs dans E, c'est-à-dire les fonc - tions ayant $\vec{f}$ comme fonction dérivée. Toute fonction sur $\mathbb{R}_1$ à valeurs dans E , et de classe C^1 , est intégrale indéfinie de sa fonction dérivée , par rapport à la mesure de Lebesgue dx.</u>

<u>Démonstration.</u> Le théorème montre qu'une intégrale indéfinie quelconque est une primitive , Mais la différence entre 2 primitives est une constante de $\vec{E}$ (théorème 22 du chapitre III) donc toute primitive diffère d'une intégrale indéfinie d'une constante, donc est aussi une intégrale indéfinie (théorème 82).

<u>Remarques</u> 1°/ On <u>sait qu'une origine du calcul intégral a été la recherche des primitives.</u> <u>C'est depuis longtemps l'inverse qui est vrai : ce sont les primitives qui servent à calculer les intégrales; si F est une primitive de la fonction continue $\vec{f}$, on a :</u>

$$\overrightarrow{F(b) - F(a)} = \int_a^b \vec{f}(x)\, dx . \tag{IV,9;38ter}$$

2°/ Considérons la fonction $\vec{F}$ de 2 variables : $(x,y) \longmapsto \int_x^y \vec{f}(t)\, dt$, et supposons $\vec{f}$ partout continue.

Alors F admet évidemment une dérivée totale, donnée, en notation différentielle, par

$$(\mathrm{IV},9;39) \quad \overrightarrow{dF} = \left(F'(x,y)\right)\cdot d(x,y) = \vec{f}(y)\,dy - \vec{f}(x)\,dx\,.$$

En effet, on a, d'après la relation de Chasles :

$$(\mathrm{IV},9;40) \quad \vec{F}(x,y) = \int_a^y \vec{f}(t)\,dt - \int_a^x \vec{f}(t)\,dt\,,$$

et ces 2 intégrales, l'une fonction de x seul et l'autre de y seul, ont pour dérivées $\vec{f}(y)$ et $\vec{f}(x)$.

3°/ Une fonction dérivable, mais à dérivée non partout continue, n'est pas nécessairement intégrale indéfinie de sa dérivée, qui d'ailleurs n'est pas nécessairement localement intégrable pour la mesure de Lebesgue.

4°/ Le corollaire relie d'ailleurs deux notions assez différentes. La primitive est liée à la dérivée. La recherche de la primitive se généralise comme suit : soient E et F des espaces affines normés, Ω un ouvert de E ; soit $\vec{f}$ une fonction continue sur Ω, à valeurs dans $\mathcal{L}(\vec{E};\vec{F})$: trouver une fonction sur Ω, à valeurs dans F, qui admette $\vec{f}$ comme fonction dérivée. L'intégrale indéfinie est liée à la notion de mesure de Radon; on peut parler des intégrales indéfinies d'une mesure $\vec{\mu}$, ou d'une fonction $\vec{f}$ par rapport à une mesure $\lambda \geq 0$. On compare ici les primitives de $\vec{f}$, et les intégrales indéfinies de $\vec{f}$ par rapport à dx.

<u>Corollaire 2</u> - Soit M <u>une application de classe C^1 d'un intervalle $\mathbb{R}_1$ de $\mathbb{R}$ dans un espace affine norme E. Elle définit un chemin rectifiable, et, pour tout intervalle $[a,b]$ de $\mathbb{R}_1$, $a<b$, la longueur de 1 arc</u> $M_{[a,b]}$ <u>est</u>

$$(\mathrm{IV},9;40\,bis) \quad s(b) - s(a) = \int_{[a,b]} \left\|\overrightarrow{M'(t)}\right\| dt$$

En effet, d'après le corollaire 1, M est une intégrale indéfinie de $\overrightarrow{M'(t)}\,dt$; donc elle est localement à variation bornée (théorème 87), donc définit un chemin rectifiable d'après la définition; la longueur du chemin $M_{|a,b|}$, ou variation totale de M sur $|a,b|$ est alors (IV,9;40 bis) d'après (IV,9;23 ou 35).

Théorème 89 bis - Si la fonction M, définie sur $\mathbb{R}_1$, à valeurs dans l'espace affine normé E, est de classe C^1, la fonction abscisse curviligne s est aussi de classe C^1.

En tout point t_0 où $\overrightarrow{M'(t_0)} \neq \vec{0}$, le chemin possède une tangente, et un vecteur unitaire de cette tangente est donné par $\dfrac{\overrightarrow{M'(t_0)}}{\|\overrightarrow{M'(t_0)}\|}$. Au voisinage de t_0, M_0 est une fonction de classe C^1 de l'abscisse curviligne s, et on a :

(IV,9;40 ter) $$\frac{\overrightarrow{M'(t_0)}}{\|\overrightarrow{M'(t_0)}\|} = \left(\frac{\overrightarrow{dM_0}}{ds}\right)_{s=s_0},$$

vecteur unitaire de la tangente.

En outre si deux points t', t'' tendent vers le point t_0, l'arc joignant les deux points $M(t')$, $M(t'')$, a une longueur équivalente à celle de la corde qui le sous-tend.

Démonstration - Le fait que l'arc A soit une fonction de t de classe C^1 résulte tout simplement du théorème 89, car s est, d'après le corollaire 2 de ce même théorème, une intégrale indéfinie de la fonction continue $t \to \|\overrightarrow{M'(t)}\|$. Si alors, en un point t_0, cette dérivée n'est pas nulle, il résulte du théorème des fonctions implicites (théorème 29 du chapitre III) qu'il est possible, au voisinage de t_0, de calculer t comme fonction de s de classe C^1, et que sa dérivée au point t_0 est donnée par

(IV,9;40 quarto) $$\left(\frac{dt}{ds}\right)_{s=s_0} = \frac{1}{\left(\frac{ds}{dt}\right)_{t=t_0}} = \frac{1}{\|\overrightarrow{M'(t_0)}\|}.$$

11 en résulte que la fonction $M_0(s)$, qui peut toujours se calculer **sans** qu'aucune hypothèse soit **nécessaire** sur la **dérivabilité,** est **dérivable** par rapport **à** s , aux polntsb voisins de $s(t_0)$, et que sa **dérivée** est bien donnée par la formule

(IV,9;40quinto) $$\left(\frac{d\overrightarrow{M_0}}{ds}\right)_{s=s_0} = \frac{d\overrightarrow{M}}{dt}\frac{dt}{ds} = \frac{\overrightarrow{M'(t_0)}}{\|\overrightarrow{M'(t_0)}\|} .$$

Z

Dans l'énoncé, nous avons parlé de tangente au chemin en un point. ~~Ce n'est pas du tout la tangente à l'ensemble image du chemin, au sens de la page 202. Car un même point tel que $M(t_0)$ est aussi peut être l'image par~~ M d'une infinité d'autres points $t \neq t_0$, **correspondant** à diverses "branches de courbe" passant. en ce même point, et la seule considération de la dérivée de M en t_0 ne peut donner aucuns indication sur la tangente en ce point. Nous appelons tangente au chemin au point t_0 la limite, **si** elle existe, de la droite joignant $M(t_0)$ à $M(t_0+\Delta t)$, lorsque $\Delta t \neq 0$ tend vers 0 . S'il y a une dérivée $\neq \vec{0}$, on peut affirmer qu'une telle tangente existe, et que $\overrightarrow{M'(t_0)}$ est un vecteur de cette tangente •

Comme alors Ici $\frac{d\overrightarrow{M_0}}{ds}$ est un vecteur de norme 1, il définit un vecteur unitaire de la tangente. On dit aussi que ce vecteur est le vecteur unitaire de la demi-tangente correspondant "**au** sens des t croissants" parce que cette demi-tangente est la limite de la demi-droite d'origine M(t), passant par $M(t+\Delta t)$, lorsque Δt tend vers 0 par valeurs $>$ 0 .

Soient alors t', t'' deux nombres tendant vers t_0 ; les valeurs correspondantes s', s'' , de l'abscisse **curviligne, tendent** vers s_0 D'après la formule des accroissements finis (corollaire'1 du théorème 13 du chapitre III), on peut écrire:

(IV,9;40sexto) $$\overrightarrow{M_0(s'') - M_0(s')} = \overrightarrow{M_0'(s_0)}\,(s''-s') + \vec{\omega}\,|s''-s'| ,$$

où $$\|\vec{\omega}\| \leq \sup_{\theta\in[s',s'']} \|\overrightarrow{M_0'(\theta)} - \overrightarrow{M_0'(s_0)}\| .$$

• De la même manière on peut définir le contingent en un point d'une variété paramétrique de classe C' , qui n'a rien à voir avec le contingent de l'image de la **variété** paramétrique.

On en déduit, puisque $\| \overrightarrow{M'(s_0)} \| = 1$:

(IV, 9;40 septimo) $$\left| \| \overrightarrow{M_0(s'') - M_0(s')} \| - |s'' - s'| \right|$$

$$\leq \| \overrightarrow{M_0(s'') - M_0(s')} - \overrightarrow{M_0'}(s_0)(s'' - s') \|$$

$$\leq \| \vec{\omega} \| \; |s'' - s'| .$$

La continuité de la dérivée $\frac{d\overrightarrow{M_0}}{ds}$ nous assure que $\|\vec{\omega}\|$ tend vers 0 quand A' et s'' tendent vers s_0, ce qui prouve bien l'équivalence de la longueur de l'arc, $|s'' - s'|$, et de celle de la corde, $\| \overrightarrow{M_0(s'') - M_0(s')} \|$.

Naturellement ce théorème 89 donne une condition suffisante pour qu'une intégrale indéfinie F de $\vec{f}$ soit dérivable; F peut parfaitement avoir des dérivées en certains points où $\vec{f}$ est discontinue. D'ailleurs il peut arriver que la fonction $\vec{f}$, localement intégrable pour la mesure dx, soit discontinue en tous les points de la droite, comme le montre l'exemple de la fonction définie à la page 482 ; cependant la fonction F admet comme dérivée $\vec{f}(x)$ nécessairement pour certaines valeurs de x, comme le montre le théorème suivant, que nous admettrons, car sa démonstration est difficile, et qui est dû à Lebesgue :

Théorème 90 - Si $\vec{f}$ est une fonction définie sur un intervalle $\mathbb{R}_1$ de $\mathbb{R}$, à valeurs dans un espace de Banach $\vec{E}$, et localement intégrable pour la mesure de Lebesgue dx, alors toute intégrale indéfinie F de $\vec{f}$ par rapport à dx admet $\vec{f}(x)$ comme dérivée, pour dx-presque toutes les valeurs de x.

Naturellement, on ne pouvait pas s'attendre à mieux que dx-presque partout, puisque, si on change $\vec{f}$ sur un ensemble de mesure nulle, on ne change pas F.

Corollaire 1 - Toute fonction F sur $\mathbb{R}_1$, à valeurs dans un espace affine normé E, ayant partout une dérivée bornée en norme par k, est lipschitzienne, de coefficient de Lipschitz k. Toute fonction F sur $\mathbb{R}_1$, à valeurs dans E de dimension finie, lipschitzienne, de coefficient de Lipschitz k, admet dz-presque partout une dérivée, de norme $\leq k$.

La première partie est le corollaire 2 du théorème 13 du chapitre III. La 2ème partie résulte du théorème 90, et du corollaire du théorème 88.

Remarque - Les 2 parties ne sont pas des réciproques exactes l'une de l'autre. Si on sait qu'une fonction F est lipschitzienne, on ne peut pas en déduire qu'elle admette partout une dérivée (exemple : $x \longrightarrow |x|$ est lipschitzienne, de coefficient de Lipschitz 1, mais n'a pas de dérivée à l'origine); si on sait seulement qu'une fonction F a dx-presque partout une dérivée de norme $\leq k$, elle n'est pas nécessairement lipschitzienne, ni même continue (la fonction d'Heaviside (formule (IV,9;5)) a dx-presque partout une dérivée nulle).

Z

Corollaire 2 - Soit $t \longrightarrow M(t)$ un chemin rectifiable d'un espace affine normé E de dimension finie, ds la mesure de Radon dont l'abscisse curviligne s est intégrale indéfinie. Alors le chemin a ds-presque partout une tangente.

Ici encore nous disons bien "tangente au chemin" et non "tangente à l'image du chemin".

Démonstration - Considérons la représentation paramétrique intrinsèque de la courbe, $s \longrightarrow M_0(s)$ et nous avons

page 622 que $d\overrightarrow{M_0} = \vec{\tau}\, ds$, $\|\vec{\tau}\| = 1$ ds-presque partout. Mais alors le théorème de Lebesgue 90 dit que M_0 a ds-presque partout pour dérivée $\vec{\tau}$; et nous avons justement vu page 627 qu'il y a une tangente au chemin en tout point où il y a dérivée $\neq \vec{0}$; donc le chemin a ds-presque partout une tangente, de vecteur unitaire $\vec{\tau}$.

Primitives successives d'une fonction continue sur la droite

Soit, comme précédemment, $\mathbb{R}_1$ un intervalle de la droite $\mathbb{R}$. Soit $\vec{f}$ une fonction continue sur $\mathbb{R}$, à valeursdans un espace de Banach $\vec{E}$. Nous avons vu comment l'intégration nous permet de trouver une primitive de $\vec{f}$.

Nous allons voir que les primitives successives de $\vec{f}$, qui nécessitent a priori plusieurs intégrations successives, peuvent toujours se calculer par une seule intégration.

Théorème 91 - Si $\vec{f}$ est une application continue d'un intervalle $\mathbb{R}_1$ de $\mathbb{R}$ dans un espace de Banach $\vec{E}$, il existe une primitivem-ième $\vec{F}_m$ de $\vec{f}$ et une seule, qui s'annule pour un l ordre = c de $\mathbb{R}_1$, ainsi que ses dérivées jusqu'à $m-1$ inclusivement, et cette primitive est donnée P ' tégrale:

$$(\text{IV},9;41) \qquad \vec{F}_m(x) = \int \frac{(x-\xi)^{m-1}}{(m-1)!} \vec{f}(\xi)\, d\xi$$

Démonstration - L'unicité d'une primitive d'ordre m satisfaisant aux conditions données est évidente. En effet, d'après le théorème 22 du chapitre III, la différence entre deux primitives d'ordre m est nécessairement un polynôme de degré $\leq m-1$; mais, si l'on impose à ces primitives d'être nulles ainsi que leurs dérivées jusqu'à l'ordre $m-1$ au point c , ce polynôme devra être nul ainsi que ses dérivées jusqu'à l'ordre m -1 au point c et par conséquent identiquement nul, d'après la formule'de Taylor pour les polynômes. Il nous faut donc seulement montrer que la formule (IV,9;41) définit effectivement une primitive d'ordrem répondant à la question.

Le théoréme est évidemment vrai pour $m = 1$; supposons-le donc vrai relativement aux primitives d'ordrem -1 ,et ddmontrons-le relativement aux primitives d'ordre $m \geq 2$. Autrement dit, nous supposons que la fonction:

* Cette formule demande la plus grande attention :

ξ est la variable muette d'intégration, et x , qui n'est pas une variable muette, figure à la fois dans $\int^x$ et sous le signe d'intégration !

(IV,9;42) $$\vec{F}_{m-1}(x) = \int_c^x \frac{(x-\xi)^{m-2}}{(m-2)!} \vec{f}(\xi)\,d\xi ,$$

est une primitive d'ordre $m-1$ de $\vec{f}$, s'annulant au point c ainsi que ses dérivées d'ordre $\leqslant m-2$.

Alors il existe une primitive F_m d'ordre m de f, s'annulant au point c ainsi que ses dérivées jusqu'à l'ordre $m-1$, et elle est donnée par la formule :

(IV,9;43) $$\vec{F}_m(x) = \int_c^x \vec{F}_{m-1}(t)\,dt = \int_c^x dt \int_c^t \frac{(t-\xi)^{m-2}}{(m-2)!} \vec{f}(\xi)\,d\xi .$$

Supposons d'abord $x > c$ Alors il est permis de remplacer la succession de deux intégrations précédentes par l'intégrale unique:

(IV,9;44) $$\iint_A \frac{(t-\xi)^{m-2}}{(m-2)!} \vec{f}(\xi)\,dt\,d\xi ,$$

en appelant A l'ensemble du plan $\mathbb{R}^2$ défini par $c \leqslant \xi \leqslant t \leqslant x$ *,

$x \in \mathbb{R}_1$ En effet, cette dernière intégrale a un sens, puisqu'il s'agit de l'intégrale d'une fonction continue sur un compact, et alors le théorème de Fubini montre bien que (IV,9;43) est égal à (IV,9;44).

Mais alors nous pouvons calculer cette intégrale double en commançant par une intégration en t, et en suivant par une intégration en ξ. Autrement dit nous effectuons sur (IV,9;43) une interversion de l'ordre des intégrations. Compte tenu encore de la méthode indiquée dans la formule (IV,8;35), on obtient:

(IV,9;45) $$F_m(x) = \int_c^x \vec{f}(\xi)\,d\xi \int_\xi^x \frac{(t-\xi)^{m-2}}{(m-2)!}\,dt ,$$

et comme la dernière Intégrale vaut $\frac{(x-\xi)^{m-1}}{(m-1)!}$, et qu'un raisonnement identique pour $x < c$ donne la même formule, ceci démontre le résultat.

* On peut indifféremment mettre $\leqslant$ ou $<$, cela ne change la fonction à intégrer que sur un ensemble de mesure nulle.

Corollaire 1 - Soit $\vec{f}$ une fonction vérifiant les conditions du théorème. Alors il existe une primitive F_m et une seule d'ordre m de $\vec{f}$ dans $\mathbb{R}_1$, à valeurs dans l'espace affine E d'espace vectoriel associé $\vec{E}$, dont les dérivées jusqu'à l'ordre $m-1$ inclusivement prennent des valeurs-données $F_m(c), \vec{F}'_m(c), \ldots, \overrightarrow{F_m^{(m-1)}}(c)$, au point c. Elle est donnée par la formule :

$$(IV,9;46) \qquad F_m(x) = F_m(c) + \frac{x-c}{1!}\overrightarrow{F'_m}(c) + \cdots + \frac{(x-c)^{m-1}}{(m-1)!}\overrightarrow{F_m^{(m-1)}}(c) + \int_c^x \frac{(x-\xi)^{m-1}}{(m-1)!}\vec{f}(\xi)\,d\xi .$$

Démonstration - Comme dans le cas du théorème, l'unicité est évidente. Si alors nous remarquons que la fonction F_m définie par (IV,9:46) est somme d une primitive d'ordre m dont les dérivées jusqu'à l'ordre $m-1$ au point c sont nulles, et d'un polynôme, donné par son développement de Taylor, et dont les dérivées jusqu'à l'ordre $m-1$ au point c prennent les valeurs données, le résultat est évident.

Corollaire 2 - Nouvelle formule de Taylor pour les fonctions d'une variable réelle.

Soit f une fonction définie sur un intervalle $\mathbb{R}_1$ de $\mathbb{R}$, à valeurs dans un espace affine normé F, ayant des dérivées d'ordre $\leq m+1$ continues, alors on a la formule:

$$(IV,9;47) \qquad f(x) = f(c) + \frac{x-c}{1!}\overrightarrow{f'(c)} + \cdots + \frac{(x-c)^m}{m!}\overrightarrow{f^{(m)}}(c) + \int_c^x \frac{(x-\xi)^m}{m!}\overrightarrow{f^{(m+1)}}(\xi)\,d\xi,$$

et aussi la formule :

$$(IV,9;48) \qquad f(x+h) = f(x) + \frac{h}{1}\vec{f}'(x) + \cdots + \frac{h^m}{m!}\overrightarrow{f^{(m)}}(x) + h^{m+1}\int_0^1 \frac{(1-t)^m}{m!}\vec{f}^{(m+1)}(x+th)\,dt .^{(*)}$$

* A priori, il faudrait supposer F complet, comme nous l'avons fait dans le théorème, pour être sûrs que l'intégrale existe. Mais ici ce n'est pas nécessaire, elle existe sûrement, puisque f est donnée.

Démonstration - La première ormule n'est autre que la formule (IV,9;46), dans laquelle on a remplacé $\vec{f}$, $\vec{F}_m$ et m par $\vec{f}^{(m+1)}$, $\vec{f}$ et $m+1$; elle est bien applicable, parce que $\vec{f}$ est bien une primitive d'ordre $m+1$ de la fonction $\vec{f}^{(m+1)}$ supposée continue, et dont les dérivées d'ordre $0,1,2,\ldots,m$, prennent les valeurs $\vec{f}(c), \vec{f}'(c), \vec{f}''(c),\ldots,\vec{f}^{(m)}(c)$, au point c.

La deuxième formule s'obtient par changement de variable, ou encore en appliquant la première à l'intervalle $[0,1]$, pour la fonction $t \longrightarrow \vec{f}(x+t\vec{h})$.

Corollaire 3 (Nouvelle expression générale de la formule de Taylor)

Soit $\vec{f}$ ~~une application de~~ classe C^{m+1} d'un ouvert Ω d'un espace affine normé E dans un espace affine normé F. Si alors tout le segment $[x, x+\vec{h}]$ est contenu dans Ω, on a la formule:

$$(\mathrm{IV},9;49)\qquad \vec{f}(x+\vec{h}) = \vec{f}(x) + \frac{\vec{f}'(x)}{1!}\cdot\vec{h} + \cdots + \frac{\vec{f}^{(m)}(x)}{m!}\cdot\vec{h}^m + \int_0^1 \frac{(1-t)^m}{m!}\left(\vec{f}^{(m+1)}(x+t\vec{h})\cdot\vec{h}^{m+1}\right)dt$$

Démonstration - Il suffit en effet d'appliquer la formule de Taylor (IV,9;47) à l'intervalle $[0,1]$ et à la fonction $t \longrightarrow \vec{f}(x+t\vec{h})$; la dérivée d'ordre k de cette fonction au point t, est $\vec{f}^{(k)}(x+t\vec{h})\cdot\vec{h}^k$, d'après le théorème des fonctions composées. Comme on le voit, cette formule de Taylor est tout-a-fait remarquable et au fond beaucoup plus intéressante que celles qui ont été données au chapitre III.

Elle donne en effet une expression explicite du terme complémentaire. Par contre, évidemment, elle suppose un peu plus, à savoir que la fonction $\vec{f}$ admet une dérivée d'ordre $m+1$ continue dans Ω. Moyennant cette hypothèse supplémentaire, les différentes formes de la formule de Taylor données au chapitre III peuvent toutes être retrouvées.

Si par exemple f est a valeurs réelles, on peut appliquer le théorème de la moyenne sous la forme (IV,1;31); il existe donc un nombre θ, $0 < \theta < 1$, tel que l'on ait

(IV,9;50) $$\int_0^1 \frac{(1-t)^m}{m!}\left(f^{(m+1)}(x+t\vec{h}).\vec{h}^{m+1}\right)dt =$$

$$\left(f^{(m+1)}(x+\theta\vec{h}).\vec{h}^{m+1}\right)\int_0^1 \frac{(1-t)^m}{m!}dt$$

$$= \frac{f^{(m+1)}(x+\theta\vec{h})}{(m+1)!}.\vec{h}^{m+1},$$

et ceci redonne la formule de Taylor (III,7;1 bis). SI maintenant l'on ne fait aucune hypothèse sur F , on a de toutes façons, d'après la formule (IV, 1 ; 24), la majoration du terme complémentaire

(IV,9;51) $$\left\|\int_0^1 \frac{(1-t)^m}{m!}\left(f^{(m+1)}(x+t\vec{h}).\vec{h}^{m+1}\right)dt\right\|$$

$$\leq \|\vec{h}\|^{m+1}|||f^{(m+1)}|||\int_0^1 \frac{(1-t)^m}{m!}dt$$

$$= |||f^{(m+1)}|||\frac{\|\vec{h}\|^{m+1}}{(m+1)!}$$

et ceci redonne la formule de Taylor (III,7;2).

Jais., si L est une application $(m+1)$-linéaire, continue, de E dans $\vec{F}$, on peut aussi écrire, compte tenu de ce que $\int_0^1 \frac{(1-t)^m}{m!}dt = \frac{1}{(m+1)!}$:

(IV,9;52) $$\int_0^1 \frac{(1-t)^m}{m!}\left(f^{(m+1)}(x+t\vec{h}).\vec{h}^{m+1}\right)dt$$

$$= \frac{L}{(m+1)!}.\vec{h}^{m+1} + \int_0^1 \frac{(1-t)^m}{m!}\left(\left(f^{(m+1)}(x+t\vec{h})-L\right).\vec{h}^{m+1}\right)dt,$$

et la dernière intégrale est majorée par:

$$\text{(IV,9;53)} \qquad \sup_{0 \leq t \leq 1} \left\| f^{(m+1)}(x + t\vec{h}) - L \right\| \; \|\vec{h}\|^{m+1} \int_0^1 \frac{(1-t)^m}{m!} \, dt$$

$$= \sup_{0 \leq t \leq 1} \left\| f^{(m+1)}(x + t\vec{h}) - L \right\| \; \frac{\|\vec{h}\|^{m+1}}{(m+1)!} \; ,$$

et ceci redonne la formule de Taylor sous la forme (III,7;9).

Ainsi, en définitive, toutes les expressions antérieurement trouvées de la formule de Taylor sont des conséquences de celle-ci. Toutefois, répétons encore qu'elles ont été démontrées avec des hypothèses moins restrictives, et, comme nous l'avons vu dans le calcul différentiel, ces hypothèses moins restrictives étaient très utiles dans certains théorèmes (voir par exemple la démonstration du théorème 14 du chapitre III).

Formule de l'intégration par parties

Théorème 92

1°/ Cas élémentaire -

Soient $\vec{u}$ et $\vec{v}$ des fonctions de classe C^1, définies sur un intervalle fermé $[a,b]$ * de $\mathbb{R}$, à valeurs dans des espaces de Banach $\vec{E}$ et $\vec{F}$ respectivement. Soit B une application bilinéaire continue de $\vec{E} \times \vec{F}$ dans un espace de Banach $\vec{G}$.

Alors, on a la formule d'intégration par parties:

$$\text{(IV,9;54)} \qquad \int_a^b B\left(\vec{u}(x), \vec{v}'(x)\right) dx = B\left(\vec{u}(b), \vec{v}(b)\right) - B\left(\vec{u}(a), \vec{v}(a)\right) - \int_a^b B\left(\vec{u}'(x), \vec{v}(x)\right) dx .$$

2°/ Cas général

Soient $\vec{\mu}$ et $\vec{\nu}$ deux mesures de Radon, définies sur un intervalle fermé $[a,b]$ de $\mathbb{R}$, à valeurs dans des espaces de Banach $\vec{E}$ et $\vec{F}$ respectivement, et admettant des bases ≥ 0-

* Dans cette formule, a et b sont quelconques sur $\mathbb{R}$: $a < b$, $a = b$ ou $a > b$.

Soit B une application bilinéaire continue de $\vec{E} \times \vec{F}$ dans un espace de Banach, $\vec{G}$. Soient $\vec{M}$ et $\vec{N}$ des intégrales indéfinies de $\vec{\mu}$ et $\vec{\nu}$, à valeurs dans $\vec{E}$ et $\vec{F}$. Si $\vec{\mu}$ et $\vec{\nu}$ n'ont pas de masse ponctuelle commune, $\int_a^b B(\vec{M}, d\vec{\nu})$ et $\int_a^b B(d\vec{\mu}, \vec{N})$ ont un sens, et on a la formule générale d'intégration par parties:

$$\text{(IV,9;55)} \qquad \int_a^b B(\vec{M}, d\vec{\nu}) = B(\vec{M}(b), \vec{N}(b)) - B(\vec{M}(a), \vec{N}(a)) - \int_a^b B(d\vec{\mu}, \vec{N}) .$$

Démonstration - La première formule est immédiate; elle résulte de la formule (IV,9;38 ter) appliquée à $\vec{F} = B(\vec{u},\vec{v})$, compte tenu de la formule de dérivation d'une application bilinéaire continue (théorème 12 du chapitre III) :

$$\text{(IV,9;56)} \qquad \vec{F}' = \vec{f} = B(\vec{u}', \vec{v}) + B(\vec{u}, \vec{v}') .$$

On voit pourquoi le cas général est effectivement une généralisation du cas particulier.

11 suffit en effet de poser $d\vec{\mu} = \vec{u}'\,dx$, et $d\vec{\nu} = \vec{v}'\,dx$. Alors $\vec{M} = \vec{u}$ et $\vec{N} = \vec{v}$ sont des intégrales indéfinies de $\vec{\mu}$ et $\vec{\nu}$; $\vec{\mu}$ et $\vec{\nu}$ sont bien de base $dx \geq 0$; comme $\vec{\mu}$ et $\vec{\nu}$ n'ont pas de masse ponctuelle, elles n'ont certainement pas de masse ponctuelle commune, alors la formule (IV,9;55) (compte tenu de la définition (IV,5;13)) redonne la formule (IV,9;54).

Démontrons donc maintenant le cas général. Par suite de la relation $\int_a^b = -\int_b^a$, on peut toujours supposer $a < b$

Soient μ_0 et ν_0 des mesures réelles ≥ 0 , telles que $\vec{\mu}$ soit de base μ_0 et que $\vec{\nu}$ soit de base ν_0 .

Si alors on pose $\lambda_0 = \mu_0 + \nu_0$, on voit que μ_0 et ν_0 sont toutes deux majorées par λ_0 , donc (théorème 53 de Dunford-Pettis) de base λ_0 .

On peut donc écrire $d\mu_0 = p_0\, d\lambda_0$, et $d\nu_0 = q_0\, d\lambda_0$, où p_0 et q_0 sont localement λ_0-intégrables. Comme on a aussi $d\vec{\mu} = \vec{p}\, d\mu_0$, $d\vec{\nu} = \vec{q}\, d\nu_0$, on aura finalement (corollaire du théorème 51) :

$$d\vec{\mu} = (\vec{p}\,p_0)\, d\lambda_0 = \vec{f}\, d\lambda_0 \,, \quad d\vec{\nu} = (\vec{q}\,q_0)\, d\lambda_0 = \vec{g}\, d\lambda_0 \,. \tag{IV,9;57}$$

Considérons alors l'intégrale double :

$$\iint_{[a,b[\,\times\,[a,b[} B\left(\vec{f}(x),\ \vec{g}(y)\right) d\lambda_0(x)\, d\lambda_0(y) \tag{IV,9;58}$$

Elle a un sens d'après le théorème 79, puisque $\vec{f}$ et $\vec{g}$ sont $d\lambda_0$- intégrables sur le <u>compact</u> [a.&] .Le même théorème montre que cette intégrale est égale à :

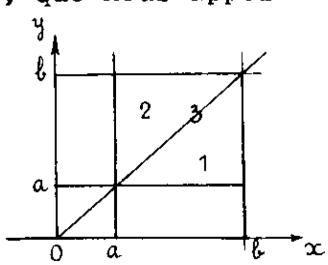
$$B\left(\int_a^b \vec{f}\, d\lambda_0 \,,\ \int_a^b \vec{g}\, d\lambda_0\right) = B\left(\vec{M}(b) - \vec{M}(a), \vec{N}(b) - \vec{N}(a)\right). \tag{IV,9;59}$$

On peut alors décomposer le carré $[a,b[\times [a,b[$ en réunion de trois ensembles disjoints (et sûrement λ_0- mesurables, car intersections d'ouverts et de fermés) :

1°) L'ensemble $a \leq y < x < b$, marqué 1 sur la figure,

2°) L'ensemble $a \leq x < y < b$, marqué 2 sur la figure,

3”) La diagonale du carré, c'est-à-dire l'ensemble $a \leq x = y < b$, que nous appellerons 3.



Fig

Calculons alors l'intégrale analogue à (IV,9;59), sur le premier ensemble.

D'après le théorème de Fubini-Lebesgue (théorème 77), on peut écrire:

(IV,9;60) $$\iint_{(1)} = \int_{]a,b[} d\lambda_0(x) \int_a^x B\left(\vec{f}(x), \vec{g}(y)\right) d\lambda_0(y)$$

Mais la fonction $\vec{g}$ est λ_0- intégrable sur l'intervalle $[a, x[$, et, comme l'application $\vec{g} \to B\left(\vec{f}(x), \vec{g}\right)$ est une application continue de $\vec{F}$ dans $\vec{G}$, on peut écrire :

(IV,9;61) $$\int_a^x B\left(\vec{f}(x), \vec{g}(y)\right) d\lambda_0(y) = B\left(\vec{f}(x), \int_a^x \vec{g}\, d\lambda_0\right).$$

Mais l'intégral: $\int_a^x \vec{g}\, d\lambda_0$ n'est autre que $\vec{N}(x) - \vec{N}(a)$; d'autre part $\int_a^x \vec{g}\, d\lambda_0$ est nulle pour $x = a$, donc la deuxième intégration $\int_{]a,b[}$ peut se remplacer par $\int_{[a,b[} = \int_a^b$, d'où la valeur de cette première intégrale; compte tenu de (IV,5;13) :

(IV,9;62) $$\iint_{(1)} = \int_a^b B\left(\vec{f}(x), \vec{N}(x) - \vec{N}(a)\right) d\lambda_0(x)$$

$$= \int_a^b B\left(\overrightarrow{d\mu}(x), \vec{N}(x) - \vec{N}(a)\right),$$

l'intégrale ayant nécessairement un sens.

Un calcul exactement analogue relatif à l'intégrale sur l'ensemble (2), donnera le résultat:

(IV,9;63) $$\iint_{(2)} = \int_a^b B\left(\vec{M}(x) - \vec{M}(a), \overrightarrow{d\nu}(x)\right).$$

Montrons maintenant que l'intégrale sur l'ensemble 3, c'est-à-dire sur la diagonale, est nulle.

Nous pouvons encore la calculer par le théorème de Fubini. Si nous fixons x, nous devons d'abord calculer $\int B(\vec{f}(x), \vec{g}(y))\, d\lambda_0(y)$ sur l'ensemble $\{x\}$. Cette intégrale est nulle, sauf si la mesure λ_0 possède une masse ponctuelle au point x. S'il en est ainsi, et si nous appelons $\lambda_0(\{x\})$ cette masse ponctuelle, l'intégrale précédente est égale a $B(\vec{f}(x), \vec{g}(x))\, \lambda_0(\{x\})$, et par conséquent l'intégrale sur la diagonale est donnée par la formule:

$$\iint_{(3)} = \int_a^b B\left(\vec{f}(x), \vec{g}(x)\right) \lambda_0\left(\{x\}\right) d\lambda_0(x) . \tag{IV,9;64}$$

Mais remarquons que la masse de $\vec{\mu}$ au point x, c'est-à dire l'intégrale par rapport à $\vec{f}\, d\lambda_0$ de la fonction caractéristique de l'ensemble $\{x\}$ réduit à x, n'est autre que $\vec{f}(x)\lambda_0(\{x\})$; de même la masse de $\vec{\nu}$ au point x est $\vec{g}(x)\lambda_0(\{x\})$. Alors l'hypothèse que $\vec{\mu}$ et $\vec{\nu}$ n'ont pas de masse ponctuelle commune, revient à écrire que, pour tout point x où λ_0 possède-une masse ponctuelle, l'une au moins des deux fonctions $\vec{f}$, $\vec{g}$ est nulle; donc la fonction $x \longrightarrow B(\vec{f}(x), \vec{g}(x))\, \lambda_0(\{x\})$ est identiquement nulle, et l'intégrale (IV,9;64) est bien nulle.

On a donc, en comparant les résultats (IV,9;59), (IV,9;62), (IV,9;63), le résultat:

$$\begin{aligned} & B\left(\vec{M}(b) - \vec{M}(a), \vec{N}(b) - \vec{N}(a)\right) \\ & = \int_a^b B\left(\overrightarrow{d\mu}(x), \vec{N}(x) - \vec{N}(a)\right) + \int_a^b B\left(\vec{M}(x) - \vec{M}(a), \overrightarrow{d\nu}(x)\right). \end{aligned} \tag{IV,9;65}$$

En développant, on en déduit, en permutant, comme nous l'avons déjà fait à (IV,9;61), l'intégration et une application linéaire continue :

$$(\text{IV},9;66) \quad B\left(\vec{M}(b), \vec{N}(b)\right) - B\left(\vec{M}(a), \vec{N}(b)\right) - B\left(\vec{M}(b), \vec{N}(a)\right) + B\left(\vec{M}(a), \vec{N}(a)\right)$$

$$= \int_a^b B\left(d\vec{\mu}, \vec{N}\right) - B\left(\int_a^b d\vec{\mu}, \vec{N}(a)\right)$$

$$+ \int_a^b B\left(\vec{M}, d\vec{\nu}\right) - B\left(\vec{M}(a), \int_a^b d\vec{\nu}\right)$$

Et, comme $\int_a^b d\vec{\mu} = \vec{M}(b) - \vec{M}(a)$, $\int_a^b d\vec{\nu} = \vec{N}(b) - \vec{N}(a)$,

on trouve, après simplification par suppression des termes égaux dans les 2 membres, la formule cherchée (IV,9;55).

Changement de variable dans le calcul des intégrales simples

Théorème 93 - Soit $t \longrightarrow x = \xi(t)$ une fonction réelle de classe C^1 sur l'intervalle $[\alpha,\beta]$ de $\mathbb{R}$, avec $\xi(\alpha) = a$, $\xi(\beta) = b$ *
Soit $\vec{f}$ une fonction continue sur l'intervalle $[a,b]$, à valeurs dans un espace de Banach $\vec{F}$.

Alors on a la formule de "changement de variable" :

$$(\text{IV},9;67) \quad \int_a^b \vec{f}(x)\, dx = \int_\alpha^\beta \vec{f}\left(\xi(t)\right) \xi'(t)\, dt .$$

Démonstration - Considérons la fonction Φ définie par

$$(\text{IV},9;68) \quad \vec{\Phi}(u) = \int_a^u \vec{f}(x)\, dx \ ;$$

* Ici encore, α, β a, b sont quelconques sur $\mathbb{R}$; $\alpha \lesseqgtr \beta$, $a \lesseqgtr b$.

c'est une fonction de classe C^1, primitive de $\vec{f}$, puisque $\vec{f}$ est supposée continue. En outre, elle vérifie $\vec{\Phi}(a) = \vec{0}$, et $\vec{\Phi}(b)$ est le premier membre de (IV,9;67). Si alors nous considérons la fonction $s \to \vec{\Psi}(s) = \vec{\Phi}(\xi(s))$, c'est une fonction de classe C^1, d'après le théorème des fonctions composées. Elle vérifie $\vec{\Psi}(\alpha) = \vec{\Phi}(a) = \vec{0}$, et sa dérivée est donnée par

$$\vec{\Psi}'(s) = \vec{\Phi}'(\xi(s))\, \xi'(s) = \vec{f}(\xi(s))\, \xi'(s) \tag{IV,9;69}$$

Il en résulte que l'on peut écrire, pour toute valeur de s,

$$\vec{\Psi}(s) = \int_\alpha^s \vec{f}(\xi(t))\, \xi'(t)\, dt \; ; \tag{IV,9;70}$$

mais, pour $s = \beta$, on a $\vec{\Psi}(\beta) = \Phi(b)$, premier membre de (IV,9;67); ~~et~~ comme alors le 2ème membre de (IV,9;70) est aussi celui de (IV,9;67), le théorème est démontré.

La formule que nous utilisons la n'est pas généralisable à des intégrales multiples. parce qu'elle utilise le concept, tout-a-fait particulier a la droite, d'intégrale $\int_a^b$. Nous avons Intérêt à pouvoir aussi l'utiliser sous une autre forme, utilisant le symbole $\int_{[a,b]}$.

Soit 1 1' intervalle $[a,b]$ ou $[b,a]$, suivant celle des deux bornes qui est la plus grande; et soit de même J l'intervalle $[\alpha,\beta]$ ou $[\beta, a]$.

Si. aux hypothèses du théorème 93, nous ajoutons la suivante :

<u>ξ est une fonction monotone sur J</u>, alors bien évidemment la formule (IV,9;67) peut aussi s'écrire :

$$\int_I \vec{f}(x)\, dx = \int_J \vec{f}(\xi(t))\, |\xi'(t)|\, dt \; . \tag{IV,9;71}$$

En effet, si ξ est croissante, alors $|\xi'(t)|$ n'est autre que $\xi'(t) \geq 0$; d'autre part, les intégrales $\int_I$ et $\int_J$ ne sont autres que $\varepsilon\int_\alpha^b$ et $\varepsilon\int_\alpha^\beta$, où $\varepsilon = \pm 1$; la valeur de ε est la même dans les deux cas, puisque ξ est supposée croissante. Si au contraire la fonction ξ est décroissante, alors $|\xi'(t)|$ est $-\xi'(t)$; mais les valeurs à prendre pour ε dans les deux intégrales précédentes sont opposées l'une à l'autre.

Nous allons ddmontrer la même formule (IV,9;71) dans un cas un peu différent.

Théorème 94 - Soient $\mathcal{O}$ et Ω deux ouverts de la droite réelle $\mathbb{R}$, soit ξ un homéomorphisme de Ω sur $\mathcal{O}$, de classe C^1 ainsi que son homéomorphisme réciproque.
soit $\vec{f}$ une fonctionantinue sur $\mathcal{O}$, à valeurs dans un espace de Banach $\vec{F}$ et à support compact.

Alors on a la formule

$$\int_{\mathcal{O}} \vec{f}(x)\,dx = \int_{\Omega} \vec{f}(\xi(t))\,|\xi'(t)|\,dt\,. \tag{IV,9;72}$$

Démonstration Considérons un point quelconque du support de $\vec{f}$. Comme $\mathcal{O}$ est un voisinage de ce point, il existe au moins un Intervalle ouvert ayant pour centre ce point, d'adhérence entièrement contenue dans $\mathcal{O}$. L'ensemble de tous les intervalles ouverts, ainsi formés pour tous les points du support de $\vec{f}$, est un recouvrement de ce support, et comme ce support est compact, il suffit d'un nombre fini de ces Intervalles, soit $]a_i, b_i[$, $i \in I$, pour recouvrir ce support.

Soit alors γ_i une partition de l'unité relative au recouvrement de ce support par les $]a_i, b_i[$

On a alors

$$\int_{\mathcal{O}} \vec{f}(x)\,dx = \sum_{i\in I} \int_{]a_i,b_i[} \gamma_i(x)\,\vec{f}(x)\,dx\,. \tag{IV,9;73}$$

Mais la fonction ξ^{-1} doit être un homéomorphisme de $]a_i, b_i[$ sur son image. Par conséquent la fonction ξ^{-1} est nécessairement strictement monotone dans $]a_i, b_i[$ * ,ainsi que son application réciproque\ (théorème 37 du chapitre II). Si alors nous posons $a_i = \xi(\alpha_i)$, $b_i = \xi(\beta_i)$, comme ξ est monotone, on peut appliquer la formule (IV,9;67),et écrire

$$(IV,9;74) \quad \int_{|a_i,b_i|} \gamma_i(x)\vec{f}(x)\,dx = \int_{|\alpha_i,\beta_i|} \gamma_i(\xi(t))\,\vec{f}(\xi(t))\,|\xi'(t)|\,dt$$

d'où l'on déduit alors

$$(IV,9;75) \quad \int_{\mathcal{O}} \vec{f}(x)\,dx = \sum_{i\in I} \int_{|\alpha_i,\beta_i|} \gamma_i(\xi(t))\vec{f}(\xi(t))\,|\xi'(t)|\,dt$$

$$= \sum_{i\in I} \int_{\Omega} \gamma_i(\xi(t))\,\vec{f}(\xi(t))\,|\xi'(t)|\,dt = \int_{\Omega} \vec{f}(\xi(t))\,|\xi'(t)|\,dt$$

et le théorème est démontré.

<u>Corollaire 1 - Dans les conditions de l'énoncé du théorème, l'image, par l'application ξ de Ω dans $\mathcal{O}$, de la mesure $|\xi'(t)|\,dt$ sur Ω , est la mesure de Lebesgue dx sur $\mathcal{O}$.</u>

• Sinon, il existerait 3 points u, v, w, de cet intervalle , tels que $u < v < w$, et $\xi^{-1}(u) \leq \xi^{-1}(v)$, $\xi^{-1}(w) \leq \xi^{-1}(v)$, ou $\xi^{-1}(u) \geq \xi^{-1}(v)$, $\xi^{-1}(w) \geq \xi^{-1}(v)$. Plaçons - nous par exemple dans la 1ère hypothèse, avec $\leq$. En fait on aurait même $<$, puisque ξ^{-1} est bijective. Mais alors ξ^{-1} prendrait au moins 2 fois toute valeur strictement comprise entre $\xi^{-1}(v)$ et $\mathrm{Sup}(\xi^{-1}(u), \xi^{-1}(w))$, une fois dans l'intervalle $[u, v[$, et une fois dans l'intervalle $]v, w]$, d'après le corollaire du théorème 33 du chapitre II;ce qui est absurde, puisqu'elle est bijective.

Il suffit de prendre pour f une fonction scalaire continue à support compact sur $\mathcal{O}$, et on tombe sur la définition de l'image d'une mesure.

Corollaire 2 - Si ξ vérifie les conditions de l'énoncé du théorème, et si $\vec{f}$ est une fonction quelconque sur $\mathcal{O}$, à valeurs dans un espace de Banach $\vec{F}$, $\vec{f}$ est intégrable par rapport à dx sur $\mathcal{O}$, si et seulement si la fonction $t \longrightarrow \vec{f}(\xi(t))\,|\xi'(t)|$ est intégrable par rapport à dt sur Ω , et on a alors (IV,9;72).

Car, d'après le corollaire 1, l'image par ξ de $|\xi'(t)|dt$ est dx . Alors $\vec{f}$ est intégrable pour dx , si et seulement si $\vec{f}(\xi(t))$ est intégrable pour $|\xi'(t)|\,dt$, en vertu du théorème 60; et ceci a lieu si et seulement si $\vec{f}(\xi(t))|\xi'(t)|$ est intégrable pour dt , d'après le théorème 51. Les théorèmes 60 et 51 prouvent en outre (IV, 9;72).

Intégrales impropres sur la droite

soit $]a,b[$, $a<b$, un intervalle de $\mathbb{R}$, ne contenant pas son extrémité b . Soit $\vec{\mu}$ une mesure de Radon sur $]a,b[$, à valeurs dans un espace de Banach $\vec{E}$, et de base $\geqslant 0$. On dit que l'intégrale impropre $\int_{]a,\to b[} d\vec{\mu}$ est convergente, si $\int_{]a,b'[} d\vec{\mu}$ a un sens pour tout b' de $]a,b[$, et a une limite dans $\vec{E}$, lorsque b' tend vers b par valeurs dans $]a,b[$. Nous avons écrit $\int_{]a,\to b[}$ (où $\to b$ veut dire qu'on prend b' dans $]a,b[$, et que b' tend vers b) et non $\int_{]a,b[}$, car rien ne dit que cette dernière expression ait un sens en tant qu'intégrale d'une fonction intégrable. Dire que $\int_{]a,\to b[}$ est convergente, signifie que, pour toute intégrale indéfinie M de $\vec{\mu}$, à valeurs dans un espace affine normé E, d'espace vectoriel associé $\vec{E}$, M (b') a une limite lorsque b' tend vers b par valeurs dans $]a,b[$.

Soit λ une mesure $\geqslant 0$, $\vec{f}$ une fonction à valeurs dans l'espace de Banach $\vec{F}$, localement λ - intégrable.

Alors on dit que **l'intégrale** impropre $\int_{|a,\rightarrow b[} \vec{f}\, d\lambda$ de $\vec{f}$ par rapport **à** λ est convergente, si $\int_{|a,\rightarrow b[} d\vec{\mu}$, $\vec{\mu} = \vec{f}\lambda$, **est** convergente. On dit qua $\int_{|a,\rightarrow b[} \vec{f}\, d\lambda$ **est absolument** ou **normalement** convergente, **si** $\int_{|a,\rightarrow b[} \|\vec{f}\|\, d\lambda$ est convergente. On a une définition **analogue** pour $\int_{]a\leftarrow, b|}$.

On peut aussi considérer $\int_{]a\leftarrow, \rightarrow b[}$; on dit que cette Intégrale est convergente, si $\int_{]a', b'[}$ a une limite quand a' **et** b' tendent respectivement (indépendamment l'un de l'autre) vers a et b , en restant dans $]a, b[$. SI, par exemple, $a < b$, cela veut dire que, quel que soit $\varepsilon > 0$, **il** existe $a'_0 > a$ et $b'_0 < b$ tels que $a < a' \leq a'_0$, $b'_0 \leq b' < b$, **entraîne** $\left\| \int_{]a\leftarrow, \rightarrow b[} - \int_{]a', b'[} \right\| \leq \varepsilon$.

Cela signifie simplement que, pour c arbitraire **dans** $]a, b[$, $\int_{|c, \rightarrow b[}$ et $\int_{]a\leftarrow, c|}$ ont tous deux un sens.

Assez fréquemment a ou b valent $\pm\infty$; en particulier, on peut **étudier** $\int_{]-\infty\leftarrow, \rightarrow +\infty[}$. On introduira aussi les **symboles** $\int_a^{\rightarrow b}$, $\int_{a\leftarrow}^{b}$, $\int_{a\leftarrow}^{\rightarrow b}$.

Théorème 97 - Soient λ une mesure de Radon ≥ 0 sur $|a, b[$, $\vec{f}$ une fonction **définie** sur $[a, b[$, **à** valeurs **dans un** espace de Banach F , λ-**intégrable** sur tout $|a, b'[$, $b' \in]a, b[$ Pour que l'intégrale $\int_{|a,\rightarrow b[} \vec{f}\, d\lambda$ soit absolument **convergente, il faut et il suffit**bque $\int_{|a, b'[} \|\vec{f}\|\, d\lambda$ n é . **lorsque** b' **varie dans** $|a, b[$; **il faut et** il suffit aussi que la fonction $\vec{f}$ soit **intégrable** sur l'intervalle $|a, b[$ relativement **à** la **mesure** λ ; et dans ce cas, **l'intégrale** $\int_{|a,\rightarrow b[} \vec{f}\, d\lambda$ est a fortiori convergente et est **égale à** $\int_{|a, b[} \vec{f}\, d\lambda$

Ce théorème montre bien, comme nous l'avions déjà dit. que <u>la notion d'intégraoilité au sens de Lebesgue correspond à une généralisation de la notion d'intégrale impropre de Riemann absolument convergente, vue en Mathématiques Spéciales. Par contre les intégrales impropres de Riemann semi - convergentes ne sont pas des integrales de fonctions intégrables au sens de Lebesgue</u>; <u>et alors</u>, $\int_{|a,b[} \vec{f}\, d\lambda$ <u>n'ayant pas de sens, l'écriture</u> $\int_{|a,\to b[} \vec{f}\, d\lambda$ <u>est</u>, <u>théoriquement, absolument indispensable</u> * . <u>Ce théorème montre d'autre part que la notion d'intégrale impropre absolument convergente est, au stade actuel, sans intérêt pour nous, puisqu'elle n'apporte rien de nouveau par rapport à l'intégrale de Lebesgue. Seules sont nouvelles les intégrales impropres semi-convergentes, c'est-k-dire convergentes mais non absolument convergentes.</u>

<u>Démonstration</u> Appelons φ_b, la fonction caractéristique de l'intervalle $[a, b'[$.

1°/ Puisque $b' \to \int_{|a,b'[} \|\vec{f}\|\, d\lambda$ est une fonction croissante, elle a une limite finie lorsque b' tend vers b , si et seulement si elle est bornée. Donc il est bien exact que $\int_{|a,\to b[} \|\vec{f}\|\, d\lambda$ est convergente, si et seulement si $\int_{|a,b'[} \|\vec{f}\|\, d\lambda$ est bornée quand b' varie.

2°/ $\vec{f}$ est mesurable puisque localement intégrable (note (*) page 523); donc elle est intégrable si et seulement si $\int^*_{|a,b[} \|\vec{f}\|\, d\lambda = \int \|\vec{f}\| \varphi_b\, d\lambda < +\infty$; mais, d'après le théorème 36 de Fatou, ceci est la limite de $\int \|\vec{f}\| \varphi_{b'}\, d\lambda = \int^*_{|a,b'[} \|\vec{f}\|\, d\lambda = \int_{|a,b'[} \|\vec{f}\|\, d\lambda$, donc $\vec{f}$ est bien intégrable si et seulement si $\int_{|a,\to b[} \|\vec{f}\|\, d\lambda$ a un sens-

* Nous disons "théoriquement" indispensable. <u>La plupart du temps, on ne met pas la flèche</u> ! Par exemple, on écrit toujours $\int_0^{+\infty} \frac{\sin x}{x}\, dx = \frac{\pi}{2}$, alors que $\frac{\sin x}{x}$ n'est pas intégrable par rapport à dx sur $]0, +\infty[$; on devrait écrire

$$\int_{|0,\to+\infty[} \frac{\sin x}{x}\, dz \quad \text{ou} \quad \int_0^{\to+\infty} \frac{\sin x}{x}\, dx .$$

3°/ Ces conditions étant supposées réalisées, $\vec{f}\varphi_{b'}$ converge simplement vers $\vec{f}\varphi_b$ lorsque b' tend vers b, et on a la majoration $\|\vec{f}\varphi_{b'}\| \leq \|\vec{f}\|$, fonction $\geqslant 0$ intégrable fixe: alors le théorème **35** de Lebesgue dit bien que $\int_{[a,b'[} \vec{f}\, d\lambda = \int \vec{f}\varphi_{b'}\, d\lambda$ converge vers $\int_{[a,b[} \vec{f}\, d\lambda = \int \vec{f}\varphi_b\, d\lambda$ lorsque b' tend vers b, donc $\int_{[a,\to b[} \vec{f}\, d\lambda$ est bien convergente, et égale à $\int_{[a,b[} \vec{f}\, d\lambda$.

<u>Corollaire 1</u> - Sur le borné $[a,b]$, $a<b$, contenant l'origine, la fonction $\frac{1}{|x|^\alpha}$ est intégrable pour la mesure de Lebesgue, si et seulement si $\alpha < 1$; sur l'intervalle $[a,+\infty[$ ou $]-\infty,a]$ ne contenant pas l'origine, la fonction $\frac{1}{|x|^\alpha}$ est intégrable, si et seulement si $\alpha \geq 1$.

La fonction $\frac{1}{|x|^\alpha}$ est continue dans $\complement 0$, donc localement intégrable dans $]-\infty,0[$ et dans $]0,+\infty[$.

Soit, par exemple, $a \leqslant 0 \leqslant b$, $a < b$.

La fonction $\frac{\varepsilon}{1-a}\frac{1}{|x|^{\alpha-1}}$, $\varepsilon = +1$ pour $x > 0$, -1 pour $x < 0$, est une primitive de $\frac{1}{|x|^\alpha}$ dans $\complement 0$, pour $\alpha \neq 1$; $\varepsilon \log|x|$ est une primitive pour $\alpha = 1$.

Alors, pour $0 < c < d < +\infty$:

$$(\text{IV}, 9; 99) \qquad \int_{[c,d]} \frac{dx}{|x|^\alpha} = \begin{cases} \frac{1}{1-\alpha}\left(\frac{1}{d^{\alpha-1}} - \frac{1}{c^{\alpha-1}}\right). \\ \log d - \log c, \text{ pour } \alpha = 1. \end{cases}$$

La fonction $\frac{1}{x^\alpha}$ étant $\geqslant 0$, elle est intégrable dans $]0,b]$ si et seulement si l'intégrale impropre $\int_{]0\leftarrow,b]}$ est convergente. c'est-à-dire si et seulement si ses primitives ont une limite finie quand $x > 0$ tend vers 0.

Pour cela, il faut et il suffit que $\alpha < 1$. **Alors**

$\int_{]0,b]} \frac{dx}{x^\alpha} = \int_{]0\leftarrow,b]} \frac{dx}{x^\alpha} = \frac{1}{1-\alpha}\frac{1}{b^{\alpha-1}}$. De même pour l'intervalle $[a,0]$.

Mais $\frac{1}{|x|^\alpha}$ est intégrable dans $|a,b|$ si et seulement si elle l'est dans $[a,0]$ et dans $[0,b]$, c'est-à-dire si et seulement si $\alpha < 1$.

Soit maintenant $a > 0$. Alors $\frac{1}{x^\alpha}$ est intégrable dans $|a,+\infty[$, si et seulement si l'intégrale impropre $\int_{|a,\to+\infty[} \frac{dx}{x^\alpha}$ converge, c'est-à-dire si et seulement si $\alpha > 1$. Alors $\int_{|a,+\infty[} \frac{dx}{x^\alpha} = \int_{|a,\to+\infty[} \frac{dx}{x^\alpha} = \frac{1}{\alpha-1}\frac{1}{a^{\alpha-1}}$. Même raisonnement pour l'intervalle $]-\infty,a|$, $a<0$.

Corollaire 2 - Pour α complexe, et x réel, >0 appelons $\frac{1}{x^\alpha}$ le nombre $e^{\alpha \operatorname{Log} x}$. Alors la fonction $x \to \frac{1}{x^\alpha}$ est intégrable pour la mesure de Lebesgue dans $]0,a|$, $a>0$, si et seulement si $\mathcal{R}\alpha < 1$; elle est intégrable dans $|a,+\infty[$, si et seulement si $\mathcal{R}\alpha > 1$. En outre $\int_{]0\leftarrow,a|} \frac{dx}{x^\alpha}$ et $\int_{|a,+\infty[} \frac{dx}{x^\alpha}$ ne sont convergentes que si elles le sont absolument.

En effet, f est continue, donc mesurable. Alors elle est intégrable si et seulement si sa valeur absolue $\frac{1}{x^{\mathcal{R}\alpha}}$ est intégrable, et nous sommes ramenés au corollaire 1.

Pour la convergence des intégrales, nous utiliserons la même primitive $\frac{1}{1-\alpha}\frac{1}{x^{\alpha-1}}$ ou $\operatorname{Log} x$. A partir de la formule (IV,9;99), on voit encore qu'il n'y a convergence que pour $\alpha<1$ pour $\int_{]0\leftarrow,a|}$, $\alpha>1$ pour $\int_{|a,+\infty[}$, c'est-à-dire quand il y a convergence absolue ou intégrabilité.

Nous allons maintenant étendre au calcul des intégrales impropres, par rapport à la mesure dx, les méthodes d'intégration par parties et de changement de variable.

Théorème 98 (Intégration par parties)

1°/ Cas **élémentaire**

Soient $\vec{u}$ et $\vec{v}$ des fonctions **définies** sur l'intervalle $[a,b[$ **de** $\mathbb{R}$, à valeurs dans des espaces de Banach $\vec{E}$ et $\vec{F}$ **respectivement,** et de classe $\mathcal{C}^1$. Soit B une application bilinéaire continue de $\vec{E}\times\vec{F}$ **dans un espace** de Banach $\vec{G}$. Si l'expression $B(\vec{u}(x),\vec{v}(x))$ admet une limite **lorsque** $x<b$ tend vers b, **alors** $\int_a^{\to b} B(\vec{u},\vec{v}')\,dx$ **a un sens, si et** seulement si $\int_a^{\to b} B(\vec{u}',\vec{v})\,dx$ a un **sens, et l'on a la** formule

$$\int_a^{\to b} B(\vec{u},\vec{v}')\,dx = \lim_{b'\to b} B\left(\vec{u}(b'),\vec{v}(b')\right) - B\left(\vec{u}(a),\vec{v}(a)\right) - \int_a^{\to b} B(\vec{u}',\vec{v})\,dx. \tag{IV,9;100}$$

2°/ Cas **général**(*)

Soient $\vec{\mu}$ et $\vec{\nu}$ des mesures de Radon **sur** $[a,b[$, à valeurs dans des espaces de Banach $\vec{E}$ et $\vec{F}$ respectivement, ayant chacune une base ≥ 0 et n'**ayant** pas de masse ponctuelle **commune. Soit** B une application bilinéaire continue de $\vec{E}\times\vec{F}$ dans un espace de Banach $\vec{G}$.

Soient $\vec{M}$ **et** $\vec{N}$ **des intégrales indéfinies de** $\vec{\mu}$ **et** $\vec{\nu}$, à **valeurs dans** $\vec{E}$ **et** $\vec{F}$. **Si alors** $B(\vec{M}(x),\vec{N}(x))$ **a une** limite **quand** x **tend vers** b, **alors** $\int_a^{\to b} B(\vec{M},d\vec{\nu})$ **a un sens, si** et seulement si $\int_a^{\to b} B(d\vec{\mu},\vec{N})$ a un sens et l'on a la formule :

$$\int_a^{\to b} B(\vec{M},d\vec{\nu}) = \lim_{b'\to b} B\left(\vec{M}(b'),\vec{N}(b')\right) - B\left(\vec{M}(a),\vec{N}(a)\right) - \int_a^{\to b} B(d\vec{\mu},\vec{N}). \tag{IV,9;101}$$

(*) On utilise souvent une formule dite "seconde formule de la moyenne"; elle est un cas particulier de celle-ci.

La démonstration est immédiate, il suffit d'écrire la formule pour le couple a, b', et de faire tendre b' vers b .

Corollaire - Critère d'Abel pour la semi-convergence des intégrales impropres.

Soient $\vec{u}$ et $\vec{v}$ des fonctions définies sur l'intervalle $[a,b[$ de $\mathbb{R}$, $-\infty < a < b \leqslant +\infty$, à valeurs dans des espaces de Banach $\vec{E}$ et $\vec{F}$. Soit B une application bilinéaire continue de $\vec{E} \times \vec{F}$ dans un espace de Banach $\vec{G}$. On suppose que $\vec{u}$ et $\vec{v}$ ont les propriétés suivantes :

1°/ $\vec{u}$ est à variation bornée sur $[a,b[$, et $\vec{u}(x)$ converge vers 0 quand $x < b$ tend vers b .

2°/ $\vec{v}$ est localement dx-intégrable, à intégrales indéfinies bornées, autrement dit les quantités $\vec{\sigma}_{c,d} = \int_{|c,d|} \vec{v}(x)\,dx$ sont bornées en norme, indépendamment de c et de d . Alors, si l'espace $\vec{G}$ est de dimension finie * , l'intégrale $\int_{|a,\to b[} B(\vec{u}(x),\vec{v}(x))\,dx$ a un sens. En outre, si on appelle $U(c)$ la variation totale de $\vec{u}$ sur $]c,b[$, et si l'on appelle $V(c)$ la borne supérieure des quantités $\|\vec{\sigma}_{c,d}\|$ pour $d \in [c,b[$, on a la majoration

$$(\mathrm{IV}, 9; 102) \qquad \left\| \int_{|c,\to b[} B\big(\vec{u}(x),\vec{v}(x)\big)\,dx \right\| \leqslant \|B\| \; U(c)\,V(c), \qquad a \leqslant c < b .$$

Nous allons, pour simplifier, faire la démonstration dans le cas simple où $\vec{E}, \vec{F}, \vec{G}$, sont le corps $\mathbb{C}$, et B est la multiplication. On a une intégrale $\int_{|c,\to b[} u\,v\,dx$ On suppose u à variation bornée, $U(c) =$ variation totale de u dans] $c, b[$, et en outre on suppose que $u(x)$ tend vers 0 pour $x < b$ tendant vers b; on suppose v localement dx-intégrable, et à intégrales indéfinies bornées, $V(c) = \sup_{c \leqslant d < b} \left| \int_{|c,d[} v(x)\,dx \right|$. Alors $\int_{|c,\to b[} u\,v\,dx$ doit avoir un sens, et on doit avoir la majoration du

* Cette restriction est nécessaire si on veut utiliser le théorème 88, et en déduire que u est intégrale indéfinie d'une mesure $\vec{\mu}$ de base $\geqslant 0$. Mais, si l'on sait à l'avance qu'il en est ainsi (par exemple si $\vec{u}$ est de classe C^1 , corollaire 1 du théorème 89), cette restriction n'est plus nécessaire.

reste :

(IV,9;103) $$\left|\int_{]c,\to b[} u\, v\, dx\right| \leq U(c)\, V(c)$$

Comme nous allons le voir, il n'y a là rien d'autre qu'une intégration par parties.

Puisque u, que nous appellerons M, est à variation bornée, on a $dM = \mu$, mesure de Radon sur $[a, b[$, et en outre μ est de norme finie (théorème 88); plus précisément $U(c) = \int_{[c,b[} d|\mu|$. Considérons ensuite la mesure $v\,dx = \nu$, et soit N une intégrale indéfinie de? ; prenons par exemple $N(x) = \int_c^x v(t)\,dt$, de sorte que $N(c) = 0$; N est continue bornée, et $V(c) = \operatorname{Sup}_{c \leq x < b} |N(x)|$. Nous avons alors à calculer $\int_{]c,\to b[} M(x)\, dN(x)$; les mesures μ et ν n'ont sûrement pas de masse ponctuelle commune, puisque ν n'en a pas. On peut donc appliquer le théorème 98 : l'intégrale a un sens si et seulement la suivante en a un, et alors elles sont égales :

(IV,9;104) $$\left(\lim_{x \to b} M(x) N(x)\right) - M(a) N(a) - \int_{]a,\to b[} N(x)\, dM(x)$$

Or $M(x)$ tend vers 0 quand x tend vers b, et $N(x)$ est supposée bornée, donc $M(x)\,N(x)$ tend vers 0. En outre l'intégrale du 2ème membre est absolument convergente, puisque $\mu = dM$ est de norme finie et que N est bornée. Cela montre bien que l'intégrale étudiée est convergente (<u>mais non nécessairement absolument</u>), et, en remplaçant a par c, compte tenu de $N(c) = 0$, on a immédiatement

(IV,9;104 bis) $$\int_{]c,\to b[} u\, v\, dx = -\int_{[c,b[} N(x)\, d\mu(x) \text{, donc } \left|\int_{]c,\to b[} u\, v\, dx\right|$$

$$\leq \int_{[c,b[} |N(x)|\, d\,|\mu(x)| \leq V(c)\, U(c)\,,$$

ce qui montre aussitôt l'inégalité (IV,9;103).

Dans le cas encore plus particulier où $u = M$ est de classe C^1, la formule (IV,9;104) s'écrit

(IV,9;104 ter)
$$\int_{|a,\to b[} u\, v\, dx = \int_{|a,\to b[} M(x)\, v(x)\, dx$$
$$= \left(\lim_{x\to b} M(x) N(x)\right) - M(a) N(a)$$
$$- \int_{|a,b|} M'(x) V(x)\, dx\ ; \quad \text{et}$$
$$\int_{|a,\to c[} u\, v\, dx = - \int_{|a,c|} M' V\, dx \quad \text{donc}$$
$$\left| \int_{|a,\to c[} u\, v\, dx \right| \leq V(c) \int_{|a,c|} |M'(x)|\, dx = U(c) V(c).$$

Remarque Nous voyons ainsi, ce qui n'était pas apparu de façon aussi évidente dans le cas du théorème d'Abel pour les séries, que l'essentiel du théorème d'Abel n'est pas autre chose que la formule d'intégration par parties. Le théorème d'Abel sur les séries est d'ailleurs un cas particulier du théorème d'Abel sur les intégrales. Si en effet, étant donné les deux suites $\vec{u}_n$ et $\vec{v}_n$ dans les conditions du théorème 63 du chapitre II, on appelle $\vec{u}$ et 3 les fonctions définies sur la demi-droite $[0, +\infty[$ par

(IV,9;105)
$$\begin{cases} \vec{u}(x) = \vec{u}_n & \text{pour} \quad n \leqslant x < n+1\,, \\ \vec{v}(x) = \vec{v}_n & \text{pour} \quad n \leqslant x < n+1\,, \end{cases}$$

alors on voit que la suite des $\vec{u}_n$ est à variatinn bornée si et seulement si la fonction $\vec{u}$ sur $\mathbb{R}$ est à variation bornée, et que, pour la suite $\vec{v}_n$, les sommes partielles $\vec{\sigma}_{m,n}$ sont bornées si et seulement si les intégrales $\int_{|c,d|} \vec{v}(x)\, dx$ sont 'bornées; et alors le théorème que nous venons de démontrer redonne complétement le théorème 63 du chapitre II et la majoration (II,14;31), en prenant $c = m+1$.

Exemples d'application du critère d'Abel

Soit u une fonction définie sur une demi-droite $[a, +\infty[$, à valeurs complexes, a variation bornée, et convergeant vers 0 a l'infini.

Alors l'intégrale

(IV,9;106) $$\int_{|a \to +\infty[} e^{i\lambda x}\, u(x)\, dx$$

est convergente si λ est réel $\neq 0$, et en outre on a la **majoration**

(IV,9;107) $$\left\| \int_{|c, \to +\infty[} e^{i\lambda x}\, u(x)\, dx \right| \leq U(c)\, \frac{2}{|\lambda|}\,.$$

Si en effet ici on prend $v(x) = e^{i\lambda x}$, on voit que l'intégrale $\sigma_{c,d} = \int_{|c,d|} v(x)\, dx = \dfrac{e^{i\lambda d} - e^{i\lambda c}}{i\lambda}$ admet la majoration $|\sigma_{c,d}| \leqslant \dfrac{2}{|\lambda|}$. *

Naturellement le résultat ne subsiste pas pour. $\lambda = 0$.
De même si l'on considère les intégrales analogues

(IV,9;108) $$\int_{|a, \to +\infty[} \cos \lambda x\; u(x)\, dz, \qquad \int_{|a, \to +\infty[} \sin \lambda x\; u(x)\, d z\,,$$

le même raisonnement peut être fait. En outre la dernière a même un sens pour $\lambda = 0$ puisqu'alors elle est nulle ** .

* ☼♏♐♋♓⬧□■⬩ sur cet exemple l'intégration par parties, qui est la **démonstration** du **théorème** d'Abel. On a :

$\displaystyle\int_{[c,d[} e^{i\lambda x}\, u(x)\, dx = \left[\frac{e^{i\lambda x}\, u(x)}{i\lambda} \right]_c^d - \int_{[c,d[} \frac{e^{i\lambda x}}{i\lambda}\, du(x)$. En faisant tendre d vers $+\infty$, on volt bien qu'il y a une limite.

** Si u est réelle $\geqslant 0$, **décroissante**, et tend vers 0 **à** l'infini, le théorème des **séries** alternées redonnerait la convergence. Car $\displaystyle\int_{\left[\frac{n\pi}{\lambda}, \frac{(n+1)\pi}{\lambda}\right]} u(x) \sin\lambda x\, dx$ est du signe de $(-1)^n$, **décroît** en valeur **absolue quand** n **croît** (parce que $u\left(x + \frac{\pi}{\lambda}\right) \leqslant u(x)$), **et tend vers 0 pour** n **infini (cette intégrale est majorée en module par** $\frac{\pi}{\lambda}\, u\left(\frac{n\pi}{\lambda}\right)$) **. C'est donc le terme général d'une série convergente. Cela prouve que la suite des intégrales** $\int_{\left|a, \frac{n\pi}{\lambda}\right|}$ **a une limite pour** n **infini. Mais,**

pour $\dfrac{n\pi}{\lambda} \leqslant b' \leqslant \dfrac{(n+1)\pi}{\lambda}$, $\int_{|a, b'|}$ est compris entre $\int_{\left|a, \frac{n\pi}{\lambda}\right|}$ et $\int_{\left|a, \frac{(n+1)\pi}{\lambda}\right|}$, donc l'**intégrale** est bien convergente.

Dans les trois cas considérés, si la fonction u n'est pas intégrable sur $[a, +\infty[$, on a bien affaire à des intégrales impropres semi-convergentes.

Un cas particulièrement intéressant est celui des intégrales

$$(\text{IV},9;109)\qquad \int_{]0,\to+\infty[} \frac{e^{i\lambda x}}{x^\alpha}\,dx\,,\quad \int_{]0,\to+\infty[} \frac{\cos\lambda x}{x^\alpha}\,dx\,,\quad \int_{]0,\to+\infty[} \frac{\sin\lambda x}{x^\alpha}\,dx\,.$$

On ne se trouve pas immédiatement dans les conditions d'application du théorème d'Abel, parce que la fonction u , à cause de sa singularité à l'origine $x = 0$, n'est pas à variation bornée. Mais il suffira de séparer l'intégrale en somme de deux intégrales, $\int^1$ et $\int^{\to+\infty}$.

La première a un sens parce qu'on a affaire à une fonction intégrable-Lebesgue, si l'on suppose $\alpha < 1$ dans le cas des deux premières intégrales, $\alpha < 2$ dans le cas de la troisième.

Il reste donc à étudier l'intégrale $\int_1^{\to+\infty}$, et, pour celle-ci, nous sommes exactement dans les conditions d'application du théorème d'Abel, car $\frac{1}{x^\alpha}$ est décroissante et bornée, donc à variation bornée, et tend vers 0 à l'infini, dès que $\alpha > 0$.

Les intégrales considérées sont donc bien convergentes, les deux premières pour $0 < \alpha < 1$, la troisième pour $0 < \alpha < 2$. Remarquons qu'ici U est de classe C^∞ dans $]0, +\infty[$,et que nous nous trouvons en présence du cas le plus classique d'intégration par parties ! Ces intégrales jouent toutes un rôle très important en analyse et en physique, nous apprendrons plus tard à les calculer à partir d'une fonction nouvelle, la fonction Γ . Donnons dès maintenant deux d'entre elles :

$$(\text{IV},9;110)\qquad \int_{]0,\to+\infty[} \frac{\sin\lambda x}{x}\,dx = \frac{\pi}{2}\,,\ \text{pour } \lambda > 0\,.$$

$$(\text{IV},9;110\,bis)\qquad \int_{]0,\to+\infty[} \frac{\cos\lambda x}{\sqrt{x}}\,dx = \int_{]0,\to+\infty[} \frac{\sin\lambda x}{\sqrt{x}}\,dx = \sqrt{\frac{\pi}{2\lambda}}\,,\ \text{pour } \lambda > 0\,.$$

Théorème 99 (Changement de variable)

Soient $[\alpha,\beta[$ et $[a,b[$ deux intervalles bornés ou non de $\mathbb{R}$. Soit ξ une application de classe C^1 de $[\alpha,\beta[$ dans $[a,b[$, avec $\xi(\alpha) = a$, et $\lim\limits_{\beta'\to\beta} \xi(\beta') = b$ *. Dans ces conditions, l'intégrale $\int_a^{\to b} \vec{f}(x)\,dx$, où $\vec{f}$ est une fonction sur $[a,b[$, à valeurs dans un espace de Banach $\vec{F}$, localement dx-intégrable, est convergente, si et seulement si l'intégrale $\int_\alpha^{\to\beta} \vec{f}(\xi(t))\,\xi'(t)\,dt$ est convergente, et l'on a alors

$$\int_a^{\to b} \vec{f}(x)\,dx = \int_a^{\to\beta} \vec{f}(\xi(t))\,\xi'(t)\,dt \tag{IV, 9,111}$$

La démonstration est évidente, il suffit d'écrire la formule dans un intervalle $[\alpha, \beta'[$ (corollaire 2 du théorème **94)** et de faire tendre β' vers β **.

Exemple - Intégrales de Fresnel

On appelle ainsi les intégrales

$$\int_{[0,\to+\infty[} \cos\lambda x^2\,dx\ , \qquad \int_{[0,\to+\infty[} \sin\lambda x^2\,dx \tag{IV,9;112}$$

Il n'est pas du tout évident que ces intégrales soient convergentes, car les fonctions à intégrer ne convergent même pas vers 0 quand x tend vers $+\infty$. Cependant l'examen des courbes représentatives *** montre qu'elles sont très oscillantes et qu'il est parfaitement normal que les intégrales considérées soient convergentes. Il suffit

* Les positions relatives de α et β, ainsi que de a et b, sont arbitraires.

** On en déduit aisément des formules correspondant à (IV,9;71), quand ξ est monotone.

*** Les zéros de $\sin\lambda x^2$ sont les points $\lambda x^2 = n\pi$, ou $x = \sqrt{\frac{n\pi}{\lambda}}$, $n = 0,1,2,\ldots$. La distance entre deux zéros consécutifs, $\sqrt{\frac{\pi}{\lambda}}\left(\sqrt{n+1}-\sqrt{n}\right) \sim \sqrt{\frac{\pi}{\lambda}}\,\frac{1}{2\sqrt{n}}$, tend vers 0 pour n infini. Le théorème des séries alternées montrerait directement la convergence de ces intégrales (voir note ** page **653**).

en effet de faire le changement de variable $x = \sqrt{t}$ pour se ramener aux intégrales (IV,9;110 bis) antérieurement calculées. En admettant donc ces formules, qui, comme nous l'avons dit, seront démontrées plus tard, on a les résultats:

$$\int_{[0,\to+\infty[} \cos \lambda x^2 \, dx = \int_{[0,\to+\infty[} \sin \lambda x^2 \, dx = \frac{1}{2}\sqrt{\frac{\pi}{2\lambda}} . \qquad \text{(IV, 9;113)}$$

Valeur principale de Cauchy

Soit $|a,b|$ un intervalle de $\mathbb{R}$, c un point de $]a,b[$. Soit $\vec{f}$ une fonction sur $|a,b| - \{c\}$, à valeurs dans un espace de Banach $\vec{F}$, intégrable pour la mesure de Lebesgue sur $|a,c-\delta|$ et sur $|c+\delta',b|$, quels que soient δ et $\delta' > 0$, **mais** non nécessairement sur $|a,b|$. Quand on dit que $\int_{|a,b|} f \, dz$ est convergente, en tant qu'intégrale impropre, on veut dire que $\int_{|a,\to c[} \vec{f} \, dx$ et $\int_{]c\leftarrow,b|} \vec{f} \, dx$ sont toutes deux convergentes. Cela revient à dire que $\int_{|a,c-\delta|} + \int_{|c+\delta',b|}$ a une limite, lorsque δ et δ' tendent vers 0 indépendamment l'un de l'autre.

Il peut arriver que cette **limite** n'existe pas, mais qu'elle existe si $\delta' = \delta$, δ tendant vers 0. On dit alors <u>que l'intégrale de $\vec{f}$ est convergente en valeur principale de Cauchy,</u> et on note :

$$vp \int_{|a,b|} \vec{f(x)} \, dx = \lim_{\delta\to 0} \left(\int_{|a,c-\delta|} + \int_{|c+\delta,b|} \right).$$

<u>Théorème 100 - Pour que</u> $vp\int_{|a,b|} \vec{f(x)} \, dx$ <u>**existe,** il faut et il suffit que</u> $\vec{f}$ <u>soit, au voisinage de</u> $x = c$, <u>la somme d'une fonction antisymétrique</u> $\vec{f_1}$ $\left(\vec{f_1}(c+u) = -\vec{f_1}(c-u)\right)$, <u>et d'une fonction symétrique</u> $\vec{f_2}$ $\left(\vec{f_2}(c+u) = \vec{f_2}(c-u)\right)$ <u>telle que</u> $\int_{]c\leftarrow,b|} \vec{f_2}(x)dx$ <u>soit convergente.</u>

si $c = 0$, symétrique et antisymétrique s'appellent paire et impaire. Par le changement de variable $x = c + u$, on peut toujours se ramener à ce cas, nous supposerons donc $c = 0$. Toute fonction $\vec{f}$ est, d'une manière unique, somme d'une fonction paire et d'une fonction impaire, au voisinage de $x = 0$:

(IV, 9; 114) $$\vec{f}(x) = \frac{\vec{f}(x) + \vec{f}(-x)}{2} + \frac{\vec{f}(x) - \vec{f}(-x)}{2} = \vec{f}_1(x) + \vec{f}_2(x).$$

$\vec{f}_1$ étant impaire, son intégrale dans un intervalle symétrique est nulle, donc, si$q$ est choisi de façon que $a \leqslant -\alpha < 0 < \alpha \leqslant b$, on a :

(IV, 9; 115) $$\int_{|-\alpha,\delta|} \overrightarrow{f_1(x)}\,dx + \int_{|+\delta,+\alpha|} \overrightarrow{f_1(x)}\,dx = 0$$

Donc la convergence au sens $v\,p$ dans $|a, b|$, ou, ce qui revient au même dans $[-\alpha, +\alpha]$, est la même pour $\vec{f}$ et pour $\vec{f}_2$. Par ailleurs, $\vec{f}_2$ étant paire :

(IV, 9; 116) $$\int_{|-\alpha,-\delta|} \overrightarrow{f_2(x)}\,dx + \int_{|+\delta,+\alpha|} \overrightarrow{f_2(x)}\,dx = 2\int_{|\delta,\alpha|} \overrightarrow{f_2(x)}\,dx$$

de sorte que l'intégrale de $\vec{f}$ existe en $v\,p$, si et seulement si $\int_{]0\leftarrow,\alpha|} \overrightarrow{f_2(x)}\,dx$ est convergente, ce qui démontre le théorème.

Exemple : Supposons que $\vec{f}$ admette, au voisinage de $x = c$, un développement limité $\overrightarrow{f(x)} = \frac{\vec{c}_{-1}}{x-c} + \vec{c}_0 + \vec{c}_1(x-c) + \cdots$ l'intégrale existe en $v.\,p.$, puisque $\frac{1}{x-c}$ est antisymétrique.

Si d'ailleurs nous calculons $\int_{|c+\delta,b|} \vec{f}(x)\,dx$, il s'écrit $\vec{c}_{-1}(\mathrm{Log}(b-c) - \mathrm{Log}\,\delta) + \int_{|c+\delta,b|} \left(\vec{f}(x) - \frac{\vec{c}_{-1}}{x-c}\right)dx$; la seconde expression a une limite pour δ tendant vers 0 (et même $\vec{f}(x) - \frac{\vec{c}_{-1}}{x-c}$ est intégrable dans $|a, b|$) mais pas la première, à cause de $-\mathrm{Log}\,\delta$.

Si maintenant nous calculons $\int_{|a,c-\delta|} \vec{f}(x)\,dx$, il vaut $\vec{c}_{-1}(\text{Log}\,\delta - \text{Log}\,(c-a)) + \int_{|a,c-\delta|}(\vec{f}(x) - \frac{\vec{c}_{-1}}{x-c})dx$; la seconde expression a une limite quand δ tend vers 0, mais pas la première, à cause du terme $\text{Log}\,\delta$. Mais si on fait la somme $\int_{|a,c-\delta|} + \int_{|c+\delta,b|}$, les termes en $\pm\,\text{Log}\,\delta$ se détruisent, et la somme a une limite pour δ tendant vers 0, ce qui donne :

(IV,9; 117) $$v.p. \int_{|a,b|} \vec{f}(x)\,dx = \vec{c}_{-1}\,\text{Log}\left(\frac{b-c}{c-a}\right) + \int_{|a,b|}\left(\vec{f}(x) - \frac{\vec{c}_{-1}}{x-c}\right)dx$$

Il existe de même une notion de $v.p.$ pour l'intervalle $]-\infty,+\infty[$ si $\vec{f}$ est intégrable sur tout intervalle fini, au sens : $v.p. \int_{]-\infty,+\infty[} \vec{f}(x)\,dx = \lim_{B\to\infty} \int_{|-B,+B|}$. Pour que cette $v.p.$ existe, il faut et il suffit que $\vec{f} = \vec{f}_1 + \vec{f}_2$, $\vec{f}_1$ impaire, $\vec{f}_2$ paire, et que $\int_{|0,\to+\infty[} \vec{f}_2(x)\,dx$ soit convergente.

On fera aisément les calculs suivants :

(IV,9; 118) $$v.p. \int_{|a,b|} \frac{dx}{x} = \text{Log}\left|\frac{b}{a}\right|$$, pourvu que $a \neq 0$, $b \neq 0$.

(IV,9; 119) $$v.p. \int_{]-\infty,+\infty[} \frac{dx}{x} = 0$$ (fonction impaire)

<u>Remarque</u> On introduit aussi la valeur principale de Cauchy pour des séries où l'ensemble des indices est l'ensemble $\mathbb{Z}$ des entiers de signe quelconque. Si $\sum_{n=-\infty}^{+\infty} \vec{u}_n$ est une telle série, la notion de convergence n'est pas définie a priori puisque $\mathbb{Z} \neq \mathbb{N}$; seule existe a priori la notion de convergence commutative (équivalente à la convergence absolue si les $\vec{u}_n$ prennent leurs valeurs dans un espace de dimension finie (remarque 2°/ page 125)). <u>Toutefois, on dit généralement que la série</u> $\sum_{n=-\infty}^{+\infty} \vec{u}_n$ <u>converge, si chacune des 2 séries</u> $\sum_{n=0}^{\infty} \vec{u}_n$, $\sum_{n=0}^{\infty} \vec{u}_{-n}$ <u>converge</u>. Cela revient à dire que $\sum_{n=-N_1}^{n=+N_2} \vec{u}_n$ a une limite quand N_1 et N_2 tendent vers $+\infty$, indépendamment l'un de

l'autre. Mais on pourra considérer que $\sum_{n=-\infty}^{+\infty} \vec{u}_n$ converge en valeur principale de Cauchy, si $\sum_{n=-N}^{n=+N} \vec{u}_n$ converge, quand N tend vers $+\infty$. Si, par exemple, $\vec{u}_n = -\vec{u}_{-n}$, il en sera toujours ainsi, et la somme trouvée sera nulle.

On peut aussi introduire la valeur principale de Cauchy pour des intégrales multiples. Si $\vec{f}$ est une fonction sur $\mathbb{R}^n$, sommable dans tout ensemble $|x| \geq \delta > 0$, on dit que son intégrale est convergente en $v.p.$ de Cauchy, si $\iint\cdots\int_{|x|\geq\delta} \vec{f}(x)dx$ a une limite quand δ tend vers 0. Pour $n = 1$, cela redonne bien la $v.p.$ sur $\mathbb{R}$.

Voici un exemple où Intervient la valeur principale de Cauchy, et qui est particulièrement important en mécanique quantique.

Théorème 101. Soit $]a,b[$ un intervalle de $\mathbb{R}$, avec $a < 0 < b$ soit $\vec{f}$ une fonction définie sur $]a,b[$, à valeurs dans un espace de Banach F, et intégrable. Si $\vec{f}$ est dérivable à l'origine, ou, plus généralement, si la quantité $\frac{\vec{f}(x)-\vec{f}(0)}{|x|^\alpha}$, $\alpha > 0$, $x \neq 0$, reste bornée lorsque x tend vers 0, alors $\int_{]a,b[} \frac{\vec{f}(x)}{x}dx$ a un sens en valeur principale de Cauchy relativement à l'origine; en outre l'intégrale

$$\int_{]a,b[} \frac{\vec{f}(x)}{x+iy}\,dx \tag{IV,9;120}$$

converge, lorsque y tend vers 0 par valeurs de signe ε ($\varepsilon = \pm 1$), vers

$$v.p.\int_{]a,b[} \frac{\vec{f}(x)}{x}\,dx - \varepsilon i\pi\vec{f}(0) \tag{IV,9;121}$$

On se serait attendu à trouver le terme $v.p.\int$, mais non le 2ème terme $-\varepsilon i\pi \vec{f}(0)$. C'est pourtant justement la présence de ce 2ème terme qui est essentielle.

Démonstration Tout d'abord $\frac{\vec{f}(x)}{x+iy}$ est intégrable dans $]a,-A]$ et dans $[+A,+b[$, $A>0$, car elle est mesurable et majorée en norme par $\frac{1}{A}\|\vec{f}(x)\|$. Par ailleurs, lorsque y tend vers 0, $\int_{[A,+b[} \frac{\vec{f}(x)}{x+iy}dx$ converge vers $\int_{[A,+b[} \frac{\vec{f}(x)}{x}dx$, en vertu du théorème 35 de Lebesgue. Il reste donc à montrer que $v.p.\int_{[-A,+A]} \frac{\vec{f}(x)}{x}dx$ existe, et que $\int_{[-A,+A]} \frac{\vec{f}(x)}{x+iy}dx$ converge vers $v.p.\int_{[-A,+A]} \frac{\vec{f}(x)}{x}dx - \varepsilon i\pi \vec{f}(0)$.

On a alors la décomposition :

$$\int_{[-A,+A]} \frac{\vec{f}(x)}{x+iy}dx = \vec{f}(0)\int_{[-A,+A]} \frac{dx}{x+iy} + \int_{[-A,+A]} \frac{\vec{f}(x)-\vec{f}(0)}{x+iy}dx \,. \tag{IV,9;122}$$

Examinons d'abord le premier terme.

Supposons, pour fixer les idées, $\varepsilon = +1$.

Considérons, dans l'ouvert du plan complexe, complémentaire de la demi-droite $\{z = x+iy\,;\, y=0, x\leq 0\}$, la fonction Log_3 définie page 156. On pose $z = re^{i\varphi}$, avec $-\pi<\varphi<\pi$, et on pose $\mathrm{Log}\, z = \mathrm{Log}\, r + i\varphi$.

On sait qu'alors la fonction logarithmique, ainsi définie, est de classe C^1, et que sa dérivée est $\frac{1}{z}$; par conséquent, a fortiori, la fonction $x\to \mathrm{Log}_3(x+iy)$ est dérivable sur la droite, pour $y\neq 0$, et sa dérivée est la fonction $x\longrightarrow \frac{1}{x+iy}$. Il en résulte que la première intégrale de (IV,9;122) est égale à :

$$\vec{f}(0)\left[\mathrm{Log}\,(x+iy)\right]_{x=-A}^{x=+A} \tag{IV,9;123}$$

Le logarithme du module disparaît, parce que $|x+iA|=|x-iA|$; et (IV,9;123) vaut donc $\vec{f}(0)\,i(\varphi_2-\varphi_1)$, où l'on appelle φ_2 l'argument du point $A+iy$, et φ_1 l'argument du point $-A+iy$. Si alors nous faisons tendre y vers 0 , ces arguments tendent respectivement vers 0 et π , et par conséquent la première intégrale converge vers $-i\pi\vec{f}(0)$.

Considérons maintenant la deuxième intégrale.

Lorsque y tend vers 0 , la fonction $x \longrightarrow \dfrac{\vec{f}(x)-\vec{f}(0)}{x+iy}$ converge simplement vers la fonction $x \longrightarrow \dfrac{\vec{f}(x)-\vec{f}(0)}{x}$, sauf à l'origine où d'ailleurs cette dernière n'est pas définie. On peut donc dire qu'elle converge simplement presque partout vers cette fonction. Par ailleurs, elle est majorée en norme par la fonction $x \longrightarrow \dfrac{\|\vec{f}(x)-\vec{f}(0)\|}{|x|}$, qui est $\geqslant 0$ et intégrable, car majorée par $\dfrac{\|\vec{f}(x)-\vec{f}(0)\|}{|x|^{\alpha}}\,\dfrac{1}{|x|^{1-\alpha}} \leqslant$ constante $\times \dfrac{1}{|x|^{1-\alpha}}$, et $1-\alpha<1$. Le théorème de Lebesgue nous dit donc que le passage à la limite est légitime pour la deuxième intégrale, et qu'elle converge, pour y tendant vers 0, vers l'intégrale

$$\int_{]-A,+A[} \frac{\vec{f}(x)-\vec{f}(0)}{x}\,dx\ . \tag{IV,9;124}$$

Comme par ailleurs la fonction à intégrer est intégrable, son intégrale coïncide avec son intégrale en valeur principale de Cauchy; mais celle-ci est

$$\begin{aligned} &v.p.\int_{]-A,+A[} \frac{\vec{f}(x)}{x}\,dx - \vec{f}(0)\ v.p.\int_{]-A,+A[}\frac{dx}{x} \\ &= v.p.\int_{]-A,+A[} \frac{\vec{f}(x)}{x}\,dx\ , \end{aligned} \tag{IV,9;125}$$

parce que $\dfrac{1}{x}$ est impaire.

En additionnant alors les 2 limites trouvées pour les deux intégrales de (IV,9;122), on obtient bien le résultat cherché.

Si y tend vers 0 par valeurs < 0, φ_2 tend toujours vers 0, mais φ_1 vers $-\pi$, et (IV,9;123) tend vers $+i\pi$ 3).

Remarque Si on applique ceci à une fonction f à valeurs scalaires, continue, à support compact sur la droite réelle, on peut être tenté de dire que la mesure $\frac{dx}{x+iy}$, dépendant du paramètre y, converge vaguement, lorsque y tend vers 0 avec le signe ε, vers la mesure $v.p.\frac{dx}{x} - \varepsilon i\pi\delta$. Mais évidemment il est impossible d'employer ce langage, car la fonction $\frac{1}{x}$ n'est pas localement intégrable par rapport à dx, par conséquent $v.p.\frac{dx}{x}$ n'est pas une mesure de Radon. Par ailleurs la convergence précédente n'a pas lieu pour toutes les fonctions f continues à support compact, mais seulement pour celles qui sont dérivables ou vérifient une certaine condition restrictive à l'origine.

Mous verrons plus tard, dans l'étude des distributions, comment néanmoins on peut interpréter le passage à la limite précédent.

§ 10 INTÉGRALES MULTIPLES SUR $\mathbb{R}^n$ LONGUEURS, AIRES, VOLUMES, DANS LES ESPACES EUCLIDIENS AFFINES DE DIMENSION FINIE. CHANGEMENTS DE VARIABLES DANS LES INTÉGRALES MULTIPLES SUR

La formule (IV,9;72) est susceptible d'une généralisation aux intégrales multiples sur $\mathbb{R}^n$, relativement à la mesure

$$dx = dx_1\, dx_2 \cdots dx_n .$$

Théorème 102 - Soient $\mathcal{O}$ et Ω deux ouverts de $\mathbb{R}^n$, et ξ un homéomorphisme de Ω sur $\mathcal{O}$, de classe C^1 ainsi que son homéomorphisme réciproque. Soit $\vec{f}$ une fonction continue sur Ω, à valeurs dans un espace de Banach F et à support compact.

Alors on a la formule

$$\text{(IV,9;76)}^{(1)} \qquad \iint\cdots\int_{\mathcal{O}} \vec{f}(x)\,dx = \iint\cdots\int_{\Omega} \vec{f}(\xi(t))\,\left|\det(\xi'(t))\right|\,dt ,$$

* Le début de ce § 10 avait été placé initialement au § 9; le changement n'a été fait que très tardivement dans la rédaction. Mais nous n'avons pas voulu changer la numérotation des formules, à cause des références qui y sont faites ultérieurement. Aussi les formules du début de ce § continueront à être numérotées (IV,9;...). Mille excuses !

dans laquelle dx veut dire $dx_1\, dx_2 \ldots dx_n$, et dt veut dire $dt_1\, dt_2, \ldots dt_n$; quant à $\det \xi'(t)$, il veut dire le déterminant jacobien de ξ au point t.

Démonstration Le théorème ayant été démontré au théorème 94 pour $n = 1$, car alors le déterminant jacobien n'est autre que la dérivée, nous allons le démontrer par récurrence.

Supposons qu'il soit vrai pour une intégrale multiple d'ordre $n-1$, et démontrons le pour une intégrale multiple d'ordre $n \geqslant 2$.

Démonstration dans un cas particulier La fonction ξ est définie par un système de n fonctions $\xi_1, \xi_2, \ldots \xi_n$, de n variables $t_1, t_2, \ldots, t_n$, nous supposerons, dans ce cas particulier, que la fonction ξ_i est $\xi_i(t) = t_j$, et nous allons dans ce cas démontrer le théorème.
D'après le théorème de Fubini, on peut écrire

(IV,9;77) $$\iint\cdots\int_{\mathcal{O}} \vec{f}(x)\,dx = \int da \int\cdots\int_{\mathcal{O}(a)} \vec{f}(x_1, x_2, \ldots, x_{i-1}, a, x_{i+1}, \ldots x_n)\,dx_1\,dx_2 \ldots dx_{i-1}\,dx_{i+1} \ldots dx_n,$$

où $\mathcal{O}(a)$ est l'ensemble des points $(x_1, x_2, \ldots x_{i-1}, x_{i+1}, \ldots x_n)$ de $\mathbb{R}^{n-1}$ tels que $(x_1, x_2, \ldots, x_{i-1}, a, x_{i+1}, \ldots x_n) \in \mathcal{O}$. Cet ensemble $\mathcal{O}(a)$ est un ouvert de $\mathbb{R}^{n-1}$ (section de $\mathcal{O}$, par l'hyperplan $x_i = a$ Identifié à $\mathbb{R}^{n-1}$); et la fonction à intégrer sur $\mathcal{O}$ (a) a son support compact, car il est contenu-dans la section, par l'hyperplan, du support compact de $\vec{f}$ dans $\mathcal{O}$, qui est compacte.

Alors l'hypothèse de récurrence nous permet de transformer l'intégrale $(n-1)$-uple sur $\mathcal{O}(a)$ par un changement de variables. L'application $(t_1, t_2, \ldots, t_{j-1}, t_{j+1}, \ldots, t_n)$

$\longrightarrow (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$, avec

$x_k = \xi_k(t_1, t_2, \ldots, t_{j-1}, a, t_{j+1}, \ldots, t_n)$, $k = 1, 2, \ldots, i-1, i+1, \ldots, n$, est un homéomorphisme de classe C^1 ainsi que son homéomorphisme réciproque, de $\Omega(a)$ sur $\mathcal{O}(a)$, où $\Omega(a)$ est l'ouvert de $\mathbb{R}^{n-1}$, ensemble des $(t_1, t_2, \ldots, t_{j-1}, t_{j+1}, \ldots, t_n)$ tels que $(t_1, t_2, \ldots, t_{j-1}, a, t_{j+1}, \ldots, t_n) \in \Omega$.

Alors on a :

(IV,9;78)
$$\int\cdots\int_{\Theta(a)} \vec{f}(x)\, dx_1\, dx_2 \ldots dx_{i-1}\, dx_{i+1} \ldots dx_n$$

$$= \int\cdots\int_{\Omega(a)} \vec{f}(\xi(t)) \left| \frac{D(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)}{D(t_1, t_2, \ldots, t_{j-1}, t_{j+1}, \ldots, t_n)} \right| dt_1\, dt_2 \ldots dt_{j-1}\, dt_{j+1} \ldots dt_n$$

Il reste à porter cette expression dans (IV,9;77). Mais

(IV,9;79)
$$\frac{D(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)}{D(t_1, t_2, \ldots, t_{j-1}, t_{j+1}, \ldots, t_n)} = (-1)^{i+j} \frac{D(x_1, x_2, \ldots, x_n)}{D(t_1, t_2, \ldots, t_n)}$$

puis-que $x_i = t_j$, donc les modules de ces 2 déterminants sont les mêmes.

Finalement

(IV,9;80)
$$\iint\cdots\int_{\Theta} \vec{f}(x)\, dx =$$

$$\int da \int\cdots\int_{\Omega(a)} \vec{f}(\xi(t_1, t_2, \ldots t_{j-1}, a, t_{j+1}, \ldots t_n)) \left| \det \xi'(t_1, t_2, \ldots t_{j-1}, a, t_{j+1}, \ldots t_n) \right|$$

$$dt_1\, dt_2 \ldots dt_{j-1}\, dt_{j+1} \ldots dt_n = \iint\cdots\int_{\Omega} \vec{f}(\xi(t)) \left| \det \xi'(t) \right| dt\,,$$

d'après Fubini, et la formule (IV,9;76) est bien démontrée dans ce cas particulier.

<u>Démonstration dans le cas général</u>. Soit τ un point de Ω. Comme, en ce point, le déterminant jacobien $\det \xi'(t)$ est $\neq 0$ puisque ξ et ξ^{-1} sont de classe C^1 (corollaire 4 du théorème 11 du chapitre III), il existe au moins un indice i et un indice j tels que la dérivée partielle $\frac{\partial \xi_i}{\partial t_j}(\tau)$ soit $\neq 0$.

Considérons alors l'application θ_τ définie par les formules

(IV,9;81)
$$\begin{cases} y_l = t_l \quad , \quad l \neq j \\ y_j = \xi_i \quad (t_1, t_2, \ldots t_n) \end{cases}$$

Son déterminant jacobien au point7 n'est alors autre que la dérivée partielle $\frac{\partial \xi_i}{\partial t_j}$, par suite il est $\neq 0$. Donc, d'après le théorème 29 des fonctions implicites, il existe un ouvert Ω_τ contenant 7, tel que l'application θ_τ soit un homéomorphisme, de classe C^1 ainsi que son homéomorphisme réciproque, de Ω_τ sur un ouvert $\mathcal{U}_\tau$ de $\mathbb{R}^n$. Pour $y=(y_1, y_2, \ldots, y_n)$ dans $\mathcal{U}_\tau$, on peut résoudre les équations (IV,9;81) par $t_\ell = y_\ell$, $\ell \neq j$, et $t_j = \eta_j(y_1, y_2, \ldots, y_n)$, η_j de classe C^1. Soit $\mathcal{O}_\tau$ l'image de Ω_τ par ξ.
Lorsque τ varie, les ouverts $\mathcal{O}_\tau$ forment un recouvrement de $\mathcal{O}$ et en particulier un recouvrement du support compact K de $\vec{f}$; mais alors il existe nécessairement un nombre fini d'entre eux qui suffit à recouvrir K; appelons-les $\mathcal{O}_{\tau_i}$, $i \in I$.
Soit d'autre part γ_i une partition de l'unité relative à ce recouvrement fini de K. Nous pouvons écrire

$$(\text{IV},9;82) \qquad \iint\cdots\int_{\mathcal{O}} \vec{f}(x)\,dx = \sum_{i\in I} \iint\cdots\int_{\mathcal{O}_{\tau_i}} \gamma_i(x)\,\vec{f}(x)\,dx$$

Alors l'application ξ de Ω_{τ_i} dans $\mathcal{O}_{\tau_i}$ peut se décomposer en succession de deux C^1-difféomorphismes, à savoir $\xi = (\xi \circ \theta_{\tau_i}^{-1}) \circ \theta_{\tau_i}$.

Le premier, θ_{τ_i}, est défini par (IV,9;81); le suivant, $\zeta_i = \xi \circ \theta_{\tau_i}^{-1}$, peut s'écrire $(y_1, y_2, \ldots, y_n) \longrightarrow (x_1, x_2, \ldots, x_n)$, avec

$$(\text{IV},9;83) \qquad \begin{cases} x_k = \xi_k(y_1, y_2, \ldots, y_{j-1}, \ \eta_j(y_1, y_2, \ldots, y_n), y_{j+1}, \ldots, y_n), \ k \neq i; \\ x_i = y_j \end{cases}$$

On peut alors utiliser le cas particulier déjà étudié, pour l'homéomorphisme $\xi \circ \theta_{\tau_i}^{-1}$, car il est bien de la forme traitée dans ce premier cas.

On peut donc écrire la formule

$$(\text{IV},9;84) \qquad \iint\cdots\int_{\mathcal{O}_{\tau_i}} \gamma_i(x)\vec{f}(x)\,dx = \iint\cdots\int_{\mathcal{U}_{\tau_i}} \gamma_i(\zeta_i(y))\,\vec{f}(\zeta_i(y))\,\left|\det \zeta_i'(y)\right|\,dy\,.$$

Mais l'homéomorphisme θ_{τ_i} de Ω_{τ_i} sur $\mathcal{U}_{\tau_i}$ est lui aussi de cette forme, et même encore plus simple puisqu'il n'existe de changement que relativement à une seule des variables.

On a donc la formule

(IV,9;85) $$\iint\cdots\int_{\mathcal{O}_{\tau_i}} \gamma_i(x)\vec{f}(x)\,dx = \iint\cdots\int_{\Omega_{\tau_i}} \gamma_i(\xi(t))\vec{f}(\xi(t))\left|\det \zeta_i'(\theta_{\tau_i}(t))\right| \left|\det \theta'_{\tau_i}(t)\right| dt\,.$$

La formule de multiplication des déterminants jacobiens, corollaire 3 du théoréme 11 du chapitre II, nous montre alors que l'on a $\left|\det \zeta_i'(\theta_{\tau_i}(t))\right| \left|\det \theta'_{\tau_i}(t)\right| = \left|\det \xi'(t)\right|$, de sorte que

(IV,9;86) $$\iint\cdots\int_{\mathcal{O}_{\tau_i}} \gamma_i(x)\vec{f}(x)\,dx = \iint\cdots\int_{\Omega_{\tau_i}} \gamma_i(\xi(t))\vec{f}(\xi(t))\left|\det \xi'(t)\right| dt\,.$$

En sommant pour les diverses valeurs de i, on trouve bien (IV,9;76), et le théorème est complètement démontré.

<u>Remarque</u> Il est bien évident que l'hypothèse de bijectivité de ξ est indispensable, et que, par exemple, on a, dans le cas d'une variable :

(IV,9;87) $$\int_{]-1,+1[} 2|t|\,dt = 2 \neq 1 = \int_{[0,1[} dx\,,$$

parce que l'application $t \longrightarrow x = t^2$ n'est pas un homéomorphisme de $]-1,+1[$ sur $[0,1[$. Voir à ce sujet formule (IV,9;98 bis). Au contraire, dans le cas très particulier d'une variable démontré au théorème 93, il n'était pas nécessaire de supposer que ξ fût un homéomorphisme, et par exemple, on avait la formule :

(IV,9;88) $$\int_{-1}^{+1} 2t\,dt = \int_{+1}^{+1} dx = 0\,,$$

en posant $x = t^2$.

<u>Corollaire 1</u> - <u>Dans les conditions du théorème, l'image, par l'homéomorphisme</u> ξ, <u>de la mesure</u> $|\det \xi'(t)|\,dt$ <u>sur</u> Ω <u>est la mesure</u> $dx = 1(x)\,dx$ sur $\mathcal{O}$ * .

* Il peut paraître étrange d'employer des lettres différentes, x et t, d'autant plus que $\mathcal{O}$ et Ω sont tous les deux dans le même espace, $\mathbb{R}^n$. C'est simplement plus commode, à cause de la formule (IV,9;76), parce que ξ a été écrit : $t \longrightarrow x = \xi(t)$. Mais plus loin, au corollaire 4 et après le corollaire 5, nous utiliserons la même lettre. Voir note * page 429 .

Démonstration Si en effet on applique la formule (IV,9;76) à une fonction f scalaire continue à support compact, on tombe exactement sur la définition de l'image d'une mesure.

Remarque Naturellement, ξ étant un homéomorphisme, l'image de dx par ξ^{-1} est $|\det \xi'(t)|\, dt$. L'image de dt par ξ est $|\det (\xi^{-1})'(x)|\, dx$.

Conformément aux notations de la page 545, on se permettre d'écrire $dx = d(\xi(t)) = |\det \xi'(t)|\, dt$.

Corollaire 2 - Soit $\vec{f}$ une fonction définie sur $\mathcal{O}$, à valeurs dans un espace de Banach $\vec{F}$. Pour que $\vec{f}$ soit intégrable pour la mesure dx, il faut et il suffit que la fonction $t \to \vec{f}(\xi(t))\,|\det \xi'(t)|$ soit intégrable par rapport à la mesure dt, et alors on a la formule (IV,9;76) *.

Ceci est une considérable généralisation du théorème, dans lequel on supposait f continue à support compact.

Démonstration - Le corollaire 1 nous permet d'appliquer le théorème 60 : $\vec{f}(x)$ est dx-intégrable, si et seulement si $\vec{f}(\xi(t))$ est $(|\det \xi'(t)|\, dt)$-intégrable; et le théorème 51 nous dit qu'il en est ainsi, si et seulement si $\vec{f}(\xi(t))\,|\det \xi'(t)|$ est dt-intégrable.

Corollaire 3 - Soit A une partie de Ω, mesurable pour la mesure de Lebesgue de $\mathbb{R}^n$. Son image $\xi(A) \subset \mathcal{O}$ est mesurable pour la mesure de Lebesgue, et a pour mesure

$$\int_{\xi(A)} dx = \int_A |\det(\xi'(t))|\, dt .$$

* Cette fois-ci nous prenons les mesures dx et dt, c'est-à-dire la même mesure de $\mathbb{R}^n$.

Démonstration Il suffit d'appliquer (IV,9;76) **à** la fonction caractéristique de l'ensemble ξ **(A)** .

Corollaire 4 - La mesure de Lebesgue $dx = dx_1\, dx_2 \ldots dx_n$, sur l'espace vectoriel $\mathbb{R}^n$, est invariante par toute translation de $\mathbb{R}^n$; la mesure d'une partie **A** de $\mathbb{R}^n$ est égale à la mesure de ses translatées. Autrement dit, quel que soit l'élément $\vec{a}$ de $\mathbb{R}^n$, et la translation $\tau_{\vec{a}}$: $x \longrightarrow x + \vec{a}$, on a $\tau_{\vec{a}}(dx) = dx$.

Démonstration Le déterminant jacobien d'une translation, $\xi = \tau_{\vec{a}}$ ou $\xi(t) = t + \vec{a}$, est égal à 1 en tout point de $\mathbb{R}^n$.

Nous verrons ultérieurement que la mesure dx est, à un facteur près, la seule qui soit invariante par translation : Toute mesure de Radon complexe sur $\mathbb{R}^n$, invariante par translation. est de la forme $c\, dx$. où c est une constante complexe; plus géndralement, toute mesure de, Radon sur $\mathbb{R}^n$, à valeurs dans un espace vectoriel normé E, invariante par toutes les translations, est de la forme $\vec{c}\; dx$, où $\vec{c}$ est un vecteur fixe de $\vec{E}$.

Corollaire **5** - Soit ξ une application linéaire de $\mathbb{R}^n$ dans lui-même. Si alors A est un ensemble de $\mathbb{R}^n$ mesurable pour la mesure de Lebesgue de $\mathbb{R}^n$, son image ξ (A) est aussi mesurable pour la mesure de Lebesgue, et en outre (mesure de ξ (A)) $= |\Delta|$ (mesure de A), où Δ est le déterminant de l'application linéaire ξ .

Démonstration L'application linéaire a un déterminant jacobien constant, égal à A , et il suffit d'appliquer le corollaire 3.

Ce corollaire suppose que ξ soit un homéomorphisme, donc $A \neq 0$; mais le résultat subsiste si $\Delta = 0$, avec mes $(\xi(A)) = 0$, **car** $\xi(A)$ est contenu dans un sous-espace vectoriel de dimension $< n$ de $\mathbb{R}^n$ (voir par exemple le corollaire 1 du théorème 102 bis).

Nous voyons là une nouvelle interprétation tout-à-fait remarquable du module du déterminant d'une application linéaire : $|\Delta| = \dfrac{\text{mesure de } \xi(A)}{\text{mesure de } A}$, pour tout A mesurable pour la mesure de Lebesgue. En employant le langage du corollaire 1, on voit aussi que l'image par ξ de la mesure $|A|\,dx$ est la mesure dx, ou encore que l'image de la mesure dx par l'application linéaire ξ, si le déterminant $|\Delta|$ de ξ n'est pas nul, est la mesure $\frac{1}{|\Delta|}\,dx$. (Si A est un ensemble mesurable de $\mathbb{R}^n$ pour la mesure de Lebesgue, sa mesure pour dx doit être égale à la mesure de son image ξ (A) pour la mesure Image $\xi\,(dx)$, ce qui est exaet, puisque $\xi(dx)=\frac{1}{|\Delta|}\,dx$ et que la mesure de $\xi(A)$ pour dx est $|\Delta|$ fois celle de A).

Si ξ est une homothétie de rapport k, la mesure de ξ (A), pour la mesure de Lebesgue, est $|A| = |k|^n$ fois celle de A:

$$\text{mesure de } k\,A = |k|^n \times \text{mesure de } A\,,$$

<u>Corollaire 5 bis - La mesure de Lebesgue d'un-n **parallèlépipède * de** $\mathbb{R}^n$, de sommet a et défini par les vecteurs</u> $\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n$, <u>est le module du **déterminant** de ces vecteurs.</u>

soit $X_{i,j}$, $j=1,2,\ldots,n$, les coordonnées de $\vec{X}_i$. Le déterminant des vecteurs est, par définition, le déterminant des $X_{j,i}$

Considérons alors l'application linéaire ξ qui transforme $\vec{e}_i$, i-ième **vecteur** de la base canonique **de** $\mathbb{R}^n$, en $\vec{X}_i$; comme les $\vec{e}_i$ définissent un n- parallélépipède

* Dans un espace **affine** E sur le corps des réels, on **appelle** p**-parallélépipède de sommet** a, défini par les **vecteurs** $\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_p$, l'**ensemble** des **éléments de** E de **la forme** $a + t_1\vec{X}_1 + t_2\vec{X}_2 + \cdots + t_p\vec{X}_p$, avec $t_i \in [0,1], i=1,2,\ldots,p$. Rien n'oblige ces vecteurs à être indépendants; p **peut** donc dépasser la dimension de E. Dans le présent corollaire, on an vecteurs **dans** $\mathbb{R}^n$, s'ils sont dépendants, le corollaire s'applique encore, et la mesure du n-parallélépipède est nulle.

de volume 1 (ce parallélépipède est simplement le produit $[0,1] \times [0,1] \times \cdots \times [0,1]$, donc sa mesure est la puissance n-ième de la mesure de Lebesgue de $[0,1]$ sur $\mathbb{R}$, c'est-à-dire 1), le corollaire 5 dit que le volume cherché est $|\det \xi|$, qui est précisément le module du déterminant de la matrice des X_{ij} ou le module du déterminant des n vecteurs $\vec{X}_i$.

<u>Corollaire 6 - Soient</u> a <u>un point de</u> Ω, A <u>une partie de</u> Ω, <u>mesurable pour la mesure de Lebesgue. Alors le rapport des mesures de Lebesgue de</u> $\xi(A)$ <u>et de</u> A <u>converge vers</u> $|\det \xi'(a)|$, <u>lorsque l'ensemble</u> A <u>converge uniformément vers le point</u> a.

<u>Nous disons qu'un ensemble variable</u> **A** <u>converge uniformément vers</u> a, <u>si la borne supérieure de la distance de</u> a **<u>aux</u>** <u>points de</u> A <u>converge vers</u> 0.

<u>Démonstration</u> La différence $\varepsilon = \dfrac{\text{mesure de } \xi(A)}{\text{mesure de } \mathbf{A}} - |\det \xi'(a)|$

peut en effet s'écrire sous la forme

$$\varepsilon = \frac{\iint\cdots\int_A |\det(\xi'(t))|\,dt}{\iint\cdots\int_A dt} - |\det \xi'(a)| \tag{IV,9;89}$$

$$= \frac{\iint\cdots\int_A \big(|\det(\xi'(t))| - |\det(\xi'(a))|\big)\,dt}{\iint\cdots\int_A dt.}$$

Elle admet par conséquent la majoration :

$$|\varepsilon| \leqslant \sup_{t \in A} \Big| |\det(\xi'(t))| - |\det(\xi'(a))| \Big| , \tag{IV,9;90}$$

ce qui démontre le théorème, compte tenu de ce que ξ est supposée de classe C^1, et que par conséquent la fonction $t \longrightarrow |\det \xi'(t)|$ est continue.

<u>Remarque-</u>Ce corollaire donne une interprétation remarquable de $|\det(\xi'(a))|$. On le comprend intuitivement, pour la raison suivante. Soit $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$, la base canonique de $\mathbb{R}^n$. Son image par l'application dérivée $\xi'(a)$ est formée des vecteurs $\overrightarrow{\frac{\partial \xi}{\partial t_1}}(a)$, $\overrightarrow{\frac{\partial \xi}{\partial t_2}}(a), \ldots, \overrightarrow{\frac{\partial \xi}{\partial t_n}}(a)$. Si on prend

pour A un petit n-parallélépipède de sommet a, défini par les vecteurs $k\vec{e}_i$, k petit, son image est "approximativement" le n-parallélépipède de sommet ξ (a), défini par les vecteurs $k\frac{\partial\xi}{\partial t_i}(a)$. Le rapport des volumes de ces n-parallélépipèdes est exactement le module du déterminant des $\frac{\partial\xi}{\partial t_i}(a)$, c'est-à-dire $|\det(\xi'(a))|$.

Mais nous sommes là très loin d'une démonstration rigoureuse du corollaire **6.**

Exemple de changement de variables, calcul d'une intégrale en coordonnées polaires.

Occupons nous d'abord des coordonnées polaires planes. Nous avons affaire à une application $P:(r,\varphi)\longrightarrow(x,y)$ de $\mathbb{R}^2$ dans $\mathbb{R}^2$, étudiée au chapitre III, page **226** :

(IV,9;91) $$x = r\cos\varphi \;,\quad y = r\sin\varphi \,.$$

Le déterminant jacobien de P est égal à r.

Si par ailleurs nous restreignons l'application à l'ouvert Ω défini par $r > 0$, $0<\varphi<2\pi$, **P** est un homéomorphisme, de classe C^1 ainsi que son homéomorphisme réciproque, de Ω sur son image $\mathcal{O}$, qui est dans $\mathbb{R}^2$, le complémentaire de l'ensemble $y=0$, $x\geq 0$. Soit alors $\vec{f}$ une fonction $(x,y)\rightarrow\vec{f}(x,y)$, définie sur $\mathbb{R}^2$, à valeurs dans un espace de Banach $\vec{F}$. Elle sera intégrable sur $\mathbb{R}^2$, si et seulement si elle est intégrable sur $\mathcal{O}$, et alors ses intégrales sur $\mathbb{R}^2$ et sur $\mathcal{O}$ seront les mêmes, parce que le complémentaire de $\mathcal{O}$ est de mesure nulle dans $\mathbb{R}^2$, pour la mesure de Lebesgue.

L'après le corollaire 2, elle sera intégrable dans $\mathcal{O}$, si et seulement si la fonction $(r,\varphi)\longrightarrow\vec{f}(r\cos\varphi, r\sin\varphi)r$ est intégrable sur Ω, et alors on a la formule :

(IV,9;92) $$\iint_{\mathbb{R}^2}\vec{f}(x,y)\,dx\,dy = \iint_{\Omega}\vec{f}(r\cos\varphi\,,\; r\sin\varphi)\; r\,dr\,d\theta \,.$$

Naturellement, dans cette dernière formule, on peut remplacer l'intégrale sur Ω par l'intégrale sur $\overline{\Omega}$, c'est-à-dire sur l'ensemble défini par les inégalités larges $r\geq 0$, $0\leq\varphi\leq 2\pi$, parce que l'ensemble $\overline{\Omega}-\Omega=\dot{\Omega}$, défini par

$\{r=0, 0 \leq \varphi \leq 2\pi$, ou $r>0, \varphi = 0$, ou $r>0, \varphi = 2\pi\}$ est de mesure nulle pour la mesure $dr\, d\varphi$.

Naturellement il est souvent avantageux d'utiliser une méthode géométrique pour calculer le déterminant jacobien. Faisons-le précisément dans l'exemple des coordonnées polaires, dans l'espace à trois dimensions $\mathbb{R}^3$. Mous avons cette fois une application P de $\mathbb{R}^3$ dans $\mathbb{R}^3$ définie par les formules $m=(r,\theta,\varphi) \longrightarrow M=(x,y,z)$, avec :

(IV,9;93)
$$\begin{cases} x = r \sin\theta \cos\varphi \\ y = r \sin\theta \sin\varphi \\ z = r\cos\theta \end{cases}$$

On se restreindra ici à l'ouvert $\Omega = \{r,\theta,\varphi\,;\, r>0, 0<\theta<\pi, 0<\varphi<2\pi\}$; alors? est un C^1 -difféomorphisme de Ω sur son image, ensemble ouvert $\mathcal{O}$ de $\mathbb{R}^3$, complémentaire du demi-plan $\{x,y,z\,;\, y=0, x\geq 0\}$.

. Comme précédemment, $\complement\mathcal{O}$ est de mesure nulle, ainsi que $\Omega = \bar{\Omega} - \Omega$.

On voit alors que la dérivée partielle $\frac{\partial \vec{P}}{\partial r}$ au point $m=(r,\theta,\varphi)$, est le vecteur unitaire $\vec{i}$ porté par le rayon vecteur du point $M=(x,y,z)=P(m)$. La dérivée partielle $\frac{\overrightarrow{\partial P}}{\partial \theta}$ est le vecteur $r\vec{j}$, où $\vec{j}$ est le vecteur unitaire tangent au méridien de M , dans le sens des θ croissants. La dérivée partielle $\frac{\overrightarrow{\partial P}}{\partial \varphi}$ est le vecteur $r\sin\theta\,\vec{k}$, où $\vec{k}$ est le vecteur unitaire porté par le parallèle de M , dans le sens des Q croissants. On a donc, en notation différentielle, la formule

(IV,9;94)
$$\overrightarrow{dM} = P'(m).\overrightarrow{dm} = \vec{i}\,dr + r\vec{j}\,d\theta + r\sin\theta\,\vec{k}\,d\varphi$$

Alors le déterminant jacobien de P au point $m=(r,\theta,\varphi)$ est le déterminant du système des trois vecteurs $\frac{\partial \vec{P}}{\partial r}, \frac{\partial \vec{P}}{\partial \theta}, \frac{\partial \vec{P}}{\partial \varphi}$, par rapport aux vecteurs de la base de $\mathbb{R}^3$, il est donc égal au produit de $r^2 \sin\theta$ par le déterminant du système de trois vecteurs $\vec{i}, \vec{j}, \vec{k}$; et comme le trièdre formé par ces trois vecteurs unitaires est trirectangle, son déterminant (déterminant d'une matrice orthogonale) est égal à ± 1 *.

* Voir ce renvoi page 673.

Donc on a

$$\det P'(r,\theta,\varphi) = \pm r^2 \sin\theta \neq 0$$

La formule du changement de variables en coordonnées polaires dans l'espace à trois dimensions est donc

(IV, 9; 95)
$$\iiint_{\mathbb{R}^3} \vec{f}(x,y,z)\,dx\,dy\,dz$$

$$= \iiint_{\overline{\Omega}} \vec{f}(r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)\, r^2\sin\theta\, dr\, d\theta\, d\varphi .$$

Généralisation du théorème 102, et des corollaires 1, 2, 3, 6.

Le théorème 102 et les corollaires 1, 2, 3, restent vrais pourvu que ξ soit un homéomorphisme de classe C^1 ; l'homéomorphisme réciproque ξ^{-1} n'a pas besoin d'être de classe C' , autrement dit $\det \xi'(t)$ peut s'annuler pour certaines valeurs de $t \in \Omega$. Le corollaire 6 reste vrai si ξ est une application de classe C^1 , non nécessairement un homémorphisme. Nous admettrons ces résultats.

On peut alors en déduire :

Théorème 102 bis.

Soit ξ une application de classe C' d'un ouvert Ω de $\mathbb{R}^n$ dans $\mathbb{R}^n$. Si nous appelons A l'ensemble des points de Ω où le déterminant jacobien de ξ est nul, son image $\xi(A)$ est de mesure nulle dans $\mathbb{R}^n$ par rapport à la mesure de Lebesgue.

Renvoi de la page 672 -

* Cela revient à dire que le parallélépipède défini par $\vec{i}, \vec{j}, \vec{k}$, a le volume 1; bien sûr, "on le sait" , mais n'oublions pas qu'antérieurement à la théorie de l'intégration, jamais une théorie correcte de la mesure des volumes n'a été donnée ! C'est d'ailleurs ce qui sera fait bientôt, dans ce même paragraphe.

Démonstration_ Pour simplifier, bornons-nous au cas où $\Omega = \mathbb{R}^n$. Abrégeons par mes la mesure de Lebesgue d'un ensemble. Soit $a \in A$. D'après le corollaire 6, étendu comme il vient d'être dit à une application quelconque ξ de classe C^1, il existe, $\varepsilon > 0$ étant donné, une boule ouverte B_a, de centre a, telle que, pour toute partie B de B_a, mesurable pour la mesure de Lebesgue de $\mathbb{R}^n$, on ait

(IV,9;96) $$\text{mes}\ \xi(B) \leq \varepsilon\ \text{mes}\ B$$

Soit **K** un compact de **A**. Alors les $B_a, a \in K$, forment un recouvrement ouvert de **K** ; il en existe un sous-recouvrement fini, soit $B_1, B_2, \ldots, B_N$. Posons

$$B'_1 = K \cap B_1 \ , B'_2 = K \cap B_2 \cap \complement B'_1 \ , B'_3 = K \cap B_3 \cap \complement B'_1 \cap \complement B'_2 \ , \text{ etc...}$$

Alors $B'_1, B'_2, \ldots, B'_n$, sont encore mesurables comme intersections d'ensembles mesurables, **mais** en outre ils sont disjoints, et recouvrent K. **Alors** les $\xi(B'_1), \xi(B'_2), \ldots, \xi(B'_N)$ recouvrent $\xi(K)$.

On a donc :

(IV,9;97) $$\text{mes}\ \xi(K) \leq \sum_{i=1}^{N} \text{mes}\ \xi(B'_i) \leq \varepsilon \sum_{i=1}^{N} \text{mes}(B'_i) = \varepsilon\ \text{mes}\ K.$$

Comme ε est arbitraire, on a $\text{mes}\ \xi(K) = 0$

Remarquons alors que **A** est fermé comme image réciproque de $\{0\}$ par l'application continue $t \to \det \xi'(t)$ de $\mathbb{R}^n$ dans $\mathbb{R}$. Si alors on appelle K_m l'intersection de A et de la boule $\|\vec{x}\| \leq m$ de $\mathbb{R}^n$, K_m est un compact de A, **et** **A** est la réunion des K_m * ; chaque $\xi(K_m)$ est de mesure nulle, comme nous venons de le voir; alors $\xi(A)$, réunion d'une infinité dénombrable d'ensembles $\xi(K_m)$ de mesure nulle, est lui-même de mesure nulle.

* C'est là que l'hypothèse $\Omega = \mathbb{R}^n$ simplifie : elle permet de montrer immédiatement que A est réunion dénombrable de compacts . Mais c'est vrai dans le cas général.

Corollaire 1 - Soit V une variété de classe $\mathcal{C}^1$, dénombrable à l'infini. Soit H une application de classe $\mathcal{C}^1$ de V dans $\mathbb{R}^n$. Si la dimension de V est $\leqslant n-1$, son image $H(V)$ est de mesure nulle dans $\mathbb{R}^n$ pour la mesure de Lebesgue. En particulier, si V est une variété de classe $\mathcal{C}^1$ de dimension $\leqslant n-1$, de $\mathbb{R}^n$, elle est de mesure nulle.

Démonstration Soit ℓ la dimension de V, $\ell \leqslant n-1$ soit $\Phi : \mathcal{O} \longrightarrow \Phi(\mathcal{O})$ une carte de V. Alors l'application composée $H \circ \Phi$ est une application de classe $\mathcal{C}^1$ de $\mathcal{O}$ dans $\mathbb{R}^n$.

Elle peut être prolongée en une application de classe $\mathcal{C}^1$, de l'ouvert $\mathcal{O} \times \mathbb{R}^{n-\ell}$ de $\mathbb{R}^n$, dans $\mathbb{R}^n$, en posant

(IV,9;98) $$\xi(t,u) = (H \circ \Phi)(t), \quad \text{pour } t \in \mathcal{O},\ u \in \mathbb{R}^{n-\ell}$$

Cette application a un déterminant jacobien identiquement nul, car $\frac{\partial \xi}{\partial u} \equiv 0$ et par conséquent, d'après le théorème, l'image $H \circ \Phi(\mathcal{O}) = \xi(\mathcal{O} \times \mathbb{R}^{n-\ell})$ est de mesure de Lebesgue nulle. Cela signifie que l'image par H de l'image $\Phi(\mathcal{O})$ de la carte Φ est de mesure nulle. Si alors K est un compact de V, comme il peut être recouvert par un nombre fini d'images de cartes, son image par H est nécessairement de mesure nulle; et, comme V est supposée réunion dénombrable de compacts, son image par H est nécessairement de mesure nulle.

Remarque Il ne faudrait pas croire que ce résultat soit encore exact pour des variétés de classe $\mathcal{C}^0$, c'est-à-dire topologiques non différentiables. Si par exemple, nous considérons dans le plan ce qu'on appelle un arc de Jordan, c'est-à-dire un ensemble homéomorphe à un segment de droite $[a,b]$, si l'homéomorphisme H de $[a,b]$ sur cet arc est seulement continu mais non dérivable, la mesure de cet arc de Jordan pour $dx \otimes dy$ n'est pas nécessairement nulle. D'ailleurs une application continue H d'un segment de $\mathbb{R}$ dans $\mathbb{R}^n$ est un chemin de $\mathbb{R}^n$, au sens de la page 90; or un chemin, seulement continu, c'est-à-dire non dérivable, peut avoir pour image tout un cube de $\mathbb{R}^n$, de mesure > 0.

Corollaire 2 - Soit ξ une application de classe $\mathcal{C}^1$ d'un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ dans un ouvert $\mathcal{O}$ de $\mathbb{R}^n$.

Pour tout point x de $\mathcal{O}$, appelons $\nu(x)$ le nombre de points de l'image réciproque $\xi^{-1}(\{x\})$.

Alors ξ est propre relativement à la mesure $|\det \xi'(t)|\,dt$ de Ω, si et seulement si ν est localement dx-intégrable dans $\mathcal{O}$; dans ce cas, l'image par ξ de la mesure $|\det \xi'(t)|\,dt$ de Ω est la mesure $\nu(x)\,dx$ de $\mathcal{O}$. Dans ce cas, si f est une fonction sur $\mathcal{O}$, à valeurs dans un espace de Banach F, $\nu(x)\vec{f}(x)$ est dx-intégrable, si et seulement si $\vec{f}(\xi(t))\,|\det \xi'(t)|$ est dt-intégrable, et l'on a :

$$\iint\cdots\int_{\mathcal{O}} \overrightarrow{f(x)}\,\nu(x)\,dx = \iint\cdots\int_{\Omega} \overrightarrow{f(\xi(t))}\,|\det \xi'(t)|\,dt \,. \qquad (\text{IV},9,98^{\text{bis}})$$

En faisant $\nu = 1$, on retrouve les corollaires 1 et 2 du théorème 102.

Mesure des volumes dans un espace affine euclidien de dimension finie

Soit E un espace affine euclidien de dimension N. Il est évidemment très intuitif qu'il existe dans un tel espace une mesure des volumes, encore faut-il le demontrer en toute rigueur.

Soit $O, \vec{e}_1, \vec{e}_2, \ldots, \vec{e}_N$, un référentiel orthonormal de E.

On peut alors, par ce référentiel, identifier E à $\mathbb{R}^N$, et définir ainsi la mesure produit tensoriel $dx = dx_1 \otimes dx_2 \cdots \otimes dx_N$.

Pour parler un langage plus correct, on doit dire que le choix du référentiel définit une bijection linéaire P : $(x_1, x_2, \ldots, x_N) \to a + x_1\vec{e}_1 + \cdots + x_N\vec{e}_N$ de $\mathbb{R}^N$ sur E, et que, dans ces conditions, on peut définir la mesure image $P(dx)$, par P, de la mesure de Lebesgue $dx = dx_1\,dx_2 \cdots dx_N$ de $\mathbb{R}^N$. L'abus de langage consiste à écrire encore dx au lieu de $P(dx)$

Théorème 103 (Théorème fondamental de la mesure des volumes)

Soit E un espace affine euclidien de dimension finie N. La mesure définie, dans un référentiel orthonormal, par $dx = dx_1\,dx_2 \ldots dx_N$, est indépendante du référentiel choisi.

Démonstration Si a, $\vec{e}_1$, $\vec{e}_2$, ..., $\vec{e}_N$ est un premier référentiel orthonormal, b, $\vec{f}_1$, $\vec{f}_2$, ..., $\vec{f}_N$, un deuxième, les coordonnées x et y d'un même point dans ces deux référentiels sont liées par la formule de changement de coordonnées

(IV,10;1)[1]
$$\begin{cases} y_i = a_i + \sum_{j=1}^{N} m_{i,j}\, x_j \,, \\ a_i = (\overrightarrow{a-b} \mid \vec{f}_i) \,, \quad m_{i,j} = (\vec{e}_j \mid \vec{f}_i) \end{cases}$$

Alors, d'après le corollaire 1 du théorème 102 de changement de variables dans les intégrales multiples, on a $dy = |\Delta|\, dx$ ** , où Δ est le déterminant de la matrice.716 des $m_{i,j}$. Mais le déterminant d'une matrice312 orthogonale • ** est égal à ± 1, car cette matrice vérifie $\mathcal{M}\,{}^t\mathcal{M} = I$, ${}^t\mathcal{M}$ étant la transposée et I la matrice identique, d'où $\Delta^2 = 1$; sa valeur absolue est donc égale à 1, par conséquent on a $dx = dy$, et le théorème est démontré.

Ainsi, dans un espace affine euclidien, existe une mesure de Radon ≥ 0 bien déterminée; nous l'écrirons dx. Cette <u>mesure est d'ailleurs invariante par les translations et plus généralement par les déplacements de l'espace</u>, c'est-à-dire par toutes les applications affines inversibles conservant les produits scalaires: On peut lui appliquer les corollaires 5 et 5 bis du théorème 102, puisqu'elle est égale à $dx_1\, dx_2 \ldots dx_N$ pour un référentiel orthonormal.

* La numérotation normale des formules, avec (IV,10;...), commence ici.

** Avec l'abus de langage qui vient d'être indiqué. En réalité, les deux référentiels définissent 2 bijections P et Q de $\mathbb{R}^N$ sur E , et on a $P = Q \circ \mathcal{M}$, où $\mathcal{M}$ est l'application de $\mathbb{R}^N$ sur $\mathbb{R}^N$ définie par (IV,10;1). Alors $P(dx) = Q(\mathcal{M}(dx)) = Q(|\Delta^{-1}|\, dy)$ ou $|\Delta|\, P(dx) = Q(dy)$ que nous écrivons $|\Delta|\, dx \check{=} dy$.

*** Eprouver la plus grande méfiance à l'égard des démonstrations de cette propriété du déterminant des matrices orthogonales, basées sur la notion géométrique de volume ! La notion de volume n'est pas une donnée de la providence, elle doit être introduite en toute rigueur ..... par la théorie de l'intégration ! Il faut donc une démonstration algébrique de la propriété $\Delta = \pm 1$.

Si alors A est un ensemble quelconque de E , supposé mesurable, sa mesure par rapport à dx est ce qu'on appelle son volume dans l'espace euclidien. Si A est un parallélépipède rectangle de côtés $a_1, a_2, \ldots, a_N$ son volume est $a_1 a_2 \ldots a_N$, comme on le voit en choisissant un référentiel orthonormal, dont les vecteurs de base sont parallèles à ses côtés.

Théorème 103 bis - *Si $\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_N$, sont N vecteurs de même origine, le volume V du N-parallélépipède qu'ils définissent est donné par la formule*:

$$V^2 = \begin{vmatrix} (\vec{X}_1 | \vec{X}_1) & (\vec{X}_1 | \vec{X}_2) & \cdots & (\vec{X}_1 | \vec{X}_N) \\ (\vec{X}_2 | \vec{X}_1) & (\vec{X}_2 | \vec{X}_2) & \cdots & (\vec{X}_2 | \vec{X}_N) \\ \cdots & \cdots & \cdots & \cdots \\ (\vec{X}_N | \vec{X}_1) & (\vec{X}_N | \vec{X}_2) & \cdots & (\vec{X}_N | \vec{X}_N) \end{vmatrix} \tag{IV,10;2}$$

Démonstration Prenons en effet un référentiel orthonormal. Nous savons que V^2 est le carré du déterminant des N vecteurs (corollaire 5 bis du théorème 102); il suffit alors de faire le produit du déterminant par lui-même, suivant la règle du produit des déterminants, pour trouver la formule précédente.

Cette formule est évidemment plus avantageuse que celle qui utilise le déterminant des n vecteurs par rapport à un référentiel, parce qu'elle utilise seulement les produits scalaires, et est par suite indépendante de tout référentiel.

Mesure des longueurs dans un espace affine euclidien

Un espace affine euclidien est normé, donc on peut y mesurer les longueurs des courbes (voir page 618). Si $t \longrightarrow M(t)$ est une courbe paramétrique de classe C^1, t parcourant un intervalle $\mathbb{R}_1$ de $\mathbb{R}$, la longueur de cette courbe est $\int_{\mathbb{R}_1} \| \overrightarrow{M'(t)} \| \, dt$ (formule (IV,9;40 bis)), qui vaut aussi $\int_{\mathbb{R}} (\overrightarrow{M'(t)} | \overrightarrow{M'(t)})^{\frac{1}{2}} dt$. Si les x_i sont les composantes de M dans un référentiel orthonormal, cela donne $\int_{\mathbb{R}_1} \sqrt{\sum_{i=1}^{N} x_i'^2(t)} \, dt$.

Mesure des aires n - dimensionnelles dans une variété linéaire de dimension n d'un espace affine euclidien de dimension finie

Nous avons donc pu définir, dans un espace affine euclidien de dimension finie, d'une part la longueur des arcs de courbes. d'autre part la mesure des volumes. mais il existe une possibilité de définir tous les intermédiaires : des aires de surfaces et plus généralement des aires n-dimensionnelles de variétés à n dimensions, $0 \leqslant n \leqslant N$.

Occupons nous d'abord du cas où F est une variété linéaire de dimension n de l'espace affine E . Alors F est elle-même un espace affine euclidien, et par suite il existe dans F une mesure des volumes. Si on choisit dans F un référentiel orthonormal, qui l'identifie à $\mathbb{R}^n$, cette mesure est $dx_1\, dx_2 \ldots dx_n$. Mais comme F est un espace à n dimensions, nous n'emploierons pas le mot volume, mais le mot d'aire n-dimensionnelle , ou n-aire . Si F_1 et F_2 sont deux variétés linéaires de même dimensions n , toute application affine de F_1 sur F_2 , qui conserve les produits scalaires, transporte un référentiel orthonormal de F_1 sur un référentiel orthonormal de F_2 , donc conserve les n-aires. Toute application affine qui est une similitude de rapport k, c'est-à-dire qui multiplie les longueurs par k et les produits scalaires par k^2 , a un déterminant jacobien $\pm k^n$ par rapport à des référentiels orthonormaux de F_1 et F_2 , donc multiplie les n-aires par $|k|^n$.

Si $\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n$ sont n vecteurs de même origine de F , l'aire n-dimensionelle dun-parallélépipède qu'ils définissent est donnée par la formule

$$(IV,10;3) \qquad S^2 = \begin{vmatrix} (\vec{X}_1 | \vec{X}_1) & (\vec{X}_1 | \vec{X}_2) & \ldots & (\vec{X}_1 | \vec{X}_n) \\ (\vec{X}_2 | \vec{X}_1) & (\vec{X}_2 | \vec{X}_2) & \ldots & (\vec{X}_2 | \vec{X}_n) \\ \ldots & \ldots & \ldots & \ldots \\ (\vec{X}_n | \vec{X}_1) & (\vec{X}_n | \vec{X}_2) & \ldots & (\vec{X}_n | \vec{X}_n) \end{vmatrix}$$

$$= \underset{i,j}{\text{dét}} (\vec{X}_i | \vec{X}_j)$$

Exercice - Il est parfois utile d'écrire la formule précédente sous une autre forme. Supposons par exemple qu'il s'agisse de deux vecteurs X et Y, alors la formule précédente s'écrit :

(IV,10;4) $$S^2 = (\vec{X} \mid \vec{X})\,(\vec{Y} \mid \vec{Y}) - (\vec{X} \mid \vec{Y})^2,$$

qui, d'après l'identité de Lagrange, peut aussi s'écrire

(IV,10;5) $$\sum_{i,j} (X_i Y_j - X_j Y_i)^2,$$

où les X_i (resp. Y_i) sont les coordonnées de $\vec{X}$ (resp. $\vec{Y}$), par rapport à un référentiel orthonormal quelconque.

Ainsi S^2 n'est autre que la somme des carrés des aires des parallèlogrammes, projections orthogonales du parallélogramme initial sur les plans engendrés par les divers couples de 2 vecteurs de la base. Cette formule s'étend aux aires n-dimensionnelles, pour n quelconque.

Théorème 104 - L'aire n-dimensionnelle d'un n-parallèlépipède est égale au produit de l'aire $(n-1)$-dimensionnelle de l'une quelconque de ses bases par la longueur de la hauteur correspondante.

Démonstration -Le n-parallélépipède est défini par son sommet et n vecteurs $\vec{X_1}, \vec{X_2}, \dots, \vec{X_n}$. Supposons d'abord ces n vecteurs dépendants; alors l'airen-dimensionnelle du parallélépipède est nulle; mais, dans ce cas, ou bien les $n-1$ vecteurs $X_1, X_2, \dots, X_{n-1}$, sont eux aussi dépendant& alors l'aire $(n-1)$-dimensionnelle de la base est nulle, ou bien ils sont indépendants, mais alors $\vec{X_n}$ est dans le sous-espace vectoriel F_n engendré par $\vec{X_1}, \vec{X_2}, \dots, \vec{X_{n-1}}$, et c'est la hauteur qui est nulle; dans les deux cas, l'énoncé est bien exact.

Supposons donc les n vecteurs indépendants. Considérons la base * du parallélépipède, définie par le sommet, et $\vec{X_1}, \vec{X_2}, \dots, \vec{X_{n-1}}$. Prenons une base orthonormale de $\vec{F}$, dont les $n-1$ premiers éléments soient dans le sous-espace vectoriel $\vec{F_n}$ engendré par $\vec{X_1}, \vec{X_2}, \dots, \vec{X_{n-1}}$; le n-ième est un vecteur unitaire $\vec{e_n}$ normal à $\vec{F_n}$. La n - aire du n-parallélépipède est le module du déterminant des n vecteurs par rapport à cette base. On peut écrire $\vec{X_n} = \vec{\xi_n} + \vec{Y_n}$, où $\vec{\xi_n}$, projection orthogonale de $\vec{X_n}$ sur $\vec{F_n}$, est une

* Le mot base a 2 sens différents dans cette démonstration: base d'un parallélépipède, base d'un espace vectoriel.

combinaison linéaire de $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$, et où $\vec{Y}_n = (\vec{X}_n \mid \vec{e}_n)\vec{e}_n$ a pour longueur $|(\vec{X}_n \mid \vec{e}_n)|$; $\vec{Y}_n$ est précisément la hauteur du parallélépipède, associée à la base choisie. Le déterminant de $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n$ est alors aussi le déterminant de $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}, \vec{Y}_n$; mais, dans celui-ci, la dernière ligne est $0, 0, 0, \dots 0, (\vec{X}_n \mid \vec{e}_n)$; donc le déterminant est égal au produit du déterminant de $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$, dont le module est la $(n-1)$-aire de la base, par $(\vec{X}_n \mid \vec{e}_n)$, dont le module est la longueur de la hauteur $\vec{Y}_n$, ce qui démontre le théorème.

Corollaire 1 - L'aire n-dimensionnelle d'un n-parallélépipède défini par n vecteurs $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n$, est majorée par le produit des longueurs de ces vecteurs; elle est strictement majorée par ce produit, sauf si les vecteurs sont 2 à 2 orthogonaux.

Démonstration C'est évident pour $n = 1$. Supposons le corollaire vrai pour $n - 1$ vecteurs, démontrons le pour n.

La n-aire du n-parallélépipède est le produit de la $(n-1)$-aire de la base engendrée par $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$, par la longueur $\|\vec{Y}_n\|$ de la hauteur $\vec{Y}_n$, la première est majorée par $\|\vec{X}_1\| \, \|\vec{X}_2\| \cdots \|\vec{X}_{n-1}\|$, d'après l'hypothèse de récurrence, et strictement majorée si $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$, ne sont pas 2 à 2 orthogonaux; la 2ème est majorée, par $\|\vec{X}_n\|$, et strictement majorée si $\vec{X}_n$ n'est pas orthogonal à $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_{n-1}$, ce qui démontre le corollaire.

Corollaire 2 - Soit $(X_{i,j})_{\substack{1 \le i \le n \\ 1 \le j \le n}}$ un tableau carré de nombres réels à n lignes et n colonnes. Le déterminant Δ de ce tableau admet la majoration :

$$(\text{IV},10;5\,bis) \quad |\Delta| \le \sqrt{X_{1,1}^2 + X_{1,2}^2 + \cdots + X_{1,n}^2}\ \sqrt{X_{2,1}^2 + X_{2,2}^2 + \cdots + X_{2,n}^2} \cdots \sqrt{X_{n,1}^2 + X_{n,2}^2 + \cdots + X_{n,n}^2}$$

Il suffit de traduire analytiquement le corollaire 1.

Remarque - Si les $X_{i,j}$ sont complexes, cette majoration subsiste, à condition évidemment de remplacer les $X_{i,j}^2$ par $|X_{i,j}|^2$; mais elle est plus délicate à démontrer. La majoration (IV,10;5 bis) a été démontrée par J. Hadamard, et a trouvé une application spectaculaire dans la théorie des équations intégrales de Fredholm.

Théorème 105 (projection orthogonale des aires hyperplanes).

Soient F et G des hyperplans dans l'espace euclidien affine E de dimension N. L'aire $(N-1)$-dimensionnelle de la projection orthogonale sur G d'un ensemble A mesurable de F est le produit de l'aire de A par le cosinus de l'angle aigu de F et de G.

Démonstration On peut supposer F et G non parallèles sans quoi le théorème est évident. Soit H l'intersection de F et de G; sa dimension est $N-2$. Soient $\vec{f}$ et $\vec{g}$ les vecteurs unitaires de $\vec{F}$ et $\vec{G}$ perpendiculaires à $\vec{H}$, et choisis de façon à faire entre eux un angle aigu θ, qui est par définition l'angle de F et de G (cet angle est aussi celui des normales à F et G). On a donc $(\vec{f}\,|\,\vec{g}) = \cos\theta \geq 0$. Choisissons dans H un référentiel orthonormal $0, \vec{h}_1, \vec{h}_2, \dots, \vec{h}_{N-2}$. Alors $0, \vec{h}_1, \vec{h}_2, \dots, \vec{h}_{N-2}, \vec{f}$, est un référentiel orthonormal de F, et $0, \vec{h}_1, \vec{h}_2, \dots, \vec{h}_{N-2}, \vec{g}$ un référentiel orthonormal de G. Ces référentiels identifient ces espaces à $\mathbb{R}^{N-1}$.

La projection de F sur G est une application affine laissant fixe 0 donc définie par une matrice par rapport à ces référentiels. Les projections de $\vec{h}_1, \vec{h}_2, \dots, \vec{h}_{N-2}$ sont ces vecteurs eux-mêmes, la projection de $\vec{f}$ est $\vec{g}\cos\theta$, donc la matrice de la projection est diagonale, et les éléments de sa diagonale sont $1,1,1,\dots\dots \cos\theta$; le déterminant **A est** $\cos\theta$, alors le corollaire 5 du théorème 102 donne le résultat.

Aire n-dimensionnelle d'une variété paramétrique de dimension n

Cherchons maintenant à définir l'aire n-dimensionnelle d'une variété de classe C^1 à n dimensions de l'espace euclidien affine E. Nous ne ferons pas ici une théorie générale analogue à celle que nous avons faite d'une part pour les longueurs et d'autre part pour les volumes. Nous nous occuperons seulement du cas d'une variété de classe C^1.

1°/ **soit** $Y : \mathcal{O} \to Y(\mathcal{O})$ une application de classe C^1 d'un ouvert $\mathcal{O}$ de $\mathbb{R}^n$ dans E .

L'application dérivée $\Psi'(u)$ au pointu de $\mathbb{R}^n$ transforme le système des n vecteurs de base de $\mathbb{R}^n$ dans le système des n vecteurs $\frac{\partial \vec{\Psi}}{\partial u_i}(u)$ de $\vec{E}$, $i = 1, 2, \dots, n$.

Ce système de n vecteurs définit un parallélépipède, dont l'aire n-dimensionnelle $D(u)$ est donnée par la formule

$$(IV,10;6) \qquad D(u) = \sqrt{\det_{i,j}\left(\frac{\partial \vec{M}}{\partial u_i}(u) \,\Big|\, \frac{\partial \vec{M}}{\partial u_j}(u)\right)}.$$

Alors, par définition, l'aire n -dimensionnelle de la variété paramétrique de classe C^1 de dimension n définie par Ψ est donnée par l'intégrale

$$(IV,10;7) \qquad S = \int\cdots\int_{\mathcal{O}} D(u)\,du = \int\cdots\int_{\mathcal{O}} D(u_1, u_2, \dots, u_n)\,du_1\,du_2 \dots$$

La raison d'être de cette définition est claire. Partageons $\mathcal{O}$ en petits parallélépipèdes de côtés parallèles aux axes. L'image par Ψ du parallélépipède de sommet u , de côtés parallèles aux axes, de longueurs $du_1, du_2, \dots, du_n$, est "approximativement" le parallélépipède de sommet $\Psi(u)$, défini par les vecteurs $\frac{\partial \vec{\Psi}}{\partial u_1}(u)\,du_1$, $\frac{\partial \vec{\Psi}}{\partial u_2}(u)\,du_2, \dots, \frac{\partial \vec{\Psi}}{\partial u_m}(u)\,du_n$, dont l'aire est $D(u)\,du_1\,du_2 \dots du_n$; l'aire cherchée est "la somme de ces aires élémentaires", d'où l'intégrale (IV,10;7).

Théorème 106 - L'aire n -dimensionnelle de la variété paramétrique Ψ , définie par (IV,10;7), ne change pas si l'on remplace la variété par une variété équivalente. Autrement dit. si Ψ_1 et Ψ_2 sont des applications de classe C_1 de deux ouverts $\mathcal{O}_1$ et $\mathcal{O}_2$ de $\mathbb{R}^n$ dans E , et s'il existe un C^1-difféomorphisme H de $\mathcal{O}_1$ sur $\mathcal{O}_2$, de manière que $\Psi_1 = \Psi_2 \circ H$, alors on a l'égalité

$$(IV,10;8) \qquad \int\cdots\int_{\mathcal{O}_1} D_1(u)\,du = \int\cdots\int_{\mathcal{O}_2} D_2(u)\,du.$$

Démonstration Soit $\alpha_1 \in \mathcal{O}_1$, $\alpha_2 = H(\alpha_1) \in \mathcal{O}_2$, $x = \Psi_1(\alpha_1) = \Psi_2(\alpha_2)$.
Le sous-espace vectoriel engendré par les n vecteurs $\frac{\partial \vec{\Psi}_1}{\partial u_i}(\alpha_1)$ est de dimension $\leq n$, et $= n$, si ces vecteurs sont indépendants. Appelons de toutes façons F un sous-espace à n dimensions qui le contienne. Naturellement ce sous-espace est le même que celui qui est engendré par les vecteurs $\frac{\partial \Psi_2}{\partial u_i}(\alpha_2)$, puisque l'application linéaire $\Psi_1'(\alpha_1)$ est la composée $\Psi_2'(\alpha_2) \circ H'(\alpha_1)$, et que $H'(\alpha_1)$ est une bijection (Corollaire 4 du théorème 11 du chapitre III). Choisissons dans cet espace vectoriel F une base orthonormale arbitraire, qui l'identifie ainsi à $\mathbb{R}^n$.

Si alors nous posons $A_1 = D_1(\alpha_1)$, et $A_2 = D_2(\alpha_2)$, le facteur A_1 , qui est l'aire du parallélépipède engendré par les $\frac{\partial \Psi_1}{\partial u_i}(\alpha_1)$, n'est autre que le module du déterminant de l'application linéaire $\Psi_1'(\alpha_1)$ de $\mathbb{R}^n$ dans $\mathbb{R}^n$, et le facteur A_2 est le module du déterminant de l'application linéaire $\Psi_2'(\alpha_2)$ de $\mathbb{R}^n$ dans $\mathbb{R}^n$.

Alors la relation $\Psi_1'(\alpha_1) = \Psi_2'(\alpha_2) \circ H'(\alpha_1)$ nous montre que l'on a entre les déterminants la relation

$$(IV,10;9) \qquad D_1(\alpha_1) = \left|\det H'(\alpha_1)\right| D_2(\alpha_2) = \left|\det H'(\alpha_1)\right| D_2\left(H(\alpha_1)\right) .$$

Mais alors, ceci étant vrai pour tout α_1 de $\mathcal{O}_1$ la formule (IV,9;76) du changement de variables dans les intégrales multiples nous montre que l'on a exactement

$$(IV,10;10) \qquad \int\cdots\int_{\mathcal{O}_2} D_2(\alpha_2)\, d\alpha_2 = \int\cdots\int_{\mathcal{O}_1} D_2\left(H(\alpha_1)\right) \left|\det H'(\alpha_1)\right| d\alpha_1$$

$$= \int\cdots\int_{\mathcal{O}_1} D_1(\alpha_1)\, d\alpha_1 ,$$

et que par conséquent les deux intégrales définissant les aires n-dimensionnelles pour Ψ_1 et pour Ψ_2 sont bien égales.

Exemple - Cas des surfaces dans un espace affine euclidien à trois dimensions. On a coutume, dans ce cas, d'appeler u, v, les coordonnées du paramètre de $\mathcal{O} \subset \mathbb{R}^2$, x, y, z, les 3 composantes de Y sur un référentiel orthonormal de E , A , B , C les trois déterminants jacobiens

$$(\text{IV},10;11) \quad \begin{cases} A = \dfrac{\partial y}{\partial u}\dfrac{\partial z}{\partial v} - \dfrac{\partial z}{\partial u}\dfrac{\partial y}{\partial v}, \\[2ex] B = \dfrac{\partial z}{\partial u}\dfrac{\partial x}{\partial v} - \dfrac{\partial x}{\partial u}\dfrac{\partial z}{\partial v}, \\[2ex] C = \dfrac{\partial x}{\partial u}\dfrac{\partial y}{\partial v} - \dfrac{\partial y}{\partial u}\dfrac{\partial x}{\partial v}, \end{cases}$$

E , F , G les quantités

$$(\text{IV},10;12) \quad \begin{cases} E = \left\| \dfrac{\overrightarrow{\partial M}}{\partial u} \right\|^2 = \left(\dfrac{\partial x}{\partial u}\right)^2 + \left(\dfrac{\partial y}{\partial u}\right)^2 + \left(\dfrac{\partial z}{\partial u}\right)^2, \\[2ex] F = \left(\dfrac{\overrightarrow{\partial M}}{\partial u} \middle| \dfrac{\overrightarrow{\partial M}}{\partial v}\right) = \dfrac{\partial x}{\partial u}\dfrac{\partial x}{\partial v} + \dfrac{\partial y}{\partial u}\dfrac{\partial y}{\partial v} + \dfrac{\partial z}{\partial u}\dfrac{\partial z}{\partial v}, \\[2ex] G = \left\| \dfrac{\overrightarrow{\partial M}}{\partial v} \right\|^2 = \left(\dfrac{\partial x}{\partial v}\right)^2 + \left(\dfrac{\partial y}{\partial v}\right)^2 + \left(\dfrac{\partial z}{\partial v}\right)^2, \end{cases}$$

et H ce que nous avons appelé D précédemment. D'après (IV,10;4) on a

$$(\text{IV},10;13) \quad H = \sqrt{EG - F^2},$$

et, d'après (IV,10;5), on a

$$(\text{IV},10;14) \quad H = \sqrt{A^2 + B^2 + C^2}$$

2°) Revenons à la variété paramétrique V définie par Ψ .

L'expression $D(u)\,du$ définit une mesure de Radon dS sur $\mathcal{O}$. Si A est un ensemble de $\mathcal{O}$, la mesure de cet ensemble, par rapport à cette mesure de Radon, peut être considérée comme l'aire n -dimensionnelle de la restriction de Ψ à cet ensemble * . Si Ψ est propre relativement à cette mesure dS, alors l'image par Y de cette mesure de Radon est une mesure $\Psi(dS) \geq 0$ sur l'espace affine E . Généralement on appellera indifféremment dS la mesure sur $\mathcal{O}$, ou son image. Elle sert à définir des intégrales de surface. Si $\vec{f}$ est une fonction définie sur E , à valeurs dans un

* Comme pour la longueur d'un chemin, on ne définit pas ici l'aire d'un ensemble, mais d'une application Y d'une variété (ou d'une partie A d'une variété) dans E .

espace de Banach $\vec{F}$, ce qu'on appellera $\int\cdots\int_{\vec{F}} dS$ *, c'est par définition

$$(\text{IV},10;15) \qquad \int\cdots\int \vec{f}\, dS = \int\cdots\int_{\mathcal{O}} (\vec{f}\circ\Psi)\, dS = \int\cdots\int_{\mathcal{O}} \overrightarrow{f(\Psi(u))}\, D(u)\, du\,.$$

C'est l'intégrale de $\vec{f}$ par rapport a la mesure image $\Psi(dS)$, si cette dernière existe, Si par ailleurs nous reprenons la situation du théorème 106, nous voyons que, si nous appelons dS_1 et dS_2 les mesures définies sur $\mathcal{O}_1$ et $\mathcal{O}_2$ par le procédé précédent, relativement aux variétés équivalentes Ψ_1 et Ψ_2, alors dS_2 n'est autre que l'image $H(dS_1)$ de dS_1 par l'homéomorphisme H ; c'est ce qu'indique en effet le corollaire 1 du théorème 102. Alors $\Psi_1(dS_1)$ et $\Psi_2(dS_2)$ représentent la même mesure sur E (si elles existent), puisque

$$\Psi_1(dS_1) = (\Psi_2\circ H)(dS_1) = \Psi_2(H(dS_1)) = \Psi_2(dS_2)\,.$$

3°) Nous sommes maintenant en mesure de donner la définition correcte de l'aire n-dimensionnelle d'une variété paramétrique quelconque de classe C^1 et de dimension n, dans un espace affine euclidien E de dimension N.

Soit V une variété (abstraite ou plongée dans un espace affine), de classe C^1, de dimension n. Soit H une application de classe C^1 de V dans F, définissant ainsi une variété paramétrique ou singulière de classe C^1 et de dimension n.

Soit Φ une carte d'un ouvert $\mathcal{O}$ de V. Alors $\Psi = H\circ\Phi$ est une application de classe C^1 de $\mathcal{O}$ dans E, et on peut ainsi définir une aire n-dimensionnelle dS_Φ, mesure ≥ 0 dans $\mathcal{O}$. Son image par l'homéomorphisme Φ est une mesure $\Phi(dS_\Phi)$ sur l'ouvert $\Phi(\mathcal{O})$ de V. SI maintenant nous considérons deux cartes de ce type, soit Φ_1 et Φ_2, et si les ouverts $\Phi_1(\mathcal{O}_1)$ et $\Phi_2(\mathcal{O}_2)$ ont une intersection Ω non

*) Comme pour l'intégrale de courbe (note *, page 622), il faudrait écrire $\int\cdots\int_{(\Psi)} \vec{f}\, dS$, intégrale relativement à la variété paramétrique (Ψ). D'autre part, le symbole d'intégrale multiple $\int\cdots\int$ n'est nullement justifié, Puisqu'il ne s'agit Pas de l'intégrale par rapport à un Produit tensoriel de mesures; c'est simplement un usage plutôt regrettable, indiquant que c'est une intégrale sur une variété de dimension > 1

vide, nous pouvons appeler Ω_1 et Ω_2 les images réciproques de Ω par Φ_1 et Φ_2 . Alors Φ_1 et Φ_2, restreintes à Ω_1 et Ω_2 définissent des variétés paramétriques équivalentes, puisque $H_{1,2} = \Phi_2^{-1} \circ \Phi_1$ est un homéomorphisme de classe C^1 (corollaire 1 du théorème 33 du chapitre III), et que $\Phi_1 = \Phi_2 \circ H_{1,2}$ La remarque faite précédemment montre alors que la mesure dS_{Φ_2} sur $\mathcal{O}_2$ est l'image de la mesure dS_Φ sur $\mathcal{O}_1$ par $H_{1.2}$, et que les Images de ces mesures, c'est-à-dire les mesures $\Phi_1(dS_{\Phi_1})$ et $\Phi_2(dS_{\Phi_2})$, dans Ω , coïncident.

Considérons alors le système de tous les ouverts $\Phi(\mathcal{O})$, images de toutes les cartes Φ sur V ; dans chacun de ces ouverts, nous avons une mesure réelle ≥ 0 $\Phi(dS_\Phi)$, et, dans l'intersection Ω de deux de ces ouverts $\Phi_1(\mathcal{O}_1)$, $\Phi_2(\mathcal{O}_2)$, les mesures $\Phi_1(dS_{\Phi_1})$, $\Phi_2(dS_{\Phi_2})$, définies par les deux cartes correspondantes Φ_1 et Φ_2 coïncident. Nous sommes donc dans la situation du théorème du recollement des morceaux de mesures (théorème 17), et nous pouvons affirmer l'existence d'une mesure de Radon bien déterminée sur V, qui, dans toute Image de carte $\Phi(\mathcal{O})$, coïncide avec l'aire n-dimensionnelle $\Phi(dS_\Phi)$ définie par cette carte. C'est cette mesure de Radon sur V qu'on appelle la mesure des aires n-dimensionnelles sur la variété singulière, et qu'on appellera dS. Si A est une partie de V, ce qu'on appellera l'aire n-dimensionnelle de A (ou de la restriction de H à A), c'est, par définition, la mesure de A pour dS. Si alors il se trouve que dS est de norme finie, sa norme, c'est-a-dire l'intégrale $\int \cdots \int_V dS$, est l'aire n-dimensionnelle de la variété paramétrique; cette aire ne change pas si l'on remplace la variété paramétrique par une variété équivalente. Si l'application H est dS- propre, il existe une mesure image $H(dS)$, notée encore souvent dS , sur E, et dont le support est dans l'adhérence $\overline{H(V)}$ de l'image $H(V)$. Si $\vec{f}$ est une fonction sur E , à valeurs dans un espace de Banach $\vec{F}$, on pourra parler de l'intégrale $\int \cdots \int \vec{f}\, dS$, *

qui est, par définition $\int \cdots \int_V (\vec{f} \circ H)\, dS$. C'est l'intégrale de $\vec{f}$ par rapport à $H(dS)$, si cette mesure image existe.

* ou mieux $\int \cdots \int_{(H)} \vec{f}\, dS$, intégrale relativement à la variété paramétrique définie par l'application H de V dans E. Noter que nous avons introduit 3 sortes de mesures ≥ 0: dS_Φ sur $\mathcal{O}$; $\Phi(dS_\Phi)$ sur $\Phi(\mathcal{O})$, le recollement des $\Phi(dS_\Phi)$ donnant dS sur V ; enfin $H(dS)$ dans E .

On peut résumer les résultats précédents dans le théorème :

Théorème 107 - Soit $H : V \longrightarrow E$, une application de classe C^1, d'une variété V de classe C^1 de dimension n, dans un espace affine euclidien E de dimension N , définissant une variété paramétrique de E . Pour toute carte $\Phi : \mathcal{O} \longrightarrow \Phi(\mathcal{O})$ de V , o n la dS_Φ mesure $D(u)\,du \geqslant 0$ sur $\mathcal{O}$, définie par (IV,10;6). Il existe alors sur V une mesure $dS \geqslant 0$ et une seule, qui, dans chaque $\Phi(\mathcal{O})$, coïncide avec $\Phi(dS_\Phi)$; on l'appelle la mesure des aires I-L -dimensionnelles ou des n-aires sur la variété paramétrique. Si A est une partie (dS-mesurable) de V, $\int \cdot\cdot \int_A dS$ est l'aire de A ; en particulier $\int \cdot\cdot \int_V dS = \|dS\|$ est l'aire de la variété paramétrique.

Ces résultats s'appliquent naturellement si V est une variété contenue dans E , avec H = identité.

Remarques : 1°/ Si V est un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, on peut, pour carte Φ , prendre l'identité, et on a directement $dS = dS_\Phi = D(u)\,du$. C'est d'ailleurs ce cas que nous avons traité d'abord, avant le théorème 106 (à ce moment, V et H s'appelaient $\mathcal{O}$ et Ψ).

2°/ Si $n = 1$, si V est un intervalle ouvert $\mathbb{R}_1$ de la droite réelle $\mathbb{R}$, et si H est une application notée M de $\mathbb{R}_1$ dans E , on se trouve dans le cas de la remarque 1°/ ci-dessus. Alors $D(t)$ (nous écrivons t au lieu de u pour la variable de $\mathbb{R}_1$) est la longueur de $\overrightarrow{M'}(t)$, et ainsi $ds = D(t)\,dt = \|\overrightarrow{M'}(t)\|\,dt$. On retombe bien sur la longueur (formule (IV,9;40)). Mais la longueur d'une courbe peut se définir dans un espace affine normé quelconque et même dans un espace métrique quelconque (page 618) alors que les aires n -dimensionnelles, $n > 1$, n 'ont été définies que dans les espaces affines euclidiens.

3°/ si $n = N$, et si V est un ouvert Ω de E , avec H = identité, la mesure des aires N- dimensionnelles est simplement la mesure des volumes dx . En effet, on se trouve dans les conditions de la remarque 1°/, et si on prend $\Phi = H =$ identité, on a $D(u) = 1$.

4°/ Si Ω est un ouvert de $\mathbb{R}^N$, ξ une application C^1 de Ω dans E, elle définit une variété paramétrique de classe C^1 de E ; la N-aire d'une partie A de Ω peut être appelée son volume, comme dans 3°/. Mais naturellement, ce n est pas nécessairement le volume de son image $\xi(A)$, car ξ n'est pas necessairement injective. Si on choisit dans E un référentiel orthonormal, qui l'identifie à $\mathbb{R}^N$, ce volume est toujours, d'après (IV,10;7), $\iint\cdots\int_A |\text{dét.}\, \xi'(t)|\, dt$; nous avons bien vu au corollaire 3 du théorème 102 que c'est le volume de $\xi(A)$, si ξ est un homéomorphisme. Dans le cas général, si nous appelons, pour chaque point x de E, $\nu(x)$ le nombre de points de Ω dont l'image par ξ est x, $0 \leqslant \nu(x) \leqslant +\infty$, on a toujours exactement, d'après (IV,9;98 bis),

(IV,10;15 bis) $$\iint\cdots\int_A \left|\text{dét}\, \xi'(t)\right| dt = \iint\cdots\int_{\xi(A)} \nu(x)\, dx\,.$$

C'est le "volume de $\xi(A)$, en comptant chaque point autant de fois qu'il est recouvert".

5°/ supposons $n = 0$. Une variété de dimension 0 de E est un ensemble de points isolés de E. On est amené à définir l'aire 0-dimensionnelle d'une telle variété comme étant simplement le nombre (fini ou infini) de ses points. L'aire 0-dimensionnelle d'un point est égale a 1.

Corollaire 1 - Une similitude de rapport k multiplie les n-aires par $|k|^n$.

Soit V une variété de classe C', de dimension n, et $H : V \to E$ une variété paramétrique de classe C' de E. Si ξ est une similitude de rapport k de E, alors $\xi \circ H$ est une nouvelle variété paramétrique, ayant toujours V comme variété des paramètres. Comme ξ multiplie par $|k|^n$ les n-aires des n-parallèlépipèdes, on voit immédiatement que le $D(u)$ du relatif à $\xi \circ H$ est $|k|^n$ fois celui qui est relatif à H, donc que la mesure dS sur V relative à $\xi \circ H$ est bien $|k|^n$ fois celle qui est relative a H, d'où le résultat.

Corollaire 2 - Soit V une variété vraie de dimension n, de classe C^1, d'un espace affine euclidien E. Si V est fermée, la n-aire de toute-partie bornée dS -mesurable A de V est finie. Si V est fermée et bornée, son aire est finie.

En effet, l'adhérence $\bar{A}$ de A relativement à V est alors fermée dans V, elle-même fermée dans E, donc elle est fermée dans E, et bornée par hypothèse; elle est alors compacte (théorème 23 du chapitre II).

Donc sa mesure est forcément finie.

Corollaire 3 - Soit $H : V \to E$ une variété paramétrique de classe C^1 de dimension n, d'un espace affine euclidien E. Si V 'est dénombrable à l'infini, et si W est une sous-variété de V, de classe C^1, de dimension $< n$, son n-aire est nulle.

Ce corollaire généralise le corollaire 1 du théorème 102 bis.

Démonstration Soit en effet $\Phi : \mathcal{O} \longrightarrow \Phi(\mathcal{O})$ une carte de V. Comme Φ est un C^1-difféomorphisme, l'image réciproque par Φ de la varidté-intersection de W et de $\Phi(\mathcal{O})$ est une variété de $\mathcal{O}$, de classe C', de dimension $< n$ (corollaire 2 du théorème 32 du chapitre III); alors cette image réciproque est de mesure nulle pour du (corollaire du théorème 102 bis), donc aussi pour $D(u)\,du = dS_\Phi$; donc, comme Φ est un homéomorphisme, $W \cap \Phi(\mathcal{O})$ est de n-aire nulle dans V. Comme tout compact K de V est recouvert par un nombre fini d'images de cartes telles que $\Phi(\mathcal{O})$, $W \cap K$ est encore d'aire nulle. Comme enfin V est réunion dénombrable de compacts, W est réunion dénombrable d'ensembles de n-aire nulle, donc est aussi de n-aire nulle,

Exemple 1 - Aire des sphères de rayon R dans un espace affine euclidien E à 3 dimensions.

Prenons un référentiel orthonormal, Identifiant E à $\mathbb{R}^3$. Les coordonnées polaires (IV,9;93) définissent une carte Φ, ai $\mathcal{O}$ est l'ouvert $0 < \theta < \pi$, $0 < \varphi < 2\pi$ de $\mathbb{R}^2$, et si on pose

$$\text{(IV,10;16)} \quad \begin{cases} x = R \sin\theta \operatorname{coq} \\ y = R \sin\theta \sin\varphi \\ z = R \cos\theta \end{cases}$$

$\Phi(\mathcal{O})$ est l'ouvert de la sphère, complémentaire du demi-méridien $y = 0,\ x \geqslant 0$. Comme ce demi-méridien est la réunion d'un demi-cercle sans extrémités, variété de classe C^1 de dimension 1. et des 2 pôles, variétés de classe C^1 et de dimension 0 ,il est d'aire nulle; donc l'aire de la sphère est l'aire de $\Phi(\mathcal{O})$; et, si A est une partie quelconque de la sphère, son aire est l'aire de son intersection avec $\Phi(\mathcal{O})$; ainsi la seule carte Φ est entièrement suffisante pour l'étude des aires sphériques, $\Phi(\mathcal{O})$ est dS-presque toute la variété. C'est la une circonstance qui se produit très souvent dans la pratique. D'après les calculs faits page 672 , on connaît $\overrightarrow{\partial\Phi}/\partial\theta$ et $\overrightarrow{\partial\Phi}/\partial\varphi$; ils sont orthogonaux, et de longueurs R et $R \sin\theta$, donc le parallélogramme qu'ils définissent est un rectangle d'aire $D(\theta,\varphi) = R^2 \sin\theta$. L'aire d'une partie A de la sphère est donc

(IV,10;17) $$S(A) = \iint_{\Phi^{-1}(A)} R^2 \sin\theta \, d\theta \, d\varphi \ .$$

En particulier, l'aire de la sphère est

(IV,10;18) $$S = \iint_{\substack{0<\theta<\pi \\ 0<\varphi<2\pi}} R^2 \sin\theta \, d\theta \, d\varphi$$

$$= R^2\left(\int_{]0,\pi[} \sin\theta \, d\theta\right)\left(\int_{]0,2\pi[} d\varphi\right) = 4\pi R^2 \ .$$

Nous étudierons plus tard, l'aire de la sphère dans un espace euclidien de dimension quelconque.

Exemple 2 - Calcul des angles solides dans un espace affine euclidien.

Soit $H : V \to E$ une hypersurface paramétrique, c'est-à-dire une variété paramétrique de dimension N -1, de classe C' , dans un espace affine E de dimension N . Soit O un point, et proposons nous de définir "l'angle solide sous lequel du point O on voit l'hypersurface" * . Nous allons ici définir

* C'est là une notion très intuitive : le résultat (IV,10;25) se voit tout de suite. Nous l'établissons ici en toute rigueur, à titre d'exemple typique; dans la pratique, on sait qu'on peut faire comme nous le faisons ici, et on va beaucoup plus vite !

un angle solide absolu et non algébrique, ce sera toujours un nombre ≥ 0 . Par ailleurs nous devrons toujours supposer que O n'appartient pas à l'image de la variété V (si jamais il y appartient, nous appellerons V_0 le complémentaire dans V de l'image réciproque $H^{-1}(\{O\})$; $H^{-1}(\{O\})$ est fermé, donc V_0 est ouvert, donc V_0 est encore une variété, et la restriction de H à V_0 est encore une variété paramétrique dont l'image ne contient pas O . L'angle solide relatif à V sera alors, par définition, l'angle solide relatif à V_0). Soit M un point quelconque de $E - O$. La demi-droite OM rencontre alors la sphère unité de centre O en un point m. On définit ainsi une application $\xi : M \to m = \xi(M)$ de $E - O$ dans E , qui est manifestement de classe C^1 , et dont l'image est la sphère unité de centre O . Si l'on considère alors l'application $\xi \circ H$, elle définit une nouvelle variété paramétrique de classe C^1 , où V est toujours la variété des paramètres; l'image de cette variété paramétrique est contenue dans la sphère unité de centre O . Alors, par définition, l'angle solide sous lequel du point O on voit la variété paramétrique définie par H est l'aire $(N-1)$ - dimensionnelle de la variété paramétrique définie par $\xi \circ H : V \longrightarrow E$.

D'ailleurs la mesure des aires associée à la variété paramétrique $\xi \circ H$ est une mesure $d\omega \geq 0$ sur V , qu'on appellera mesure des angles solides, et on pourra parler de la mesure pour ω d'une partie A de V; l angle solide total de la variété paramétrique étant $\|d\omega\| = \int_V d\omega \leq +\infty$.

Naturellement, si V est simplement un ouvert de la sphère unité, avec $H=$ identité, alors cet angle solide est l'aire de V elle-même. En particulier, si V est toute la sphère unité ou n'importe quelle Sphère de centre O , l'on trouve pour angle solide l'aire de la sphère unité dans E. L'angle solide est invariant par homothétie de centre O .

Cherchons l'application dérivée en un point M_0 de E de l'application ξ . Pour la trouver, il nous est possible de choisir un référentiel orthonormal quelconque de E; nous le choisirons de façon que le point M_0 ait ses $N-1$ premières coordonnées nulles, la N-ième coordonnée étant simplement la distance r_0 de M_0 au point O . Les formules qui définissent la transformation $\xi : x \to y$ sont alors

$$(\text{IV},10;19) \qquad y_i = \frac{x_i}{r} = \frac{x_i}{\sqrt{x_1^2 + x_2^2 + \cdots + x_N^2}} \; ; \quad i = 1, 2, \ldots, N .$$

On en déduit, par différentiation :

$$(\text{IV},10;20) \qquad dy_i = \frac{dx_i}{r} - \frac{x_i}{r^2}\, dr \; ; \text{ avec } dr = \sum_{j=1}^{N} \frac{x_j}{r}\, dx_j \; ;$$

ce qui donne, dans la position particulière de M_0 par rapport au référentiel choisi, $dr = dx_N$, d'où

(IV, 10; 21) $$dy_i = \frac{dx_i}{r_0}, \text{ pour } i = 1, 2, \ldots, N - i \; ; \; dy_N = 0.$$

<u>Il en résulte que l'application dérivée $\xi'(M_0)$ cherchée est celle qui consiste, pour chaque vecteur $\vec{X}$, à le projeter ortnogonalement sur l'nyperplan perpenaiculaire en M_0 a" rayon vecteur $\vec{OM_0}$, et ensuite à le transformer par homothétie de rapport</u> $\frac{1}{r_0}$.

soit alors $\Phi : \mathcal{O} \to \Phi(\mathcal{O})$ "ne carte de V ; soit $u_0 \in \mathcal{O}$, et $M_0 = H(\Phi(u_0))$. Alors le facteur $D(u_0)$ relatif aux $(N-1)$- aires sur V est l'aire $(N-1)$-dimensionnelle dans E du système des vecteurs $\vec{x_i} = H'(\Phi(u_0)) \cdot \frac{\partial \Phi}{\partial u_i}(u_0)$. Au contraire, pour l'angle solide, nous avons à considérer un facteur $A(u_0)$, qui doit être l'aire $(N-1)$-dimensionnelle du système des vecteurs $\vec{Y_i}$ transformés des précédents par $\xi'(M_0)$. Mais d'après le théorème 105 de projection des aires, et la transformation des aires par homothétie (corollaire 1 du théorème 107), on a $A(u_,) = D(u_0) \frac{\cos \theta_0}{r_0^{N-1}}$,

où $\theta_0 = \theta(u_0)$. est l'angle aigu de OM_0 avec la normale en M_0 à la variété V.

Finalement, on a la formule, sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^n$:

(IV,10 ; 22) $$d\omega_\Phi = \Delta(u)\, d\ u = \frac{\cos\theta(u)}{(r(u))^{N-1}} D(u)\, du.$$

L'image par Φ de cette mesure donne la mesure des angles dans l'ouvert $\Phi(\mathcal{O})$ de V ; elle s'écrit :

(IV,10 ; 23) $$\Phi(d\omega_\Phi) = \frac{\cos\theta}{r^{N-1}} \Phi(dS_\Phi),$$

où, en un pointu de V, θ et h sont relatifs a" point $H(v)$ de E. On trouve bien la même mesure, pour tous les ouverts $\Phi(\mathcal{O})$ de V, et c'est

$$(\mathrm{IV},10;24)\quad . \quad d\omega = \frac{\cos\theta}{r^{N-1}}\, dS = \frac{\cos\theta(v)}{(r(v))^{N-1}}\, dS(v)\,.$$

L'angle solide d'une partie A dS-mesurable de V , ou de V tout entière est donc

$$(\mathrm{IV},10;25)\qquad \Omega = \int\cdots\int_A \ \text{ou}\ \int\cdots\int_V \frac{\cos\theta}{r^{N-1}}\, dS\ ;$$

comme on pouvait géométriquement s'y attendre !

Calcul d'intégrales de volumes à partir d'intégrales d'hypersurface

Soit Φ un C^1-difféomorphisme d'un ouvert $\mathcal{O}$ de $\mathbb{R}^N$ sur un ouvert Ω d'un espace affine euclidien E .

Si alors $\vec{f}$ est une fonction définie sur Ω , à valeurs dans un espace de Banach F , l'intégrale $\iint\cdots\int_\Omega \vec{f}(x)\, dx$, où dx est la mesure attachée à l'espace euclidien affine, peut se calculer, d'après le corollaire 2 du théorème 102 du changement de variables dans les intégrales multiples par la formule (IV,9;76):

$$(\mathrm{IV},10;26)\qquad \iint\cdots\int_\Omega \vec{f}(x)\, dx = \iint\cdots\int_{\mathcal{O}} \vec{f}(\Phi(u))\,\left|\det \Phi'(u)\right| du\,,$$

dans laquelle le déterminant écrit est le déterminant jacobien de la transformation, <u>lorsqu'on prend dans E un référentiel orthonormal quelconque;</u> c'est encore tout simplement le volume du parallélépipède défini par les vecteurs $\vec{X}_i = \frac{\partial \Phi}{\partial u_i}(u)$, $i = 1,2,\ldots,N$.

Posons $u = (v,w)$ où $v = (u_1,\ldots,u_{N-1})$ et $w = u_N$.

Fixons w . Alors la restriction de l'application Φ à l'ensemble des u correspondant a cette valeur de w est une hypersurface Σ_w de Ω de classe C^1. Si nous considérons les vecteurs $\vec{X}_i$ les N-1 premiers sont des vecteurs tangents à l'hypersurface au point considéré, et le N-ième $\vec{X}_N$ est au contraire en dehors de l'hyperplan tangent.

Désignons par $\vec{\nu}$ un vecteur unitaire de la normale à l'hypersurface Σ_w au point considéré $\Phi(u)$ Le volume du parallélépipède des vecteurs $\vec{X}_i$ est alors le produit de l'aire (N-1)-dimensionnelle du parallélépipède défini par les N-1 premiers, par la longueur de la projection orthogonale du dernier sur la normale,. soit $|(\vec{X}_N \mid \vec{\nu})|$. Mais l'aire (N-1)-dimensionnelle des N-1 premiers n'est autre que le facteur $D(v,w)$ * permettant le calcul des aires sur la surface Σ_w. Nous sommes ainsi en droit d'écrire

$$(\mathrm{IV},10;27)\qquad \iint\cdots\int_{\Omega} \vec{f}(x)\,dx = \iint\cdots\int_{\mathcal{O}} \vec{f}(\Phi(u))\,D(u)\left|\left(\overrightarrow{\frac{\partial\Phi}{\partial w}}(u) \mid \vec{\nu}\right)\right| du \,.$$

SI alors nous appliquons le théorème de Fubini. elle peut s'écrire, si l'on sait que $\vec{f}$ est intégrable , ou si l'on sait qu'elle est réelle ≥ 0 et mesurable,

$$(\mathrm{IV},10;28)\qquad \int dw \int\cdots\int \vec{f}(\Phi(u))\,D(v,w)\left|\left(\overrightarrow{\frac{\partial\Phi}{\partial w}}(u) \mid \vec{\nu}\right)\right| dv \,,$$

ou encore

$$(\mathrm{IV},10;29)\qquad \int dw \int\cdots\int \vec{f}(x)\left|\left(\overrightarrow{\frac{\partial\Phi}{\partial w}} \mid \vec{\nu}\right)\right| dS \,.$$

Ainsi on peut calculer cette intégrale N-uple, comme la succession d'une intégrale d'hypersurface sur l'hypersurface Σ_w dépendant du paramétre w , et d'une intégrale simple par rapport à w.

* C'est une fonction de v et W, nécessairement; pour w fixé, il ne dépend que de v . On a donc $D(v,w)\,dv = dS_w$ mesure des aires relative à la variété paramétrique

$\Sigma_w : v \longrightarrow \Phi(v,w)$. Son Image par l'homéomorphisme Φ

est la mesure des aires sur Σ_w elle-même; bien que variant avec w , puisque c'est une mesure sur Σ_w , on admet de la représenter toujours par dS (!! , les mathématiques courantes sont trop compliquées sans abus de langage !)

Plaçons-nous dans le cas particulièrement intéressant suivant. Soit h une fonction réelle de classe C^1 définie sur un ouvert Ω de l'espace affine E et dont l'application dérivée $h'(a)$ est $\neq \vec{0}$ en tout point a de Ω. Plaçons-nous précisément en un point a. Puisque l'application dérivée $h'(a)$ n'est pas nulle, il existe au moins une des dérivées partielles de h, par rapport à un système de coordonnées orthonormales dans E, qui n'est pas nulle. Supposons pour fixer les idées que ce soit $\frac{\partial h}{\partial x_N}(a)$. Alors il existe d'après le théorème des fonctions implicites (thdorème 28 du chapitre III) un voisinage ouvert Ω_a de a dans Ω, dans lequel l'équation $w = h(x_1, x_2, \dots, x_N)$ peut être résolue par rapport à x_N, sous la forme $x_N = g(x_1, x_2, \dots, x_{N-1}, w)$, g fonction de classe C^1. En outre, le même théorème nous donne, d'après la règle (III,8;23): $\partial g/\partial w = \frac{1}{\partial h/\partial x_N}$. Si alors nous nous restreignons à l'ouvert Ω_a, et si nous considérons les formules

(IV,10;30)
$$\begin{aligned} x_i &= u_i = v_i, \quad i = 1, 2, \dots, N-1; \\ x_N &= g(u_1, u_2, \dots, u_{N-1}, u_N) = g(v_1, v_2, \dots, v_{N-1}, w), \end{aligned}$$

nous nous trouvons exactement dans la situation décrite cl-dessus. Nous définissons un C^1-difféomorphisme Φ d'un ouvert $\mathcal{O}_a$ de $\mathbb{R}^N$ sur Ω_a, de telle manière que les sous-espaces w = constante deviennent les hypersurfaces Σ_w d'équation $h(x) = w$. Comme les x_i, $i \leq N-1$, ne dkpendent pas de w, le vecteur $\frac{\partial \Phi}{\partial w}(u)$ est le vecteur parallèle à l'axe des x_N, de composante $\frac{\partial g}{\partial w}(u) \left(= \frac{1}{\frac{\partial h}{\partial x_N}(x)} \right)$ sur cet axe. Alors, compte tenu de ce qu'un vecteur unitaire de la normale est, d'après le corollaire 5 du théorème 33 quarto du chapitre III, $\vec{\nu} = \pm \frac{\overrightarrow{\text{grad}}\, h}{\|\overrightarrow{\text{grad}}\, h\|}$ la projection du vecteur $\vec{\nu}$ sur l'axe des x_N est égale à $\pm \frac{\frac{\partial h}{\partial x_N}(x)}{\|\overrightarrow{\text{grad}}\, h(x)\|}$

On en déduit la formule :

(IV,10;31)
$$\left| \left(\frac{\partial \Phi}{\partial w} \middle| \vec{\nu} \right) \right| = \frac{1}{\left| \frac{\partial h}{\partial x_N} \right|} \frac{\left| \frac{\partial h}{\partial x_N} \right|}{\|\overrightarrow{\text{grad}}\, h\|} = \frac{1}{\|\overrightarrow{\text{grad}}\, h\|}$$

Alors, si $\vec{f}$ est une fonction intégrable sur Ω_a, à valeurs dans un espace de Banach $\vec{F}$, on peut calculer son intégrale de volume par l'intermédiaire d'une intégrale de surface, sous la forme suivante

$$(IV,10;32) \qquad \iint\cdots\int_{\Omega_a} \vec{f}(x)\, dx = \int dw \int\cdots\int_{\Sigma_w} \frac{\vec{f}(x)}{\|\overrightarrow{grad}\, h(x)\|}\, dS .$$

Naturellement ce calcul n'est valable que dans l'ouvert Ω_a, voisinage du point a ; mais le résultat final ne fait pas intervenir le choix particulier des coordonnées relatif au point a. La formule (IV,10;32) est entièrement intrinsèque. En recouvrant alors un compact K de Ω par un nombre fini d'ouverts tels que Ω_a, on en déduit que ce résultat est valable pour toute fonction à support compact dans Ω, à valeurs dans F, et, par un passage à la limite, on finit par obtenir, pour n'importe quelle fonction intégrable sur Ω, à valeurs dans $\vec{F}$:

$$(IV,10;33) \qquad \iint\cdots\int_{\Omega} \vec{f}(x)\, dx = \int dw \int\cdots\int_{\Sigma_w} \frac{\vec{f}(x)}{\|\overrightarrow{grad}\, h(x)\|}\, dS$$

On a donc :

Théorème 108 - Soit Ω un ouvert d'un espace affine euclidien E, et soit h une fonction réelle de classe C^1, définie sur Ω, dont la dérivée en tout point est $\neq 0$. Si alors on appelle Σ_w l'hypersurface définie par l'équation $h(x) = w$, et si $\vec{f}$ est une fonction définie sur Ω, à valeurs dans un espace de Banach $\vec{F}$ et intégrable pour la mesure des volumes dx de E, alors, pour presque toutes les valeurs de w, la fonction $\vec{f}/\|\overrightarrow{grad}\, h\|$ est intégrable sur Σ_w par rapport à la mesure superficielle dS, l'intégrale $\int\cdots\int_{\Sigma_w} \frac{\vec{f}}{\|\overrightarrow{grad}\, h\|}\, dS$, ainsi définie pour presque toutes les valeurs de w, est une fonction de w à valeurs dans $\vec{F}$, qui est intégrable par rapport à la mesure dw et l'on a la formule (IV,10;33). Si f est une fonction dS-mesurable réelle ≥ 0 non nécessairement intégrable, alors les membres de (IV,10;33) sont tous les deux égaux, à valeur finie ou infinie.

Remarque Géométriquement le résultat précédent est assez intuitif. Un élément de volume dx peut se représenter comme le produit d'un élément de surface dS de la surface Σ_w par la distance normale $d\nu$ séparant la surface Σ_w de la surface Σ_{w+dw}. Or $\overrightarrow{\text{grad}}\, h$ est un vecteur normal à Σ_w, de longueur $\left|\frac{dw}{d\nu}\right|$; d'où l'on déduit $d\nu = \frac{dw}{\|\overrightarrow{\text{grad}}\, h\|}$, et par conséquent $dx = \frac{dS\, dw}{\|\overrightarrow{\text{grad}}\, h\|}$. Naturellement ce raisonnement manque de rigueur, et c'est ce qui a été donné avant qui en est la justification précise.

Corollaire 1 - Soit E un espace euclidien affine, et soit r la fonction "distance au point O de E", c'est-à-dire $r(x) = \|\overrightarrow{x - O}\|$. Soit $\vec{f}$ une fonction définie sur E, à valeurs dans un espace de Banach $\vec{F}$ et intégrable pour la mesure des volumes dx. Alors, pour presque toutes les valeurs de r, la fonction $\vec{f}$ est intégrable par rapport à la mesure superficielle sur la sphère Σ_r de centre O et de rayon r.

Son intégrale $\int \cdots \int_{\Sigma_r} \vec{f}(x)\, dS$, ainsi définie pour presque toutes les valeurs de r, est intégrable par rapport à la mesure dr, et l'on a la formule

$$(\text{IV},10;34) \qquad \iint\cdots\int \vec{f}(x)\, dx = \int_{]0,+\infty[} dr \int\cdots\int_{\Sigma_r} \vec{f}(x)\, dS .$$

Si f est une fonction dx-mesurable réelle ≥ 0, non nécessairement intégrable, les deux membres de (IV,10;34) sont encore égaux, à valeur finie ou infinie.

Il suffit en effet d'appliquer le théorème au complémentaire Ω de O dans E, relativement à la fonction $h = r$. Le gradient de cette fonction est, en chaque point M, le vecteur unitaire porté par $\overrightarrow{OM}$, et par suite la longueur du gradient est identique à 1, d'où le résultat.

Corollaire 2 - soit $\vec{f}$ une fonction définie sur la demi-droite $\mathbb{R}_+$, ensemble des éléments ≥ 0 de $\mathbb{R}$. Soit r la fonction scalaire définie sur l'espace affine euclidien E de dimension N par $x \longrightarrow r(x) = \|\overrightarrow{x - O}\|$ distance de x au point O de E ; alors $\vec{f} \circ r : x \longrightarrow \vec{f}(r(x))$ est une fonction sur E, à valeurs dans $\vec{F}$ ne dépendant que de la distance de la variable à O (et qu'on abrègera aussi par $\vec{f}(r)$). La fonction $\vec{f} \circ r$ est mesurable sur E pour la mesure des volumes, si et seulement si f est mesurable sur $\mathbb{R}_+$ pour la mesure de Lebesgue. La fonction $\vec{f} \circ r$ est intégrable sur E pour la mesure des volumes, si et seulement si la fonction $t \longrightarrow t^{N-1} \vec{f}(t)$ est intégrable sur $\mathbb{R}_+$ pour la mesure de Lebesgue, et on a :

$$\iint \cdots \int_E \vec{f}(r)\, dx = S_N \int_{\mathbb{R}_+} \vec{f}(t)\, t^{N-1}\, dt\,, \tag{IV,10;35}$$

où S_N est l'aire des sphères de rayon 1 dans E . Enfin l'image, par l'application $r : x \longrightarrow r(x)$ de E. dans $\mathbb{R}_+$, de la mesure des volumes dx, est la mesure $S_N\, t^{N-1}\, dt$ sur $\mathbb{R}_+$.

Généralement, on emploie r comme variable muette dans le second membre de (IV,10 35) : $S_N \int_{[0,+\infty[} \vec{f}(r)\, r^{N-1}\, dr$.

Démonstration Si f est scalaire, continue, à support compact sur $\mathbb{R}_+$, on peut lui appliquer le corollaire 1; comme f est constante, égale à $f(r)$, sur la sphère Σ_r de rayon r, la quantité $\int \cdots \int_{\Sigma_r} f(r)\, dS$ se réduit tout simplement au produit de $f(r)$ par l'aire de la sphère de rayon r ; et, par homothétie (corollaire 1 du théorème 107), cette aire est égale au produit de r^{N-1} par l'aire S_N de la sphère de rayon 1, d'où (IV,10:35).

Cette formule montre que l'image directe de la mesure dx par l'application $x \longrightarrow r(x)$ n'est autre que la mesure $S_N\, t^{N-1}\, dt$ sur $\mathbb{R}_+$. Mais alors le théorème 60 nous dit que, si $\vec{f}$ est une fonction sur $\mathbb{R}$, la fonction $\vec{f}(r)$ est dx-intégrable, si et seulement si $\vec{f}$ est $(S_N\, t^{N-1} dt)$ - intégrable sur $\mathbb{R}_+$, et le théorème 51 nous dit que $\vec{f}$ est intégrable par rapport à ($S_N\, t^{N-1}$ dt) si et seulement si $S_N\, \vec{f}(t)\, t^{N-1}$ est dt -intégrable.

Ce corollaire nous prouve qu'on peut ramener au calcul d'une intégrale simple toute intégrale de volume d'une fonction ne dépendant que de r, pourvu qu'on connaisse l'aire S_N de la sphère unité dans $\mathbb{R}^N$. Cette aire est bien connue pour N = 1, 2, 3 :

$$S_1 = 2^{*}\,, \qquad S_2 = 2\pi\,, \qquad S_3 = 4\pi$$

Inversement, s'il existe une fonction f sur $\mathbb{R}_+$, pour laquelle on sache à la fois calculer $\iint \cdots \int_E \vec{f}(r)\, dx$ et

* Voir ce renvoi à la page 700;

$\int_{\mathbb{R}_+} \overrightarrow{f(t)}\, t^{N-1}\, dt$, cela donnera la valeur de S_N (si ces intégrales ne sont pas nulles). C'est ce que nous ferons plus tard, en utilisant $f(t) = e^{-t^2}$. On trouve (on pourra le faire à titre d'exercice) :

$$S_4 = 2\pi^2 , \quad S_5 = \frac{8}{3}\pi^2 , \quad S_6 = \pi^3 , \ldots .$$

L'aire d'une sphère de rayon R est alors $S_N R^{N-1}$, ce qui donne :

$$2 , \quad 2\pi R, \quad 4\pi R^2 , \quad 2\pi^2 R^3 , \quad \frac{8}{3}\pi^2 R^4 , \quad \pi^3 R^5 , \ldots .$$

<u>Corollaire 3</u> - <u>Dans un espace affine euclidien de dimension N , le volume d'une boule de rayon R est</u> $S_N \frac{R^N}{N}$ <u>(primitive de l'aire de la sphère</u> $S_N R^{N-1}$, <u>s'annulant pour</u> $R = 0$).

En effet ce volume se calcule d'après (IV,10;35), car c'est l'intégrale de la fonction caractéristique de la boule, fonction qui ne dépend que de r ; ici $f(r) = 0$ pour $r > R$, et $= 1$ pour $r \leqslant R$. D'où la valeur $S_N \int_{[0,R]} r^{N-1}\, dr = S_N \frac{R^N}{N}$.

On trouve ainsi, pour $N = 1, 2, 3, 4, 5, 6$:

$$2R , \pi R^2 , \quad \frac{4}{3}\pi R^3 , \quad \frac{1}{2}\pi^2 R^4 , \quad \frac{8}{15}\pi^2 R^5 , \quad \frac{1}{6}\pi^3 R^6 .$$

* Renvoi de la page 699 -

S_1 est une aire 0-dimensionnelle; la sphère de $\mathbb{R}$, de centre 0 et de rayon 1, se réduit à l'ensemble $\{-1, +1\}$ à 2 éléments, son aire est 2 (remarque 5°/ après le théorème 107). D'ailleurs, si une fonction sur la droite $\mathbb{R}$ ne dépend que de la distance à l'origine, cela veut simplement dire qu'elle est paire; alors (IV,10;35) s'écrit simplement $\int_{]-\infty,+\infty[} \overrightarrow{f(x)}\, dx = S_1 \int_{]0,+\infty[} \overrightarrow{f(x)}\, dx$, ce qui exige bien $S_1 = 2$.

corollaire 4 - Soit r la distance au point 0 dans un espace euclidien affine de dimension N.

Alors l'intégrale $\iint\cdots\int_{r\leq 1}\frac{dx}{r^\alpha}$ (resp. $\iint\cdots\int_{r\geq 1}\frac{dx}{r^\alpha}$) est finie si et seulement si $\alpha < N$ (resp. $\alpha > N$)

Ce corollaire généralise le corollaire 1 du théorème 97, qui correspondait à $N = 1$.

Démonstration

D après le corollaire 2, les intégrales précédentes sont finies, si et seulement si les intégrales simples correspondantes $\int\frac{r^{N-1}}{r^\alpha}\,dr = \int\frac{1}{r^{\alpha-N+1}}\,dr$ sont finies; cela se produit si $\alpha - N + 1 < 1$ (resp. > 1) ou $\alpha < N$ (resp. $> N$).

§ 11 FONCTIONS REPRÉSENTÉES PAR DES SÉRIES OU PAR DES INTÉGRALES

Fonctions représentées par des séries

soit $f_0, f_1, \ldots, f_n, \ldots$ une suite d'applications d'un ensemble E dans un ensemble F, et supposons que cette suite converge simplement, pour n tendant vers l'infini, vers une application f de E dans F ; alors la fonction f est ce qu'on appelle une fonction définie comme limite d'une suite de fonctions.

Nous nous proposons d'étudier la continuité de f, si E et F sont des espaces topologiques; ou l'intégrabilité de f, si $E = X$ est un espace localement compact muni d'une mesure de Radon $\mu \geq 0$, et F un espace de Banach; ou la dérivabilité de f, si E et F dont des espaces affines normés; tout cela à partir de propriétés analogues des f_n. En général, en analyse, au lieu d'avoir une suite, on a affaire à une série. Dans ce cas, E est un ensemble, $\vec{F}$ un espace de Banach, et on a une fonction $\vec{S}$ définie par une formule

$$\overrightarrow{S(x)} = \sum_{n=0}^{\infty} \overrightarrow{u_n(x)}, \quad x \in E. \qquad (IV, 11; 1)$$

On suppose que, par des critères quelconques, on ait pu démontrer la convergence simple de cette série; et on se propose d'étudier la continuité, l'intégrabilité et la dérivabilité de $\vec{S}$, à partir des propriétés analogues des fonctions $\vec{u}_n$. Il est bien évident que, dans ce cas, le point de vue des séries n'est pas différent du point de vue des suites, car $\vec{S}$ est la limite de la suite des $\vec{S}_n = \sum_{\ell=0}^{n} \vec{u}_\ell$, et qu'inversement, si $\vec{f}$ est limite d'une suite de fonctions $\vec{f}_n$, elle est la somme de la série

$$\vec{f}_0 + (\vec{f}_1 - \vec{f}_0) + (\vec{f}_2 - \vec{f}_1) + \cdots + (\vec{f}_n - \vec{f}_{n-1}) + \cdots .$$

Mais, s'il est vrai qu'au point de vue théorique les deux problèmes sont les mêmes dans le cas où $\vec{F}$ est un espace de Banach, et que le problème des suites est plus général puisqu'il s'applique dans le cas ou F est un espace topologique quelconque, dans la pratique c'est surtout par des séries que sont représentées les fonctions de l'analyse.

Si F est le corps des complexes, il pourra aussi arriver qu'on ait affaire à une fonction représentée comme un produit infini, à savoir $\Pi(x) = \prod_{n=0} u_n(x)$. On suppose alors qu'on ait pu démontrer la convergence simple de ce produit, et l'on cherche encore à déterminer si la fonction Π est continue, intégrable ou dérivable, à partir de propriétés analogues des termes du produit.

La plupart des critères ont déjà été vus, dans diverses parties des chapitres précédents, dans ce cas là nous nous contenterons de les rappeler, et ne donnerons ici que les critères nouveaux.

Continuité de la somme d'une série

On dispose du théorème 65 du chapitre II que nous nous bornons a répéter :

- Soit f_n une suite d'applications d'un espace topologique E dans un espace métrique F , convergeant vers une limite f , localement uniformément sur E ; si alors les f_n sont continues, f **est aussi continue**.

Naturellement ce théorème donné pour la limite d'une suite est aussi valable pour la somme d'une série ou la limite d'un produit Infini.

Intégrabilité de la somme d'une série par rapport à une mesure $\mu \geq 0$

Les théorèmes dont on dispose ici sont les théorèmes 34, 35, 36 et corollaires, 37, du présent chapitre.

Nous nous permettrons seulement ici de donner l'amélioration suivante :

Théorème 110 - Soit $\vec{f}_n$ une suite de fonctions μ - intégrables sur X , à valeurs dans un espace de Banach $\vec{F}$. On suppose que les $\vec{f}_n$ convergent simplement μ - presque partout vers une limite $\vec{f}$, et sont majorées en normes par une même fonction $g \geq 0$, μ - intégrable, fixe. Alors, quelle que soit la partie μ - mesurable Y de X , les $\int_Y \vec{f}_n$ convergent vers $\int_Y \vec{f}$, et uniformément par rapport à Y , lorsque Y parcourt l'ensemble des parties μ - mesurables de X .

<u>**Démonstration**</u> - Nous avons en effet vu, au théorème 35 de Lebesgue, que les quantités $\int \|\vec{f} - \vec{f}_n\| d\mu$ convergent vers 0 .

Alors, étant donné $\varepsilon > 0$, il existe un entier p tel que $n \geq p$ entraîne $\int \|\vec{f} - \vec{f}_n\| d\mu \leq \varepsilon$.

On a alors, pour toute partie mesurable Y de X , la même inégalité $\int_Y \|\vec{f} - \vec{f}_n\| d\mu \leq \varepsilon$, d'où a fortiori l'inégalité $\left\| \int_Y \vec{f}_n d\mu - \int_Y \vec{f} d\mu \right\| \leq \varepsilon$, ce qui démontre le théorème.

<u>Corollaire</u> - <u>soit $\vec{f}_n$ une suite de fonctions, définies sur un intervalle fermé $[a,b]$ de $\mathbb{R}$, à valeurs dans un espace de Banach $\vec{F}$, dx-intégrables, et convergeant uniformément pour n infini vers une fonction limite $\vec{f}$; alors les quantités $\int_\alpha^\beta \vec{f}_n(x)\,dx$ convergent vers $\int_\alpha^\beta \vec{f}(x)\,dx$, uniformément par rapport à α et β dans $[a,b]$.</u>

Dérivabilité de la somme d'une série

Ici au contraire, nous n'avons jusqu à présent encore donné aucun théorème.

Soit f_n une suite d'applications d'un espace affine normé E dans un espace affine normé F.

On suppose que ces applications soient toutes dérivables. Peut-on déduire d une convergence assez forte des f_n vers f , par exemple d'une convergence uniforme, la dérivabilité de la fonction limite f , et la convergence des f'_n vers f' ?

Il n'en est rien, parce que, d'une majoration portant sur des fonctions, on ne peut déduire aucune majoration sur leurs dérivées. Considérons en effet la suite des fonctions complexes d'une variable réelle définies par

$$(IV,11;2) \qquad f_n(x) = \frac{e^{in^2x}}{n}, \qquad n \geq 1 .$$

Cette suite converge uniformément vers 0 sur toute la droite réelle $\mathbb{R}$ pour n tendant vers l'infini; cependant les dérivées sont données par la formule

$$(IV,11;3) \qquad f'_n(x) = i\,n\,e^{in^2x}$$

et le module de la dérivée& converge vers l'infini lorsque n tend vers l'infini, pour toutes les valeurs de x. De la même manière, considérons la série de Weierstrass :

$$S(x) = \sum_{n=0}^{\infty} \frac{\sin 3^n x}{2^n} .$$

Cette série de fonctions réelles d'une variable réelle est uniformément convergente sur $\mathbb{R}$, parce que son terme général est majoré en module par $\frac{1}{2^n}$, elle représente donc une fonction continue.

Si nous considérons la série des dérivées

(IV,11;5) $$\sum_{n=0}^{\infty} \left(\frac{3}{2}\right)^n \cos 3^n x ,$$

elle n'est pas convergente : d'ailleurs précisément Weierstrass a montré que la fonctions ne possède de dérivée pour aucune valeur de x, et c'est le premier exemple qui ait ete donne d'une pareille fonction.

On devra donc faire l'hypothèse non seulement de la convergence des fonctions, mais de la convergence des dérivées; et même nous allons voir qu il suffira de suppo-ser la convergence uniforme des dérivées et la convergence des fonctions seulement en un point, pour en déduire la convergence uniforme des fonctions et la légitimité du passage à la limite pour la dérivation :

Théorème 111 - soient f_n des applications d'un ouvert Ω d'un affine normé E dans un affine normé F, dérivables (resp. de classe C^1). Si les dérivées f'_n (fonctions sur Ω à valeurs dans $\mathcal{L}(\vec{E}\,;\vec{F})$) convergent localement uniformément vers une limite g, si les $f_n(a)$ convergent vers une limite $f(a)$ pour au moins un point a de Ω, si Ω est connexe et F complet, alors les f_n convergent localement uniformément vers une limite f dans Ω ; f est dérivable (resp. de classe C^1) et sa dérivée est g.

Même si Ω n'est pas connexe et F non complet, mais si on suppose que les f'_n convergent localement uniformément vers une limite g et les f_n simplement vers une limite , alors la convergence des f_n vers f est localement uniforme, f est dérivable (resp. de classe C^1) et $f' = g$.

Démonstration - La convergence des f'_n étant uniforme locale, il existe un voisinage de a dans lequel les f'_n convergent uniformément vers g Appelons $B = B(a;\rho)$ une boule de centre a et de rayon $\rho > 0$, contenue dans ce voisinage. Alors, six est uans cette boule, nous pouvons appliquer la formule des accroissements finis (théorème 13 du chapitre III) à la fonction $\overrightarrow{f_{m,n}} = f_m - f_n$:

(IV,11;6) $$\left\| \overrightarrow{f_{m,n}(x) - f_{m,n}(a)} \right\| \leq \rho \sup_{\|\xi - a\| \leq \rho} \left\| f'_{m,n}(\xi) \right\|$$

Mais $f'_{m,n} = f'_m - f'_n$ converge uniformément vers 0 dans la boule B, pour m et n tendant vers l'infini; donc le premier membre converge uniformément vers 0 dans cette boule. Comme, par hypothèse, $f_{m,n}(a) = \overrightarrow{f_m(a) - f_n(a)}$ converge vers 0, on en déduit que $f_{m,n} = f_m - f_n$ converge uniformément vers 0 dans cette boule pour m et n infinis.

Les f_n ne sont pas nécessairement toutes bornées dans B. Mais, si nous déterminons un entier n_0 tel que, pour $n \geq n_0$, on ait $\sup\limits_{\|\xi - a\| \leq \rho} \left\| \overrightarrow{f_{n_0}(\xi) - f_n(\xi)} \right\| \leq 1$, alors les fonctions

$\overrightarrow{f_n - f_{n_0}}$ sont toutes bornées en norme par 1 dans **B**,

et $(\overrightarrow{f_m - f_{n_0}}) - (\overrightarrow{f_n - f_{n_0}})$, $m \geq n_0$, $n \geq n_0$,

converge encore uniformément vers 0 dans B pour m et n infinis. Cela prouve que les fonctions $\overrightarrow{f_n - f_{n_0}}$, $n \geq n_0$, forment une suite de Cauchy dans l'espace métrique $(\vec{F})_{cb}$; comme cet espace métrique estcomplet si $\vec{F}$ est complet (Corollaire 2 du théorème 65 du chapitre II), les $\overrightarrow{f_n - f_{n_0}}$ convergent uniformément vers une limite dans B, pour n infini, donc aussi les f_n. Si $\vec{F}$ n'est pas complet, mais si on sait à l'avance que les f_n convergent simplement vers une limite f, on en déduit encore que la convergence est uniforme dans B. En effet, quel que soit $\varepsilon > 0$, il existe un entier? tel que $m \geq p$, $n \geq p$, entraîne, pour tout x de B : $\| \overrightarrow{f_m(x) - f_n(x)} \| \leq \varepsilon$; en passant à la limite pour m infini (pour x fixé), on en déduit bien que pour $n \geq p$, on a, pour tout x de B , $\| \overrightarrow{f(x) - f_n(x)} \| \leq \varepsilon$, ce qui prouve notre affirmation.

Si on sait d'avance que les f_n convergent simplement vers une limite f_1 , le raisonnement que nous venons de faire pour a est valable pour n'importe quel point de Ω , et cela prouve que la convergence des f_n vers f est localement uniforme. Si au contraire on sait seulement que les $f_n(a)$ convergent vers $f(a)$, mais que F est complet et Ω connexe, il faut un nouveau raisonnement.

Appelons $\mathcal{E}$ l'ensemble des points x de Ω tels que la suite des $f_n(x)$ converge pour n infini vers une limite. Tout d'abord cet ensemble est ouvert. Si en effet b appartient à $\mathcal{E}$, nous pouvons pour b faire le même raisonnement que nous venons de faire pour a , et nous voyons qu'il existe toute une boule de centre b contenue dans Ω en chaque point de laquelle les f_n convergent vers une limite; donc $\mathcal{E}$ est ouvert. Montrons maintenant que $\mathcal{E}$ est est aussi fermé dans Ω . Pour cela supposons que $b_0, b_1, \ldots, b_j, \ldots$ soit une suite de points de Ω , convergeant, pour j tendant vers l'infini, vers un point limite b de Ω ; et supposons que tous les b_j appartiennent à $\mathcal{E}$; montrons que b lui aussi appartient à $\mathcal{E}$

D'après l'hypothèse, il existe un voisinage de b dans lequel les f'_n convergent uniformément vers g ; par conséquent il existe une boule $B(b; \rho')$ de centre b , entièrement contenue dans ce voisinage; comme les b_j convergent vers b pour j tendant vers l'infini, il existe un entier j_0 tel que $j \geq j_0$ entraîne $\| b_j - b \| \leq \frac{\rho'}{3}$; mais alors la boule de centre b_j et de rayon $\frac{2\rho'}{3}$ est contenue dans B $(b; \rho')$, et par conséquent, dans cette boule, les f'_n convergent uniformément vers g , en même temps que les $f_n(b_j)$ convergent vers $f(b_j)$; nous pouvons alors appliquer à ce point b_j le raisonnement que nous avons appliqué tout à l'heure au point a , et nous voyons qu'en tout point x de la boule $B(b_j; \frac{2\rho'}{3})$ les $f_n(x)$ convergent vers une limite; et le point b fait précisément partie de cette boule, donc il appartient lui aussi à l'ensemble $\mathcal{E}$, qui est bien fermé. Mais alors, Ω étant supposé connexe, $\mathcal{E}$ n'étant pas vide puisqu'il contient a , ne peut être à la fois ouvert et fermé que si c'est Ω lui-même; et nous avons donc démontré que les f_n convergent vers une limite f partout sur Ω ; et dans ce cas nous avons vu plus haut que la convergence est nécessairement localement uniforme.

Nous devons maintenant montrer que f est dérivable et de dérivée g . Comme a ne joue plus maintenant un rôle particulier, il suffit de montrer que f a au point a une dérivée égale à $g(a)$. Appelons $\vec{k}_n$ la fonction, définie sur $B = B(a;\rho)$, à valeurs dans $\vec{F}$, comme suit :

$$(\text{IV},11;7)\qquad \overrightarrow{k_n(x)} = \begin{cases} \dfrac{\overrightarrow{f_n(x) - f_n(a) - f_n'(a).\overrightarrow{x-a}}}{\|\overrightarrow{x-a}\|} & \text{, pour } x \neq a\,; \\ \vec{0} \text{ pour } x = a\,. \end{cases}$$

Ces fonctions convergent simplement dans B vers la fonction $\vec{k}$ définie par :

$$(\text{IV},11;8)\qquad \overrightarrow{k(x)} = \begin{cases} \dfrac{\overrightarrow{f(x) - f(a) - g(a).\overrightarrow{x-a}}}{\|\overrightarrow{x-a}\|} & \text{, pour } x \neq a\,; \\ \vec{0} \text{ pour } x = a \end{cases}$$

Vals la formule des accroissements finis, sous la forme du corollaire 1 du théorème 13, donne la majoration :

$$(\text{IV},11;9)\qquad \|\overrightarrow{k_m(x)} - \overrightarrow{k_n(x)}\| \leq \sup_{\|\overrightarrow{\xi-a}\|\leq\rho} \|(f_m'(\xi)-f_n'(\xi))-(f_m'(a)-f_n'(a))\|$$

donc $\vec{k}_m - \vec{k}_n$ converge uniformément vers $\vec{0}$ dans B pour m et n infinis. Mals alors le raisonnement déjà fait antérieurement pour f_n montre que $\vec{k}_n$ converge uniformément vers $\vec{k}$ dans B pour n infini. Or chaque fonction $\vec{k}_n$ est cnntlnue au point $x = a$, d'après la définition même de la dérivée $f_n'(a)$; donc, d'après le théorème 65 du chapitre II, la limite uniforme $\vec{k}$ est aussi continue au point $x = a$, ce qui prouve que f est dérivable en a et que $f'(a) = g(a)$.

si enfin les f_n sont de classe C^1, alors les f'_n sont continues. et comme elles convergent localement uniformément vers g, g est continue d'après le théorème 109, donc f est de classe C', et le théorème est complètement démontré.

Remarques 1°/ Nous aurions pu donner ce théorème au chapitre III, il n'y intervient que du calcul différentiel. Mais nous avons voulu bloquer dans le présent paragraphe du chapitre IV un certain nombre de résultats de même nature.

2°/ Bien entendu la 1ère partie de l'énoncé ne serait pas exacte si l'on ne supposait pas F complet et Ω connexe.

Donnons un contre-exemple, si Ω n'est pas connexe. Supposons que Ω soit le complémentaire de l'origine sur la droite réelle $\mathbb{R}$, et que f_n soit la fonction nulle pour $x < 0$ et égale à la constante n pour $x > 0$.

Alors les dérivées f'_n sont toutes identiquement nulles, et par suite convergent uniformément vers 0 pour n tendant vers l'infini. Par ailleurs les f_n convergent vers une limite en n'importe quel point $x \neq 0$; mais elles ne convergent pas vers une limite sur tout l'axe réel $\mathbb{R}$.

3°/ C'est la convergence uniforme locale qui est essentielle dans toutes ces questions. Même si on suppose la convergence des f'_n vers g uniforme dans Ω, on peut seulement dire que les f_n convergent vers f localement uniformément.

Considérons par exemple, sur la droite réelle $\mathbb{R}$, les fonctions $f_n(x) = \frac{x}{n}$. Elles convergent simplement vers 0 pour n infini; leurs dérivées $f'_n(x) = \frac{1}{n}$ convergent vers 0 uniformément sur $\mathbb{R}$. On en déduit, ce qui est bien visible a priori, que les f_n convergent vers 0 uniformément sur tout compact (ou localement uniformément); mais elles ne convergent pas vers 0 uniformément sur $\mathbb{R}$, car $\| f_n \| = +\infty$.

Corollaire 1 - Soient f_n des applications d'un ouvert Ω d'un affine normé E dans un affine normé F, de classe C^m. Si, pour tout $k \leq m$, les $f_n^{(k)}$ fonctions sur Ω à valeurs dans $\mathcal{L}_k(E^k; F)$ convergent localement uniformément vers une limite f_k, avec $f_0 = f$, alors f est de classe C^m, et $f_k = f^{(k)}$ pour $k \leq m$.

Evident.

<u>Corollaire 2 - Si Ω est un ouvert d'un affine normée</u>, <u>si F est un affine de Banach, alors l'espace</u> $(F^{\Omega})_{cb;m}$ <u>est complet</u>.

Rappelons que $f \in (F^{\Omega})_{cb;m}$ si f est une application de Ω dans F, continue bornée, et si, pour tout $k \leq m$, sa dérivée d'ordre k est continue bornée de Ω dans $\mathcal{L}_k(\vec{E}^k;\vec{F})$. Alors $\|f\|_m = \sum_{0 \leq k \leq m} \|f^{(k)}\|_0$ (voir page 251).

<u>Démonstration</u> - Soit f_n une suite de Cauchy de $(F^{\Omega})_{cb;m}$. Les fonctions dérivées $f_n^{(k)}$, $0 \leq k \leq m$, forment une suite de Cauchy dans $(\mathcal{L}_k(\vec{E}^k;\vec{F}))^{\Omega}_{cb;0}$, d'après la définition même des normes. Mais, F étant complet, $\mathcal{L}_k(\vec{E}^k;\vec{F})$ est complet (théorème 53 du chapitre II), donc aussi $(\mathcal{L}_k(\vec{E}^k;\vec{F}))^{\Omega}_{cb;0}$ (corollaire 2 du théorème 65 du chapitre II). Donc les $f_n^{(k)}$ convergent, uniformément sur Ω, vers une limite f_k, continue bornée sur Ω.

Mais alors, d'après le corollaire 1, la fonction f est de classe C^m, et on a $f^{(k)} = f_k$, donc f est dans $(F^{\Omega})_{cb;m}$. Par ailleurs, puisque chaque suite $f_n^{(k)}$ converge uniformément vers $f_k = f^{(k)}$, f_n converge vers f dans $(F^{\Omega})_{cb;m}$, qui est donc bien complet.

<u>Exemples</u> Naturellement nous avons énoncé le théorème précédent pour une suite, mais il s'applique, surtout en pratique, pour les séries.

Supposons donnée une série (IV,11; 1); les $\vec{u}_n$ sont des fonctions définies sur l'ouvert Ω de l'espace affine normé E, à valeurs dans l'espace de Banach $\vec{F}$. En général on sait d'avance que cette série converge simplement, et même en général localement uniformément, dans Ω. D'autre part les $\vec{u}_n$ sont en général continues, et cela démontre que $\vec{S}$ est continue. Si alors les $\vec{u}_n$ sont dérivables, on se demande si la fonction $\vec{S}$ est elle aussi dérivable, La méthode est la suivante : <u>On dérive terme à terme la série précédente, et l'on pose a priori</u> :

$$S'(x) = \sum_{n=0}^{\infty} u'_n(x) . \qquad \text{(IV,11;13)}$$

Si la série des dérivées $\sum_{n=0} u'_n$ est localement uniformément convergente dans Ω, alors $\vec{S}$ est bien dérivable, et la formule précédente est bien exacte.

En effet, dans ce cas, les $\vec{S}_m = \sum_{n=0}^{m} \vec{u}_n$ convergent vers s , en même temps que les $S'_m = \sum_{n=0}^{m} u'_n$ convergent localement uniformément vers $\sum_{n=0}^{\infty} u'_n$.

Le théorème nous dit alors précisément que $\vec{S}$ est dérivable et que sa dérivée est la limite précédente.

Considérons par exemple la série de Fourier, où les $\vec{a}_n$ et $\vec{S}(x)$ sont des vecteurs d'un espace de Banach $\vec{F}$:

(IV,11;14) $$\overrightarrow{S(x)} = \sum_{n=-\infty}^{+\infty} \vec{a}_n e^{inx}, \quad , x \text{ réel}$$

et supposons que les $\vec{a}_n$ vérifient la majoration

(IV,11;15) $$\| \vec{a}_n \| \leq \frac{A}{|n|^{p+2}}, p \text{ entier} \geq 0 \ , \ A \text{ constante} > 0 .$$

Alors la fonction? représentée par cette série, est au moins de classe C^p. En effet, si nous la dérivons k fois terme à terme, $k \leq p$, nous obtenons les formules suivantes

(IV,11;16) $$\vec{S}^{(k)}(x) = \sum_{n=-\infty}^{+\infty} (in)^k a_n e^{inx}$$

Chacune de ces séries est alors uniformément convergente et même normalement convergente, car $\| (in)^k \vec{a} \| \leq \frac{A}{|n|^2}$, et par conséquent le théorème démontre le résultat.

Par contre, si nous n'avons pas de meilleure majoration que (IV,11;15), nous ne pourrons pas en général dériver une fois de plus.

Remarques 1°/ Le fait que le théorème ne soit pas applicable n'empêche pas éventuellement la limite f des f_n ou la somme $\vec{S}$ de la série $\sum_{n=0}^{\infty} \vec{u}_n$ d'avoir une dérivée. Tout ce que l'on peut dire, si la série des dérivées n'est pas convergente, c'est que la dérivée $\vec{S}'$ de $\vec{S}$, si elle existe, n'est pas représentée par la série $\sum_{n=0}^{\infty} \vec{u}'_n$

Considérons par exemple la série trigonométrique

(IV,11;17) $$\sum_{n=1}^{\infty} \frac{\sin 2\pi n x}{\pi n} .$$

Elle converge partout, et uniformément dans tout intervalle $[\delta, 1-\delta]$ $s > 0$, d'après le théorème d'Abel (théorème 63 du chapitre II).

La somme $|\sin(2\pi x) + \sin(2\pi 2x) + \cdots + \sin 2\pi n x|$ est en effet majorée par $\frac{2}{|e^{2i\pi x}-1|}$, donc par $\frac{2}{|e^{2i\pi\delta}-1|}$ pour $\delta \leqslant x \leqslant 1-\delta$; et la suite $\frac{1}{\pi n}$ est décroissante et tend vers 0 pour n infini. Le reste admet la majoration (d'après (II,14;31)) : $|R_m| \leqslant \frac{2}{\pi(n+1)|e^{2i\pi\delta}-1|}$, ce qui montre notre assertion. Par contre la convergence n'est pas, a priori, uniforme sur $[0,1]$. Il y a donc convergence uniforme locale sur $]0\ 1[$ et la série est une fonction continue sur cet intervalle ouvert. Une dérivation terme à terme donnerait :

$$2\sum_{n=1}^{\infty} \cos 2\pi n x, \tag{IV,11;17}$$

série partout divergente

Il ne faudrait pas en déduire que la fonction S n'ait pas de dérivée.

Nous verrons plus tard que.!? est égale à

$$S(x) = \begin{cases} 0 \text{ pour } x = 0,\ x = 1; \\ \frac{1}{2} - x \text{ pour } 0 < x < 1^{(*)} \end{cases} \tag{IV,11;18}$$

C'est bien là une fonction discontinue pour $x = 0$ et $x = 1$, ce qui prouve que, malgré la convergence sur $[0,1]$, la convergence uniforme n'a lieu que dans $[\delta, 1-\delta], \delta > 0$

Mais cette fonction admet une dérivée, et est même de classe C^∞ dans tout 1 intervalle] $0,1[$. Tout ce que l'on peut dire c'est que sa dérivée, qui est la constante - 1, n'est pas représentée par la série (IV,11;17), qui est divergente. On volt par cet exemple que la méthode par laquelle Weierstrass a prouvé que la série (IV,11;4) représentait une fonction partout sans dérivée n'était pas du tout une méthode évidente. Le fait que la dérivation terme à terme aboutisse à une série divergente ne prouve absolument pas que la fonction considérée n'ait pas de dérivée. Tout au plus le rend-elle possible !

* On peut dès maintenant le voir en exercice !

Dérivabilité d'un produit infini - **Soit une fonction définie** par un produit infini de fonctions à valeurs réelles ou complexes :

(IV,11;19) $$\Pi(x) = \prod_{n=0}^{\infty} u_n(x) .$$

On peut alors se proposer de déduire de la convergence de ce produit et de la convergence de sa série dérivée logarithmique :

(IV,11;20) $$\sum_{n=0}^{\infty} \frac{u'_n(x)}{u_n(x)} ,$$

le fait que la fonction Π soit dérivable, et que sa dérivée logarithmique soit donnée par la formule

(IV,11;21) $$\frac{\Pi'}{\Pi} = \sum_{n=0}^{\infty} \frac{u'_n}{u_n} .$$

Il est bon de remarquer qu'on ne peut pas, pour le prouver, se ramener immédiatement à ce qui précède par la considération des fonctions $\operatorname{Log} u_\nu$, car il s'agit de fonctions à valeurs complexes, qui ne sont pas nécessairement dans la région $\{z = re^{i\varphi}; r > 0, -\pi < \varphi < \pi\}$ du plan complexe, et par suite il n'est pas certain que l'on puisse prendre des logarithmes.

Mais, si f est une fonction dérivable définie sur un ouvert Ω d'un espace affine norme E , a valeurs complexes, et partout $\neq 0$, on peut prendre sa dérivée logarithmique $\frac{f'}{f}$ sans nécessairement passer par les logarithmes *

On pourra alors faire les remarques suivantes :

1°/ Si f et g sont des fonctions dérivables sur Ω, à valeurs complexes, partout $\neq 0$ la dérivée logarithmique du produit (resp. du quotient) est la somme (resp. la différence) des dérivées logarithmiques

(IV,11;22) $$\frac{(fg)'}{fg} = \frac{f'}{f} + \frac{g'}{g} \quad ; \quad \frac{\left(\frac{f}{g}\right)'}{\frac{f}{g}} = \frac{f'}{f} - \frac{g'}{g} ;$$

* $f(x)$ est un élément du corps $\mathbb{K}$ ($=\mathbb{R}$ ou $\mathbb{C}$), $f'(x)$ est une application linéaire continue de $\vec{E}$ dans $\mathbb{K}$, donc $\frac{f'(x)}{f(x)}$ est aussi une application linéaire continue de $\vec{E}$ dans $\mathbb{K}$ (élément du dual de $\vec{E}$)

c'est en effet tout simplement la formule déja connue

$$(fg)' = f'g + fg' \;;\; \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2} \tag{IV, 11; 23}$$

2°/ Si f est une fonction dérivable sur Ω, à valeurs complexes, la dérivée logarithmique de la fonction e^f est la fonction $\frac{e^f f'}{e^f} = f'$, d'après le théorème des fonctions composées.

3°/ Si Ω est connexe, et si f a une dérivée logarithmique nulle dans Ω, alors elle est constante; en effet sa dérivée f' est nulle.

On en déduit que, si f et g sont deux fonctions qui ont la même dérivée logarithmique, alors leur rapport est une constante. En effet le rapport $\frac{f}{g}$ a une dérivée logarithmique nulle.

4°/ Pour des fonctions complexes f_n sur un espace topologique Ω, nous avons signalé, page 160, que la convergence uniforme des f_n vers f et la convergence uniforme des $\frac{f_n}{f}$ vers 1 (si f est partout $\neq 0$) n'étaient pas équivalentes; mais les convergences localement uniformes correspondantes sont equivalentes, si les f_n sont continues (page 160 du chapitre II).

De même, si f_n (resp. g_n) converge uniformément vers f (resp. g) pour n infini, on ne peut pas en déduire que $f_n g_n$ converge uniformément vers fg * ; mais c'est vrai pour la convergence localement uniforme, si les f_n et g_n sont continues.

* Le produit, application de $\mathbb{C} \times \mathbb{C}$ dans $\mathbb{C}$, n'est pas uniformément continu.

soit en effet $a \in \Omega$. Il existe un voisinage $\mathcal{V}'$ de a où f_n (resp. g_n) converge uniformément vers f (resp. g); un voisinage $\mathcal{V}''$ de a où f et g, continues comme limites localement uniformes de fonctions continues, sont bornées en module.

on a alors :

(IV, 11; 23bis) $$|f_n g_n - fg| \leq |f_n|\,|g_n - g| + |f_n - f|\,|g|$$

Montrons que, dans $\mathcal{V} = \mathcal{V}' \cap \mathcal{V}''$, le 2ème membre converge uniformément vers 0. Le second terme a bien cette propriété. Pour le premier, on remarquera qu'il existe un entier n_0 tel que $|f_n - f| \leq 1$ pour $n \geq n_0$; alors, $|f|$ étant bornée dans $\mathcal{V}''$, $|f_n|$ l'est aussi pour $n \geq n_0$, donc le premier terme converge aussi uniformément vers 0 dans $\mathcal{V}$.

5°/ <u>Si Ω est un intervalle de $\mathbb{R}$, et f une fonction complexe sur Ω, partout $\neq 0$ et de classe C^1, on a la formule</u> :

(IV, 11; 24) $$f(x) = f(a)\, e^{\int_a^x \frac{f'(\xi)}{f(\xi)} d\xi}$$

En effet, le premier et le deuxième membre ont la même dérivée logarithmique $\frac{f'}{f}$ d'après 2°/, donc ils sont proportionnels d'après 3°/, et comme ils coïncident au point a, ils sont égaux.

Nous en retiendrons seulement les majorations suivantes si $\left|\frac{f'}{f}\right| \leq M$:

(IV, 11; 25) $$\left|\frac{f(x)}{f(a)}\right| \leq e^{M|x-a|}$$

et

(IV, 11; 26) $$\left|\frac{f(x)}{f(a)} - 1\right| \leq e^{M|x-a|} - 1$$

SI maintenant f est une fonction complexe de classe C^1 sur un ouvert Ω d'un espace affine E, si f est partout $\neq 0$, et si sa dérivée logarithmique est majorée en norme par M, l'application de (IV,11;25 et 26) à la fonction $t \longmapsto f(a+t\,\overrightarrow{x-a})$, définie sur $[0,1]$, donne les mêmes inégalités, avec $\|\overrightarrow{x-a}\|$ au lieu de $|x-a|$.

On a alors le résultat suivant :

<u>Théorème 11; - soit f_n une suite de fonctions complexes de classe C^1, définies sur un ouvert Ω d'un espace affine normé E, partout $\neq 0$. Si leurs dérivées logarithmiques $\frac{f'_n}{f_n}$ convergent localement uniformément vers une limite g, si les $f_n(a)$ convergent vers une limite $f(a) \neq 0$ * pour au moins un point a de Ω, et si Ω est connexe, alors les f_n convergent vers une fonction limite f localement uniformément dans Ω ; la fonction f n'est jamais nulle, elle est de classe C^1 et admet g comme dérivée logarithmique. Si Ω n'est pas connexe, mais si on sait à l'avance que les f_n convergent simplement vers une fonction partout $\neq 0$ dans Ω, les conclusions subsistent.</u>

<u>Démonstration</u> Posons $f_{m,n} = \frac{f_m}{f_n}$. Soit $B = B(a;\rho)$ une boule de centre a, dans laquelle les $\frac{f'_n}{f_n}$ convergent uniformément vers g. Alors $\frac{f'_{m,n}}{f_{m,n}}$ converge uniformément vers 0 dans B, pour m et n infinis. De (IV,11;26) on déduit que $\frac{f_{m,n}(x)}{f_{m,n}(a)}$ converge uniformément vers 1 ; comme $f_{m,n}(a)$ converge vers 1, $f_{m,n} = \frac{f_m}{f_n}$ converge uniformément vers 1 pour m et n infinis.

En particulier il existe n_0 tel que $n \geq n_0$ entraîne $\left|\frac{f_n}{f_{n_0}} - 1\right| \leq \frac{1}{2}$ donc $\frac{1}{2}\left|f_{n_0}\right| \leq \left|f_n\right| \leq \frac{3}{2}\left|f_{n_0}\right|$ pour $n \geq n_0$, $x \in B$.

Alors, pour tout x de B, on a, pour $m \geq n_0$, $n \geq n_0$:

$$|f_m(x) - f_n(x)| \leq |f_n(x)|\left|\frac{f_m(x)}{f_n(x)} - 1\right| \leq \frac{3}{2}|f_{n_0}(x)|\left|\frac{f_m(x)}{f_n(x)} - 1\right|,$$ donc

$|f_m(x) - f_n(x)|$ converge vers 0 pour m et n infinis. Comme le corps des complexes $\mathbb{C}$ est complet, on en déduit que

* si $f(a) = 0$, on démontre sans peine, si Ω est connexe, que les f_n convergent localement uniformément vers 0.

les $f_n(x)$ ont une limite $f(x)$ pour tout x de B ; en outre, de $|f_n(x)| \geq \frac{1}{2}|f_{n_0}(x)|$ on déduit $|f(x)| \geq \frac{1}{2}|f_{n_0}(x)| > 0$ donc f ne s'annule pas dans B ; enfin, pour $\varepsilon > 0$ donné, si p est tel que $m \geq p$, $n \geq p$, entraîne $\left|\frac{f_m(x)}{f_n(x)} - 1\right| \leq \varepsilon$ pour tout x de B , on voit que $n \geq p$ entraîne $\left|\frac{f(x)}{f_n(x)} - 1\right| \leq \varepsilon$ pour tout x de B , donc $\frac{f}{f_n}$ converge uniformément vers 1 dans B .

Alors, par la même méthode que dans le théorème 111, on voit, si Ω est connexe (ou si l'on sait à l'avance que les f_n convergent simplement vers une fonction f partout $\neq 0$) que f_n converge simplement vers une limite f partout $\neq 0$, et que $\frac{f}{f_n}$ converge localement uniformément vers 1, ou f_n localement uniformément vers f d'après 4°/ page 714 Alors $\frac{f'_n}{f_n}$ converge localement uniformément vers g, et f_n vers f , donc $f'_n = \left(\frac{f'_n}{f_n}\right) f_n$ vers $g f$, puisqu'il s'agit de fonctions continues (remarque 4°/ page 714), donc le théorème 111 dit que f est dérivable et de dérivée $g f$, donc de dérivée logarithmique g, ce qui démontre le théorème.

<u>Corollaire-Si des fonctions u_n complexes sur Ω sont de classe C^1, et partout $\neq 0$, si la série $\sum_{n=0}^{\infty} \frac{u'_n}{u_n}$ converge localement uniformément dans Ω vers une fonction limite g , si d'autre part le produit $\prod_{n=0}^{\infty} u_n$ converge en au moins un point a de Ω , alors, si Ω est connexe, il converge dans tout Ω localement uniformément, et la fonction Π ~~qu'il~~ définit est de classe C^1, et admet g comme dérivée logarithmique. n'est pas connexe, mais si l'on sait à l'avance que le produit $\prod_{n=0} u_n(x)$ converge pour tout-w, de Ω , les conclusions subsistent.</u>

Fonctions représentées par des intégrales

Considérons l'intégrale

$$\vec{f}(x) = \int_T \vec{h}(x,t)\,d\mu(t)\,, \tag{IV,11;26 bis}$$

où μ est une mesure de Radon $\geqslant 0$ sur un espace localement compact T dénombrable à l'infini, et où $\vec{h}$ est une fonction sur $X \times T$, à valeurs dans un espace de Banach $\vec{F}$.

On suppose que, pour tout x de X, la fonction partielle $\vec{h}_x : t \longrightarrow \vec{h}(x,t)$, soit μ-intégrable; nous ne répéterons pas tout cela dans les hypothèses. L'intégrale précédente définit alors un vecteur $\vec{f}(x)$ de $\vec{F}$, et par suite une fonction $\vec{f}$: $x \to \vec{f}(x)$. Nous nous proposons d'étudier la cont'inuité, l'intégrabilité, la dérivabilité de cette fonction à partir des propriétés analogues de $\vec{h}$.

Continuité d'une fonction représentée par une intégrale

Le critère le plus important est tout simplement le théorème 35 de Lebesgue, que nous écrirons sous la forme suivante :

Théorème 114 - Si X est métrisable, si, pour k-presque toutes les valeurs de t, la fonction $\vec{h}$ est séparément continue par rapport à x au point a, si enfin il existe un voisinage $\mathcal{V}$ de a, tel que l'on ait, sur $\mathcal{V} \times T$, une majoration $\|\vec{h}(x,t)\| \leqslant k(t)$, où k est une fonction sur T, $\geqslant 0$ et μ-intégrable, alors $\vec{f}$ est continue au Point a de X.

En effet, lorsque x tend vers a, il finit par se trouver dans $\mathcal{V}$, et alors $\vec{f}(x)$ tend vers $\vec{f}(a)$ quand x tend vers a en vertu du théorème §5 de Lebesgue, et de la remarque qui le suit.

Il faut bien faire attention à ne pas confondre les diverses hypothèses.

On suppose, pour μ-presque toutes les valeurs de t, la continuité de $\vec{h}$ par rapport à x au point a. C'est naturel, puisqu'on cherche une continuité de $\vec{f}$ par rapport à x au point a. Par contre on suppose que $\|\vec{h}\|$ est majorée par une fonction-k de la variable t, μ-intégrable, ce qui est naturel puisqu'on Intègre par rapport à μ sur l'espace **T** de la variable t.

Corollaire - Si $\vec{h}$ est une fonction continue sur $X \times T$ à valeurs dans $\vec{F}$, et si, pour tout a de X, il existe un voisinage $\mathcal{V}$ de a et un compact K de T tels que, pour $x \in \mathcal{V}$ la fonction partielle $\vec{h}_x : t \longrightarrow \vec{h}(x,t)$ ait son support dans **K**, alors $\vec{f}$ est continue de X dans $\vec{F}$.

Démonstration - Lorsque x tend vers a, $\vec{h}_x$ converge vers $\vec{h}_a$ dans $(\vec{F})^K_{cb}$ (théorème 66). Donc il existe un voisinage $\mathcal{V}_0 \subset \mathcal{V}$ de a, tel que $\|\vec{h}_x\|$ soit bornée pour $x \in \mathcal{V}_0$; soit M sa borne. Alors on peut appliquer le théorème précédent, avec $k(t) = M\chi_K(t)$, χ_K étant la fonction caractéristique de **K**.

On a d'ailleurs, directement :

$$\|\vec{f}(x) - \vec{f}(a)\| \leq \mu(K) \sup_{t \in K} \|\vec{h}(x,t) - \vec{h}(a,t)\| \tag{IV,11;26 ter}$$

$$= \mu(K) \|\vec{h}_x - \vec{h}_a\| ,$$

la norme étant prise dans $(\vec{F}^K)_{cb}$, et le théorème 66 donne le résultat, sans s'appuyer sur le théorème de Lebesgue - (ce qui permet de ne pas supposer X métrisable).

Intégrabilité d'une fonction représentée par une intégrale

Le problème est le suivant :

On suppose que X soit lui-même localement compact et muni d'une mesure de Radon $\lambda \geqslant 0$. On se propose alors de chercher si $\vec{f}$ est λ-intégrable, et si son intégrale peut se calculer en intégrant d'abord $\vec{h}$ par rapport à λ pour t fixée, et en intégrant le résultat par rapport à t et à la mesure μ ; autrement dit, si l'on a le droit d'écrire :

$$\text{(IV,11;27)} \qquad \int_X d\lambda(x) \int_T \vec{h}(x,t)\, d\mu(t) = \int_T d\mu(t) \int_X \vec{h}(x,t)\, d\lambda(x).$$

Les critères essentiels pour cela sont tout simplement les théorèmes 77 et 78 de Fubini, et le corollaire du théorème 78; la formule (IV,11;27) est valable si $\vec{h}$ est intégrable sur $X \times T$ par rapport à la mesure produit tensoriel $\lambda \otimes \mu$, ou si elle est mesurable et $\geqslant 0$, ou si elle est mesurable et si l'une des quantités analogues à (IV,11;27), relatives à $\|\vec{h}\|$, est finie.

Dérivabilité d'une fonction définie par une intégrale

Ici $X = \Omega$ sera un ouvert d'un espace affine normé E

Ce que nous avons vu pour les séries nous incite à penser que $\vec{f}$ ne pourra être dérivable que si l'on suppose que l'intégrale relative à la dérivée partielle de $\vec{h}$ par rapport à x possède elle-même des propriétés convenables.

Contrairement à ce qui a été fait du théorème 111 nous ne chercherons pas à traiter le cas général où $\vec{f}$ n'est supposée exister qu'en un point a de Ω.

Théorème 115 - supposons que, pour μ-presque tout t, la fonction partielle $x \longrightarrow \vec{h}(x,t)$ soit dérivable (resp. de classe C^1) sur Ω, soit $x \longrightarrow \frac{\partial h}{\partial x}(x,t)$ sa fonction dérivée, de Ω dans $\mathcal{L}(\vec{E};\vec{F})$. supposons que, pour tout a, la fonction $t \longrightarrow \frac{\partial h}{\partial x}(a,t)$ (définie μ-presque partout sur T, à valeurs dans $\mathcal{L}(\vec{E};\vec{F})$) soit μ-mesurable *, et qu'il existe un voisinage $\mathcal{V}$ de a dans Ω tel que, sur $\mathcal{V} \times T$, on ait une majoration $\left\| \frac{\partial h}{\partial x}(x,t) \right\| \leq k(t)$ où k est une fonction sur T, ≥ 0 et μ-intégrable.

Alors la fonction f est dérivable (resp. de classe C') de Ω dans $\vec{F}$, et sa dérivée est donnée, pour tout a de Ω, par la formule :

(IV,11;28) $$f'(a) = \int_T \frac{\partial h}{\partial x}(a,t)\, d\mu(t) \in \mathcal{L}(\vec{E};\vec{F}).$$

Avant de donner la démonstration de ce théorème, faisons quelques remarques sur les hypothèses de l'énoncé :

Elles assurent l'existence du 2ème membre de (IV,11;28), car la fonction à intégrer, pour a donné, est définie μ-presque partout sur T, mesurable, et majorée par une fonction $k \geq 0$ intégrable, donc intégrable ($\mathcal{L}(\vec{E};\vec{F})$ est complet d'après le théorème 50 du chapitre II).

• Cette mesurabilité est automatiquement vérifiée si E est de dimension finie. Supposons, pour simplifier, $E = \mathbb{K}$, corps des scalaires. Alors, pour μ-presque toutes les valeurs de t, on a $\frac{\partial h}{\partial x}(x,t) = \lim_{\substack{\xi \to 0 \\ \xi \neq 0}} \left(\frac{\vec{h}(x+\xi,t) - \vec{h}(x,t)}{\xi} \right) \in \vec{F}$.
Or les fonctions $t \to \vec{h}(x+\xi,t)$ et $t \longrightarrow \vec{h}(x,t)$ sont μ-mesurables (et même μ-intégrables). Donc, pour x fixé, $t \to \frac{\partial \vec{h}}{\partial x}(x,t)$ est limite μ-presque partout d'une suite de fonctions μ-mesurables, donc elle est μ-mesurable.

Rappelons en outre que toutes les fonctions rencontrées dans la nature sont mesurables!

Ajoutons aussi que si, pour μ-presque toutes les valeurs de t, la fonction $\vec{h}_t$ est, non seulement dérivable, mais de classe C^1 dans Ω, la condition de majoration par k est précisément celle qui entraîne que $\vec{f}'$ soit continue donc $\vec{f}$ de classe C^1 sur Ω. Mais cette condition est nécessaire même pour démontrer la seule dérivabilité de $\vec{f}$.

Démonstration La seule chose à démontrer est donc que $\vec{f}$ admet bien une dérivée-donnée par (IV,11;28). Tout revient à montrer que, quand X tend vers $\vec{0}$, le nombre

(IV,11;29) $$\frac{1}{\|\vec{X}\|}\left\|\int_T\left(\vec{h}(a+\vec{X},t)-\vec{h}(a,t)-\frac{\partial \vec{h}}{\partial x}(a,t).\vec{X}\right)d\mu(t)\right\|$$

tend vers 0 avec $\vec{X}$.

Il est majoré par

(IV,11;30) $$\int_T\left(\frac{1}{\|\vec{X}\|}\left\|\vec{h}(a+\vec{X},t)-\vec{h}(a,t)-\frac{\partial \vec{h}}{\partial x}(a,t).\vec{X}\right\|\right)d\mu(t).$$

La fonction de t à intégrer converge simplement μ-presque partout vers 0 quand X tend vers $\vec{0}$, à cause de la dérivabilité partielle de $\vec{h}$. Le résultat sera donc obtenu en vertu du théorème 35 de Lebesgue, si nous pouvons la majorer par une fonction $\geqslant 0$ μ-intégrable de t. Or, si nous appelons ρ un nombre > 0 tel que la boule B $(a;\rho)$ soit contenue dans $\mathcal{V}$, alors, pour μ-presque toutes les valeurs de t, on peut appliquer la formule des accroissements finis sous la forme du corollaire 1 du théorème 13 du chapitre III, dès que $\|\vec{X}\|\leqslant\rho$, ce qui donne :

(IV,11;31) $$\frac{1}{\|\vec{X}\|}\left\|\vec{h}(a+\vec{X},t)-\vec{h}(a,t)-\frac{\partial \vec{h}}{\partial x}(a,t).\vec{X}\right\|$$

$$\leqslant \sup_{\xi\in]a,a+\vec{X}[}\left\|\frac{\partial \vec{h}}{\partial x}(\xi,t)-\frac{\partial \vec{h}}{\partial x}(a,t)\right\|\leqslant 2k(t);$$

k étant supposée $\geqslant 0$ et μ-intégrable, le théorème est démontré.

La manière de procéder dans la pratique est donc la suivante :

Quand la fonction $\vec{f}$ est définie par l'intégrale (IV,11;26 bis) pour voir si elle est dérivable, on dérive formellement par rapport à x sous le signe d'intégration, en écrivant (IV,11;28); la formule est alors justifiée si les conditions de l'énoncé du théorème, (qui justement, dans le cas où h_t est μ - presque toujours de classe C sur Ω permettent d'affirmer que la fonction représentée par le deuxième membre de (IV,11;28) est continue) sont réalisées.

Voici un corollaire très simple, mais particulièrement important :

Corollaire - Soit $\vec{f}$ une fonction définie par la formule (IV,11;26 bis), avec les hypothèses suivantes:

a) $\vec{h}$ est une application continue de $\Omega \times T$ dans $\vec{F}$;

b) h admet partout une dérivée partielle en x , et $\frac{\partial h}{\partial x}$ est une application continue de $\Omega \times T$ dans $\mathcal{L}(\vec{E};\vec{F})$;

c) pour tout point a de Ω, il existe un voisinage $\mathcal{V}$ de a dans Ω et un compact K de T , tels que, pour tout x de $\mathcal{V}$, la fonction partielle $\vec{h}_x$, fonction continue sur T à valeurs dans , ait son support dans K .

Alors $\vec{f}$ est de classe C^1 dans Ω, et sa dérivée est donnée par (IV,11;28).

Démonstration analogue à celle du corollaire du théorème 114, mais en vérifiant ici que ce sont les conditions de l'énoncé du théorème 115 qui sont vérifiées.

Application aux primitives successives d'une fonction continue sur un intervalle de $\mathbb{R}$.

Revenons à ce que nous avons dit au théorème 91.

La fonction $\vec{F}_m$ définie par la formule (IV,9;41) est une primitive d'ordre m de $\vec{f}$, s'annulant au point c ainsi que ses dérivées jusqu'à l'ordre $m-1$.

Nous l'avons démontré à ce moment par transformation de deux intégrations simples successives en une seule intégration. Nous pouvons en donner maintenant une autre démonstration par vérification directe.

La fonction $\vec{F}_m$ est-elle dérivable ? Bien que très intuitif, ce n'est pas immédiat; elle dépend de x, d'une part parce que x est la borne supérieure d'intégration, d'autre part parce que x intervient sous le signe d'intégration.

Alors $\vec{F}_m(x)$ n'est autre que la valeur pour $y = x$ de la fonction de deux variables définie par

(IV,11;32) $$\vec{G}_m(x,y) = \int_c^y \vec{f}(\xi)\frac{(x-\xi)^{m-1}}{(m-1)!}\,d\xi\ .$$

Montrons que cette fonction admet une dérivée totale. Pour cela nous allons montrer qu'elle admet des dérivées partielles continues en x et y, et il nous suffira alors d'appliquer le théorème 15 du chapitre III (voir plus loin, formule (IV,11;36)). D'abord

(IV,11;33) $$\frac{\overrightarrow{\partial G_m}}{\partial y}(x,y) = \vec{f}(y)\frac{(x-y)^{m-1}}{(m-1)!}$$

d'après le théorème 89; et celle-ci est trivialement continue en (x,y).

Par ailleurs, pour y fixé, le corollaire du théorème 115 est applicable à la fonction $(x,\xi) \longrightarrow \vec{f}(\xi)\frac{(x-\xi)^{m-1}}{(m-1)!}$ (nous avons ici x et ξ au lieu de x et t, $d\xi$ au lieu de $d\mu(t)$, $[c,y]$ pour le compact $K = T$); $\vec{G}_m$ admet donc une dérivée partielle en x donnée par

(IV,11;34) $$\frac{\overrightarrow{\partial G_m}}{\partial x}(x,y) = \int_c^y \vec{f}(\xi)\frac{(x-\xi)^{m-2}}{(m-2)!}\,d\xi\ ,\ \text{si}\ m \geq 2\ .$$

Montrons que $\frac{\overrightarrow{\partial G_m}}{\partial x}$ est elle aussi continue en (x,y) * .
Fixons $(x_0\,,\,y_0)$; et supposons par exemple $y_0 \geq c$. On peut écrire

(IV,11;35) $$\frac{\overrightarrow{\partial G_m}}{\partial x}(x,y) = \int_{\mathbb{R}_1} \varphi_y(\xi)\vec{f}(\xi)\frac{(x-\xi)^{m-2}}{(m-2)!}\,d\xi\ ,$$

* En fait, il est inutile de le démontrer, d'après la remarque 1° qui suit le théorème 15 du chapitre III.

où φ_y est la fonction caractéristique de l'intervalle $[c, y]$. Pour ξ fixé, lorsque (x, y) tend vers (x_0, y_0), $\varphi_y(\xi)\vec{f}(\xi)\frac{(x-\xi)^{m-2}}{(m-2)!}$ converge vers $\varphi_{y_0}(\xi)\vec{f}(\xi)\frac{(x_0-\xi)^{m-2}}{(m-2)!}$, sauf peut être pour $\xi = y_0$. (parce qu'alors $\varphi_y(\xi)$ ne converge pas vers $\varphi_{y_0}(\xi)$ si y tend vers y_0 par valeurs $< y_0$). Donc, pour $d\xi$ - presque toutes les valeurs de ξ (toutes sauf une, y_0), la fonction partielle $(x,y) \longrightarrow \varphi_y(\xi)\vec{f}(\xi)\frac{(x-\xi)^{m-2}}{(m-2)!}$ est continue au point (x_0, y_0)

Par ailleurs, si x et y parcourent des voisinages compacts de x_0, y_0 dans $\mathbb{R}_1$, la fonction à intégrer reste bornée, et à support dans un compact fixe, donc le théorème 114 montre bien que $\frac{\partial \vec{G}_m}{\partial x}$ est continue au point x_0, y_0.

Donc $\vec{G}_m$ est bien de classe C' sur $\mathbb{R}_1 \times \mathbb{R}_1$; mais alors le fonction $\vec{F}_m$ est elle-même de classe C' d'après le théorème des fonctions composées (corollaire 5 du théorème 11) et sa dérivée se calcule par la formule

$$\frac{d\vec{F}_m}{dx}(x) = \frac{\partial \vec{G}_m}{\partial x}(x,x) + \frac{\partial \vec{G}_m}{\partial y}(x,x) \quad , \tag{IV,11;36}$$

ce qui donne tout simplement, si $m \geq 2$:

$$\frac{d\vec{F}_m}{dx}(x) = \frac{\partial \vec{G}_m}{\partial x}(x,x) = \int_c^x \vec{f}(\xi)\frac{(x-\xi)^{m-2}}{(m-2)!}\,d\xi \quad , \tag{IV,11;37}$$

compte tenu de ce que $(x-y)^{m-2}$ s'annule pour $y = x$.

On peut ainsi continuer de proche en proche, pour toutes les dérivées jusqu'à l'ordre $m-1$ inclusivement, et obtenir les formules

$$\frac{d^k}{dx^k}\vec{F}_m(x) = \int_c^x \vec{f}(\xi)\frac{(x-\xi)^{m-k-1}}{(m-k-1)!}\,d\xi = \vec{F}_{m-k}(x), k \leq m-1 \quad , \tag{IV,11;38}$$

qui prouvent bien que $\vec{F}_m$ a toutes ses dérivées jusqu'à l'ordre $m-1$ nulles au point c.

Calculons maintenant la dérivée d'ordre-de $\vec{F}_m$.

Nous appliquons simplement à la fonction $\vec{F}_1$ la même formule, mais cette fois, nous obtenons

(IV,11;39)
$$\vec{F}_1(x) = \int_c^x \vec{f}(\xi)\,d\xi \; ,$$

$$\frac{d\vec{F}_1}{dx}(x) = \vec{f}(x) \; .$$

Nous avons donc bien montré que la dérivée d'ordre m de $\vec{F}_m$ est égale à $\vec{f}$, et que par suite la fonction $\vec{F}_m$ définie par (IV,9;41) est bien une primitive d'ordre m de $\vec{f}$, s'annulant au point c ainsi que ses dérivées jusqu'à l'ordre $m-1$ inclusivement, ce qui donne une nouvelle démonstration du théorème 91.

Cas des intégrales impropres convergentes

Soir. $\vec{h}$ une fonction sur un produit $X \times \mathbb{R}_1$, à valeurs dans un espace de Banach $\vec{F}$, $\mathbb{R}_1$ étant un intervalle $]a,b[$ de $\mathbb{R}$. Considérons 1 intégrale

(IV,11;40)
$$\vec{f}(x) = \int_{]a,\to b[} \vec{h}(x,t)\,d\mu(t)$$

On suppose que μ est une mesure de Radon sur $\mathbb{R}_1$, et que, pour tout x de X, l'intégrale impropre qui définit $\vec{f}(x)$ (voir définition page 044) est convergente. On dit que l'intégrale impropre est uniformément convergente lorsque x parcourt une partie A de X, si l'intégrale propre $\int_{]a,b'[}$ converge vers l'intégrale impropre $\int_{]a,\to b[}$ (lorsque $b' < b$ tend vers b), uniformément pour $x \in A$. Cela veut encore dire que le "reste" $\int_{]b',\to b[}$ converge vers 0 lorsque b' tend vers b, uniformément pour $x \in A$; ou encore que, quel que soit $\varepsilon > 0$, il existe $b_0 < b$ (indépendant de x), tel que $b_0 \leq b' < b$ entraîne $\left\| \int_{]b',\to b[} \right\| \leq \varepsilon$, quel que soit $x \in A$.

Les propriétés de $\vec{f}$ s'obtiennent en 2 temps. On démontre d'abord que l'intégrale propre $\int_{[a,b'[}$ a les propriétés désirées; on fait ensuite tendre b' vers b .

Sans nous occuper du cas le plus général possible, nous donnerons alors les théorèmes suivants :

Théorème 116 - Soit $\vec{f}$ une fonction sur X définie par (IV,11;40), où l'intégrale impropre est supposée convergente pour tout x de X. On suppose que :

1°/ X est un espace métrisable; pour tout $b' < b$, l'intégrale propre

$$(IV,11;41) \qquad \vec{f}_{b'}(x) = \int_{|a,b'[} \vec{h}(x,t)\,d\mu(t)$$

définit une fonction $\vec{f}_{b'}$ sur X à valeurs dans $\vec{F}$, continue au point $x = x_0$ de X (par exemple par application du théorème 114);

2°/ Il existe un voisinage $\mathcal{V}$ de x_0 tel que l'intégrale impropre soit uniformément convergente pour x dans $\mathcal{V}$.

Alors la fonction $\vec{f}$ est continue sur X au point x_0 de X

C'est évident : on suppose $\vec{f}_{b'}$ continue sur $\mathcal{V}$ au point x_0, et $\vec{f}_{b'}$ converge vers $\vec{f}_b = \vec{f}$ uniformément sur $\mathcal{V}$; il suffit donc d'appliquer le théorème 65 du chapitre II.

Théorème 117 - Soit une intégrale impropre (IV,11;40), où $X = \Omega$ est un ouvert d'un espace affine normé E. On suppose que:

1° / Pour tout $b' < b$, l'intégrale propre (IV,11;41) représente une fonction $\vec{f}_{b'}$ sur Ω, vérifiant les conditions du théorème 115, permettant d'affirmer que $\vec{f}_{b'}$ est dérivable (resp. de classe C^1) sur Ω, et que sa dérivée est donnée par la formule :

$$(IV,11;42) \qquad \vec{f}'_{b'}(x) = \int_{|a,b'[} \frac{\partial \vec{h}}{\partial x}(x,t)\,d\mu(t);$$

2°/ l'intégrale impropre

$$(IV,11;43) \qquad g(x) = \int_{|a,\to b[} \frac{\partial \vec{h}}{\partial x}(x,t)\,d\mu(t)$$

est convergente pour tout x de Ω, et localement uniformément sur Ω ;

3°/ Ω est connexe, F complet, et l'intégrale impropre (IV,11;40) est convergente pour au moins un point $x = a$ de Ω ; ou Ω et F sont quelconques, et cette intégrale impropre est convergente pour tout x de Ω .

Alors l'intégrale impropre (IV,11;40) est convergente pour tout x de Ω, cette convergence est localement uniforme sur Ω , la fonction $\vec{f}$ définie par (IV,11;40) est dérivable (resp. de classe C^1) sur Ω , et sa dérivée est la fonction $\vec{g}$.

En effet, lorsque b' tend vers b , les fonctions $\vec{f}_{b'}$ vérifient les conditions requises par le théorème 111 *

Exemple 1 - Considérons l'intégrale (où $x \in \mathbb{R}$) :

(IV,11;44) $$f(x) = \int_{\mathbb{R}} \frac{e^{itx}}{1+t^2}\,dt$$

La fonction $t \longrightarrow \frac{1}{1+t^2}$ est intégrable sur $\mathbb{R}$, alors f est continue sur $\mathbb{R}$ par application immédiate du théorème 114, avec $\mathcal{V} = \mathbb{R}$, et $k(t) = \frac{1}{1+t^2}$.

Dérivons par rapport à x sous le signe d'intégration. On obtient :

(IV,11;45) $$\int_{\mathbb{R}} \frac{ite^{itx}}{1+t^2}\,dt ,$$

qui, cette fois, n'a pas de sens, car $\frac{t}{1+t^2}$ n'est pas intégrable. On ne peut donc pas, par la seule considération des intégrales propres (théorème 115), prouver la dérivabilité de f .

Mais considérons l'intégrale impropre :

(IV,11;46) $$g(x) = \int_{]-\infty \leftarrow, \rightarrow +\infty[} \frac{it\,e^{itx}}{1+t^2}\,dt$$

* Il ne s'agit pas ici d'une suite $\vec{f}_n$, mais d'une famille $\vec{f}_{b'}$, dépendant du paramètre réel b' , et d'une limite lorsque b' tend vers b . On s'assure aisément que tout reste valable

Pour $x \neq 0$, elle est semi-convergente, par application du théorème d'Abel. En effet, la fonction $t \longrightarrow \frac{t}{1+t^2}$ (dont la dérivée est $\frac{1-t^2}{(1+t^2)^2}$) est monotone dans chaque intervalle $]-\infty,-1]$, $[-1,+1]$, $[+1,+\infty[$; elle est continue et tend vers 0 pour t infini, donc bornée, et à variation bornée d'après le corollaire du théorème **86.** Et la fonction $t \longrightarrow e^{itx}$ est "à intégrales Indéfinies bornées" (majoration (IV,9;107)) pour $x \neq 0$:

(IV,11;47) $$|\sigma_{c,d}| = \left| \int_c^d e^{itx}\,dt \right| \leq \frac{2}{|x|}$$

Alors le critère d Abel (corollaire du **théorème** 98) nous donne bien la semi-convergence de (IV,11;46) pour $x \neq 0$. <u>Cette semi-convergence est localement uniforme sur l'ouvert</u> $\Omega = \complement 0$ de $\mathbb{R}$: en effet, pour tout $\delta > 0$ donné, pour $c \geq 1$, $|x| \geq \delta$, la **majoration (avec** les notations du théorème d'Abel) :

(IV,11;47) $$\left| \int_{]-\infty\leftarrow,-c]} \right| \text{ et } \left| \int_{[+c,\rightarrow+\infty[} \right|$$

$$\leq U(c)\,V(c) = \frac{c}{1+c^2}\,{}^{*}\,\frac{2}{\delta} \;;$$

quantité qui tend verso pour c infini. 8 étant fixé.

Alors, pour tout c fini, l'intégrale propre

(IV,11;48) $$f_c(x) = \int_{[-c+c]} \frac{e^{itx}}{1+t^2}\,dt$$

est de classe C^1 sur Ω (et même sur $\mathbb{R}$) par application immédiate du théorème 115 ou même de son corollaire. Nous pouvons donc faire tendrecvers + ∞ et appliquer le théorème 117; donc f est dérivable dans $\Omega = \complement 0$, et sa dérivée est donnée par l'intégrale impropre g de (IV,11;46).

* pour $c \geq 1$, la fonction $\frac{t}{1+t^2}$ est monotone dans $[C,+\infty[$, et sa variation totale est exactement $\frac{C}{1+C^2}$

Si maintenant nous cherchons formellement la dérivée seconde en calculant

$$\int \frac{-t^2 e^{itx}}{1+t^2}\, dt \,, \tag{IV,11;49}$$

nous obtenons une expression qui n'a plus aucun sens, d'aucune manière.

Nous verrons plus tard que la fonction f est donnée par

$$f(x) = \pi e^{-|x|} \tag{IV,11;50}$$

Elle est bien une fonction continue sur $\mathbb{R}$, elle est bien de classe C' sur le complémentaire de l'origine **mais** n'admet pas de dérivée à l'origine; mais nous voyons qu'elle est aussi de classe C^∞, c'est-à-dire admet des dérivées successives de tous les ordres, sur le complémentaire de l'origine; ici, comme déjà dans la remarque 1°. page 711, les théorèmes précédents permettent de prévoir qu'une fonction représentée par une intégrale est dérivable, mais ne permettent pas de prévoir qu'elle ne l'est pas.

A titre de deuxième exemple, nous allons démontrer une formule très importante dans la théorie des séries des intégrales de Fourier :

Exemple 2 -

Théorème 118

On a la formule :

$$\int_{[0,\to+\infty[} \frac{\sin \lambda t}{t}\, dt = \frac{\pi}{2} \,, \quad \lambda > 0 \,. \tag{IV,11;51}$$

Cette formule a déjà été annoncée à (IV,9;110).

En posant $\lambda t = t'$, on se ramène au cas où $\lambda = 1$.

Pour la démontrer, nous considérerons la fonction J définie par

$$J(x) = \int_{[0,\to+\infty[} e^{-xt} \frac{\sin t}{t}\, dt \,. \tag{IV,11;52}$$

L'intégrale qui figure au deuxième membre est celle d'une fonction intégrable, si $x > 0$, elle se ramène donc à une intégrale propre (théorème 97). Pour $x = 0$ il n'en est pas ainsi, il s'agit d'une intégrale impropre convergente d'après le théorème d'Abel.

Nous allons d'abord démontrer que l'intégrale définie au 2ème membre est une Intégrale impropre qui converge uniformément pour $x \geq 0$. Si en effet, nous appliquons le théorème d'Abel à $\int_{[c,\to+\infty[}$, $c > 0$, avec

$$(IV,11;53) \qquad u(t) = \frac{e^{-xt}}{t}, \qquad v(t) = \sin t ,$$

u est décroissante, et tend vers 0 pour t infini, donc est à variation bornée, et v est à intégrales indéfinies bornées, avec $\left|\int_{[c,d]} \sin t\, dt\right| \leq 2$.

On a donc Immédiatement la majoration (IV,9;102) :

$$(IV,11;54) \qquad \left|\int_{[c,\to+\infty[} e^{-xt}\frac{\sin t}{t}\,dt\right| \leq \frac{2e^{-xc}}{c} \leq \frac{2}{c} ,$$

quantité qui tend bien vers 0 pour c infini, uniformément pour $x \geq 0$. Comme, pour c fixé fini, $J(x,c) = \int_{[0,c]}$ est Une fonction continue dex d'après le corollaire du théorème 114 , nous voyons que J est une fonction continue sur la demi-droite $\mathbb{R}_+$ des nombres réels ≥ 0 , d'après le théorème 116. Cherchons si J admet une dérivée. Dérivons formellement par rapport à x sous le signe d'intégration :

$$(IV,11;55) \qquad J'(x) = -\int_{[0,\to+\infty[} e^{-xt}\sin t\,dt .$$

L'intégrale qui est au deuxième membre est alors celle d'une fonction lntégrable pour $x > 0$, alors que, pour $x = 0$, elle est une intégrale impropre divergente. Il n'est donc plus nécessaire d'utiliser ici le théorème d'Abel, puisqu il est impossible de toute façon d'obtenir des informations sur la dérivée à l'origine. Mais on utilisera simplement le théorème 115. Pour toutes les valeurs de t , la fonction $x \to e^{-xt}\frac{\sin t}{t}$ est de classe C^1

sur $\Omega =]0, +\infty[$, de dérivée $-e^{-xt}\sin t$, et on a , pour $x \geq \delta > 0$, la majoration

$$\left|e^{-xt}\sin t\right| \leqslant e^{-\delta t} \quad \text{, fonction } \geq 0 \text{ intégrable} \tag{IV,11;56}$$

Le théorème 115 nous montre que par conséquent J est de classe C^1 sur $\Omega =]0, +\infty[$, et que sa dérivée est donnée par (IV,11;55).

Mais alors l'intégrale qui figure au deuxième membre peut se calculer élémentairement. Elle s'écrit en effet, en effectuant 2 intégrations par parties (théorème 98) :

$$J'(x) = \cancel{\left[-\frac{e^{-xt}}{-x}\sin t\right]_0^{+\infty}} - \int_{[0,\rightarrow+\infty[} \frac{e^{-xt}}{x}\cos t\, dt \tag{IV,11;57}$$

$$= \left[-\frac{e^{-xt}}{-x^2}\cos t\right]_0^{+\infty} \int_{[0,\rightarrow+\infty[} \frac{e^{-xt}}{x^2}\sin t\, dt$$

$$= -\frac{1}{x^2} - \frac{J'(x)}{x^2}$$

d'où l'on déduit

$$J'(x) = -\frac{\quad}{+x^2} \text{ , } x > 0 \text{ .} \tag{IV,11;58}$$

Remarquons d'ailleurs que $J'(x)$ admet une limite lorsque x tend vers 0 . cela prouve donc, d'après le théorème 14 du chapitré III, que J admet encore une , dérivée pour $x = 0$, et que cette dérivée est égale à-l.

Mais nous voyons que nous ne pouvions pas le prévoir, ni représenter $J'(0) = -1$ par l'intégrale $-\int_{[0,\rightarrow+\infty[} \sin t\, dt$, qui n'a aucun sens.

On en déduit que J s'obtient par la recherche de la primitive du deuxième membre de (IV,11;58), ce qui donne

$$J(x) = -\operatorname{Arg\,tg} x + C \text{ .} \tag{IV,11;59}$$

La constante s'obtient très facilement.

En effet, lorsque x tend vers $+\infty$, la fonction $t \longrightarrow e^{-xt}\frac{\sin t}{t}$ converge simplement vers 0, et comme elle reste majorée par la fonction5 0 intégrable fixe $t \longrightarrow e^{-t}\frac{|\sin t|}{t}$, dès que $x \geq 1$, le théorème de Lebesgue montre que $J(x)$ tend vers 0 lorsque x tend vers $+\infty$.

La constante du deuxième membre de (IV,11;59) est donc nécessairement $\frac{\pi}{2}$; par conséquent, nous avons, à cause de la continuité de J même à l'origine, $J(0) = \frac{\pi}{2}$, et ceci démontre le théorème. Dans la théorie des fonctions eulériennes, nous verrons une généralisation de cette formule.

Application à la divisibilité des fonctions dérivables

Théorème 119 - Soit $\vec{U}$ une fonction définie sur un ouvert Ω de $\mathbb{R}^N$, * à valeurs dans un espace de Banach F. On suppose que $\vec{U}$ s'annule sur l'intersection de Ω avec l'hyperplan $x_N = 0$, au moins à l'ordre , c'est-à-dire en même temps que ses dérivées jusqu'à l'ordre $k-1$ inclusivement.

Si alors $\vec{U}$ est de classe C^m avec $m \geq k$, il existe une fonction $\vec{V}$, déterminée d'une manière unique, définie sur Ω, a valeurs dans F et de classe C^{m-k}, telle que l'on ait l'identité

$$\vec{U}(x) = x_N^k\,\vec{V}(x)\ . \tag{IV,11;60}$$

On voit qu'on peut dire, dans un certain sens, que le fait, pour $\vec{U}$, de s'annuler sur l'hyperplan $x_N = 0$ à l'ordre k, entraîne que $\vec{U}$ soit "divisible par x_N^k". Toutefois il faut remarquer que le quotient n'est pas d'une classe de différentiabilité aussi élevée que $\vec{U}$.

Démonstration - 1°/ Si une fonction telle que $\vec{V}$ existe, elle est évidemment déterminée d'une manière unique; car, en tout point où $x_N \neq 0$, elle est donnée par le quotient $\vec{V}_0(x) = \frac{\vec{U}(x)}{x_N^k}$; elle est nécessairement continue puisqu'on

* On pourrait partout remplacer $\mathbb{R}$ par $\mathbb{C}$.

a supposé $m \geq k$ donc $m-k \geq 0$, et par suite elle est aussi déterminée en un point où $x_N = 0$ par passage à la limite * .

2°/ Montrons maintenant l'existence de $\vec{V}$. Pour simplifier, supposons que Ω soit convexe, de façon à pouvoir appliquer la formule de Taylor ** . Pour tout point x de Ω , de coordonnées $x_1, x_2, \ldots, x_N$, nous pouvons appliquer la formule de Taylor jusqu'à l'ordre $k-1$, relative au point 0 et à 1 accroissement x_N , à la fonction d'une variable $u \longrightarrow \vec{U}(x_1, x_2, \ldots, x_{N-1}, u)$. Comme les dérivées sur H, jusqu'à l'ordre $k-1$ inclusivement, sont nulles, cette formule se réduit au terme complémentaire, et nous pouvons ainsi écrire (formule (IV,9;49)) :

$$\vec{U}(x_1, x_2, \ldots, x_N) = x_N^k \int_0^1 \frac{\partial^k \vec{U}}{\partial x_N^k}(x_1, x_2, \ldots, x_{N-1}; t x_N) \frac{(1-t)^{k-1}}{(k-1)!} dt . \tag{IV,11;61}$$

Nous voyons alors que nous pouvons prendre pour $\vec{V}$ la fonction donnée par l'intégrale

$$\vec{V}(x_1, x_2, \ldots, x_N) = \int_0^1 \frac{\partial^k \vec{U}}{\partial x_N^k}(x_1, x_2, \ldots, x_{N-1}, t x_N) \frac{(1-t)^{k-1}}{(k-1)!} dt \tag{IV,11;62}$$

* C'est le théorème 45 du chapitre II, 1°/ de la démonstration.

** L'existence de $\vec{V}$ est assurée, si $\vec{V}_0(x) = \frac{\vec{U}(x)}{x_N^k}$,définie pour $x_N \neq 0$, a une limite lorsque x tend vers un point a tel que $a_N = 0$, et si la fonction, égale à $\vec{V}_0(x)$ pour $x_N \neq 0$, et à la limite précédente lorsque $x_N = 0$, est de classe C^{m-k} . Donc on peut, pour chaque point a tel que $a_N = 0$, se borner à considérer ce qui se passe dans un voisinage de a ; ce qui permet de se restreindre à un voisinage ouvert convexe.

Elle est manifestement une fonction de classe C^{m-k} dans Ω.

En effet, $\vec{U}$ Qtant supposée de classe C^m, sa dérivée $\frac{\partial^k \vec{U}}{\partial x_n^k}$ est de classe C^{m-k}, et alors nous pouvons utiliser, pour calculer les dérivées d'ordre $\leq m-k$ de $\vec{V}$, le procédé de dérivation sous le signe d'intégration; nous sommes ici dans le cas élémentaire du corollaire du théorème 115; et ceci démontre le théorème.

Remarques

1°/ Il est essentiel que $\vec{F}$ soit un Banach, c'est-a-dire soit complet. S'il est incomplet, $\vec{V}(x)$ est toujours un élément de F lorsque x n'est pas dans l'hyperplan $x_N = 0$, puisqu'on a $\vec{V}(x) = \frac{\vec{U}(x)}{x_N^k}$; mals, pour x dans l'hyperplan, l'intégrale (IV,11;62) n'a peut être aucun sens; la limite de $\vec{V}(x)$ lorsque x tend vers un point de l'hyperplan $x_N = 0$ n'a plus de raison d'exister.

2°/ On pourrait espérer que la fonction V soit d'une classe $C^{m'}$ où $m' >$l-l-b--0e. Il n'en est rien et le nombre $m-k$ qui est donné ici est le meilleur possible.

Considérons par exemple le cas particulier $N=1$ $(\mathbb{R}^N = \mathbb{R})$, et $\vec{F} = \mathbb{R}$ aussi, et considérons la fonction U définie par

(IV,11;63) $$u(x) = |x|^{m+\frac{1}{2}}$$

Elle est manifestement de classe C^m, mais n'est pas de classe C^{m+1}. Elle est nulle à l'origine à l'ordre $m+1$, donc en particulier à tout ordre $k \leq m$; or le quotient V s'écrit ici

(IV,11;64) $$V(x) = |x|^{m-k+\frac{1}{2}} \quad ;$$

il est de classe C^{m-k} mais n'est pas de classe C^{m-k+1}

<u>Corollaire</u> - <u>Soit</u> U <u>une fonction de classe</u> C^m <u>sur un ouvert</u> Ω <u>d'un espace affine</u> E <u>de dimension finie, à valeurs dans un espace de Banach</u> $\vec{F}$. <u>Soit</u> Σ <u>une hypersurface de classe</u> C^m <u>contenue dans</u> Ω, <u>et soit</u> $f(x) = 0$ <u>une équation normale de cette hypersurface</u> *. <u>Si</u> $\vec{U}$ <u>est nulle sur l'hypersurface</u>

* Voir corollaire du théorème 33 bis du chapitre III.

<u>au moins à l'ordre k</u> , <u>alors il existe une fonction</u> $\vec{V}$ <u>définie sur Ω,</u> <u>à valeurs dans $\vec{F}$</u> <u>de classe C^{m-k}</u>, <u>déterminée d'une manière unique,</u> <u>telle que l'on ait</u>

(IV,11;65) $$\vec{U}(x) = (f(x))^k \vec{V}(x)$$

<u>Démonstration</u> Ici, comme dans la démonstration du théorème (note 2°/ page 734), il nous suffit de donner, pour chaque point a de l'hypersurface, une démonstration relative à un voisinage quelconque de ce point. Choisissons un référentiel de E . Si alors nous prenons un point a de Σ pomme $f(x)=0$ est supposée être une équation normale de Σ , l'une au moins des dérivées partielles de f par rapport à $x_1, x_2, \ldots, x_N$, est $\neq 0$. Supposons par exemple que ce soit $\frac{\partial f}{\partial x_N}(a)$.

Considérons alors les fonctions définies par

(IV,11;66) $$u_1 = x_1 \ , \quad u_2 = x_2 \ , \ldots, u_{N-1} = x_{N-1} \ ,$$

$$u_N = f(x_1, x_2, \ldots, x_N)$$

Elles définissent une application Φ de classe C^m de Ω dans $\mathbb{R}^N$. Par ailleurs le déterminant jacobien de cette application est $\neq 0$ au point a , puisqu'il est égal à $\frac{\partial f}{\partial x_N}(a)$, et par suite il existe, d'après le théorime des fonctions implicites (corollaire 2 du théorème 31 du chapitre III) un voisinage ouvert ω de a dans Ω , tel que la restriction de Φ a ω soit un C^m -difféomorphisme de ω sur un ouvert ω_1 de $\mathbb{R}^n$. Désignons par Φ^{-1} son difféomorphisme réciproque. Φ nous permet alors de transporter toute la figure de ω sur ω_1 .

L'intersection de Σ avec ω vient sur une hypersurface $\Sigma_1 = \Phi(\Sigma)$, qui cette fois n'est autre que l'intersection avec ω_1 de l'hyperplan $u_N = 0$.

Le difféomorphisme Φ transport5 la fonction U sur ω à valeurs dans F sur une fonction U_1 sur ω_1 a valeurs dans? , qui est $\vec{U}_1 = \vec{U} \circ \Phi^{-1}$.

On vérifie facilement que cette fonction U_1 s'annule aussi a l'ordre k sur l'hyperplan $u_N = 0$. •

Il résulte alors du théorème qu'il existe, dans ω_1, une fonction $\vec{V}_1$ définie d'une manière unique, de classe C^{m-k}, telle que l'on ait l'identité $\vec{U}_1 = u_N^k \vec{V}_1$. Mais alors $\vec{V}_1$ est la transportée par Φ d une fonction? sur ω qui est définie par $\vec{V} = \vec{V}_1 \circ \Phi$; la fonction V sur ω est de classe C^{m-k}, et l'on a, dans ω , l'identité $\vec{U} = f^k \vec{V}$, et le théorème est démontré.

<u>Corollaire 2 - Soit Σ une hypersurfaoe de classe C^∞ d'un ouvert Ω d'un espace affine E de dimension finie, ayant une équation normale $f(x) = 0$. On obtient toutes les autres équations normales $g(x) = 0$, en posant $g = fV$, où V est une fonction scalaire de classe C^∞ sur Ω , partout $\neq 0$.</u>

<u>Démonstration</u> D'abord, si $g = fV$, et si V est de classe C^∞ et sans zéros, g est bien de classe C^∞ et $g(x) = 0$ définit bien Σ ; en outre, pour $a \in \Sigma$, on a $\overleftarrow{g'(a)} = \overleftarrow{f'(a)}\, V(a) + f(a)V'(a) = \overleftarrow{f'(a)}\, V(a) \neq 0$, donc $g(x) = 0$ est encore une équation normale de Σ .

Réciproquement, soit $g(x) = 0$ une équation normale de Σ Alors g est de classe C^∞ et nulle sur Σ , donc $g = fV$, où V est de classe C^∞, d'après le corollaire. Mais on peut aussi dire que $g(x) = 0$ est une équation normale de Σ et que f est de classe C^∞ et s'annule sur Σ , donc $f = gW$, où W est de classe C^∞ . Alors $g = fV = gWV$, d'où

• On a $\vec{U}_1 = \vec{U} \circ \Phi^{-1}$, donc les dérivées de U_1 sont données par le théorème des fonctions composées: elles sont compliquées, mais on démontre <u>simplement, par récurrence</u> sur k qu'elles sont toutes nulles jusqu'à l'ordre $k-1$.

$WV = 1$ là où g ne s'annule pas, c'est-à-dire en dehors de Σ. Comme W et V sont continues, WV l'est aussi, donc $WV \equiv 1$ partout, donc V est partout $\neq 0$, et le corollaire est démontré.

Remarques 1°/ Il est remarquable que, pour de telles propriétés de pur calcul différentiel, le calcul intégral soit l'outil utilisé.

2°/ Apartir du corollaire 2, on peut démontrer (mais c'est très difficile) que toute hypersurface Σ fermée de classe C^∞ d'un espace affine de dimension finie peut être définie tout entière par une seule équation normale $f(x) = 0$.

Pour simplifier, nous ne donnerons le théorème suivant que pour $k = 1$.

Théorème 120 - Soit Ω un ouvert de $\mathbb{R}^N$, $\vec{U}$ une fonction de classe C^m sur Ω, à valeurs dans un espace de Banach $\vec{F}$. On suppose que $\vec{U}$ s'annule sur l'intersection de Ω avec le sous-espace vectoriel à n dimensions d'équations $x_{n+1} = \cdots = x_N = 0$. Alors on peut trouver (en général d'une infinité de manières) des fonctions $V_{n+1}, V_{n+2}, \ldots, V_N$, de classe C^{m-1} sur Ω à valeurs dans $\vec{F}$, telles que l'on ait :

$$\vec{U}(x) = x_{n+1}\vec{V}_{n+1}(x) + \cdots + x_N \vec{V}_N(x) \quad \text{(IV,11;67)}$$

Démonstration La démonstration est analogue à celle du théorème 119; toutefois on applique la formule des accroissements finis (formule de Taylor pour $k-1 = 0$) à la fonction de $N-n$ variables : $(u_{n+1}, u_{n+2}, \ldots u_N) \longrightarrow \vec{U}(x_1, x_2, \ldots, x_n, u_{n+1}, \ldots, u_N)$. On obtient alors

(IV,11;68) $$U(x_1, x_2, \ldots, x_N) = \int_0^1 \left(U'(x_1, x_2, \ldots, x_n, tx_{n+1}, \ldots, tx_N) \cdot (0, \ldots, 0, x_{n+1}, \ldots, x_N) \right) dt$$
$$= \sum_{i=n+1}^{N} x_i \int_0^1 \frac{\partial \vec{U}}{\partial x_i} (x_1, x_2, \ldots, x_n, tx_{n+1}, \ldots, tx_N)\, dt$$

On peut donc prendre, pour $i = n+1, \ldots, N$:

(IV,11;69) $$\vec{V}_i(x) = \int_0^1 \frac{\partial \vec{U}}{\partial x_i} (x_1, x_2, \ldots, x_n, tx_{n+1}, \ldots, tx_N)\, dt ,$$

ce qui démontre le théorème

Remarque - Ici naturellement, si $n \leq N-2$, il n'y a pas une solution unique. On peut rajouter à $\vec{V}_{N-1}$ la fonction x_N et retrancher à V_N la fonction x_{N-1} ; cela ne fait qu'ajouter et retrancher $x_{N-1} x_N$ à $\vec{U}$

Corollaire - Soit V une variété de classe C^m d'un ouvert Ω d'un espace affine E de dimension finie. ayant un système normal d'équations $f_1(x) = 0$, $f_2(x) = 0, \ldots, f_\ell(x) = 0$. Soit $\vec{U}$ une fonction de classe C^m sur Ω , à valeurs dans un espace de Banach $\vec{F}$, nulle sur V . Alors il existe (en général d'une infinité de manières) un système de fonctions $\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_\ell$ de classe C^{m-1} sur Ω à valeurs dans F , tel que

(IV,11;70) $$\vec{U} = f_1 \vec{V}_1 + \cdots + f_\ell \vec{V}_\ell$$

Démonstration analogue à celle du corollaire 1 du théorème 119, avec quelques complications sur lesquelles nous n'insisterons pas, nécessitant une partition de l'unité.

V
ÉQUATIONS DIFFÉRENTIELLES

§ 1 POSITION DU PROBLÈME

Qu'entend-t-on par équation différentielle de la forme :

(V,1;1) $$\vec{y}' = \vec{L}(x,y)\ ?$$

On se donne un intervalle $|a,b|$, ouvert, semi-ouvert ou fermé, de la droite réelle $\mathbb{R}$ (*), un ouvert Ω d'un espace affine normé F (**), et une application continue $\vec{L}$ de $|a,b| \times \Omega$ dans F ; on cherche alors s'il existe une fonction $f : y = f(x)$, définie sur $|a,b|$, à valeurs dansa , dérivable, telle que, si f' désigne sa fonction dérivée, on ait l'identité :

(V,1;2) $$\vec{f}'(x) \equiv \vec{L}(x, f(x))\ .$$

Cela Implique obligatoirement que la fonction $\vec{f}'$ soit continue, d'après le théorème des fonctions composées, et que par conséquent la fonction f soit, non seulement dérivable, mais continûment dérivable.

(*) a et b peuvent être $-\infty$ et $+\infty$, mals alors $|a,b|$ ne les contient pas; on doit avoir $|a,b| \subset \mathbb{R}$ et non $\bar{\mathbb{R}}$.

(**) F est un espace affine sur le corps des réels ou des complexes, mals $|a,b|$ est toujours un intervalle de $\mathbb{R}$; $\overrightarrow{f'(x)}$ est le vecteur dérivé, élément de $\vec{F}$, suivant la définition (III,3;1). Si on remplace $\mathbb{R}$ par un espace affine normé E, on a une équation aux différentielles totales, dont les propriétés sont très différentes. Le cas où E est le corps des complexes sera étudié au 2ème semestre.

Une telle l'onction f s'appelle une solution ou une intégrale de l'équation différentielle.

Dans le cas particulier où F est l'espace $\mathbb{R}^n$, alors la donnée d'une fonction f est équivalente à la donnée d'un système de n fonctions scalaires $f_1, f_2, \dots, f_n$. L'application $\vec{L}$ est équivalente à un système den. fonctions continues scalaires $L_1, L_2, \dots L_n$, des variables $x, y_1, y_2, \dots y_n$, et l'équation différentielle est équivalente au "système différentiel".

$$(\mathrm{V},1;3) \qquad \begin{cases} y_1' = L_1(x, y_1, y_2, \dots, y_n), \\ y_2' = L_2(x, y_1, y_2, \dots, y_n), \\ \dots\dots\dots\dots\dots\dots\dots, \\ y_n' = L_n(x, y_1, y_2, \dots, y_n). \end{cases}$$

Une intégrale ou solution d un tel système différentiel est alors un système den fonctions scalaires $f_1, f_2, \dots f_n$ satisfaisant aux identités :

$$(\mathrm{V},1;4) \qquad f_i'(x) \equiv L_i(x, f_1(x), f_2(x), \dots, f_n(x)) \ ; \qquad i = 1, 2, \dots, n.$$

Les équations différentielles précédentes sont du premier ordre, en ce sens qu'elles font intervenir seulement la dérivée d'ordre 1. Cn peut naturellement considérer des équatio'ns d'ordre supérieur, mais elles peuvent se ramener aux équations du premier ordre.

Considérons par exemple l'équation différentielle suivante :

$$(\mathrm{V},1;4\,bis) \qquad \vec{y}^{(p)} = L(x, y, \vec{y}', \vec{y}'', \dots, \vec{y}^{(p-1)}).$$

Ici, $]a,b[$ est un intervalle de la droite réelle, F un espace affine normé, $\mathcal{U}$ un ouvert de $F \times \vec{F}^{p-1}$ et $\vec{L}$ une application continue de $]a,b[\times \mathcal{U}$ dans $\vec{F}$. On cherche une fonction f, définie sur $]a,b[$, à valeurs dans F, p fois dérivable, telle que, pour tout x de $]a,b[$, le point $(f(x), \vec{f}'(x), \ldots, \vec{f}^{(p-1)}(x))$ soit dans $\mathcal{U}$, et que l'on ait l'identité :

$$(V,1;5) \qquad \vec{f}^{(p)}(x) \equiv L(x, f(x), \vec{f}'(x), \ldots, \vec{f}^{(p-1)}(x)) .$$

Posons alors :

$$(V,1;6) \qquad f = g_0 \ , \quad \vec{f}' = \vec{g}_1 \ , \ldots , \ \vec{f}^{(p-1)} = \vec{g}_{p-1}$$

On voit que la recherche de f est exactement équivalente à la recherche du système de fonctions $g_0, \vec{g}_1, \vec{g}_2, \cdots, \vec{g}_{p-1}$, satisfaisant au système différentiel :

$$(V,1;7) \qquad \begin{cases} \vec{g}_0' = \vec{g}_1 \ , \\ \vec{g}_1' = \vec{g}_2 \ , \\ \ldots\ldots\ldots , \\ \vec{g}_{p-2}' = \vec{g}_{p-1} \ , \\ \vec{g}_{p-1}' = L(x, g_0, \vec{g}_1, \ldots, \vec{g}_{p-1}) . \end{cases}$$

Or ce système est tout simplement une équation différentielle, analogue a celle que nous avions considérée jusqu'à présent, dans laquelle l'espace F est remplacé par le produit $F \times \vec{F}^{p-1}$, et l'ouvert Ω par $\mathcal{U}$; si on pose $(y_0, y_1, y_2, \cdots y_{p-1}) = z$, nous avons simplement à résoudre l'équation différentielle :

$$(V,1;8) \qquad \vec{z}' = \vec{\mathcal{L}}(x,z) \text{ , avec } z = g(x) = (g_0(x), \vec{g}_1(x), \ldots, \vec{g}_{p-1}(x)),$$

où $\mathcal{L}$ est l'application $(x, (y_0, \vec{y}_1, \ldots, \vec{y}_{p-1})) \longrightarrow$

$(\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_{p-1}, L(x, y_0, \vec{y}_1, \ldots, \vec{y}_{p-1}))$ de $]a,b[\times \mathcal{U}$ dans $\vec{F}^{p}$.

Nous venons toujours de supposer que l'équation différentielle considérée était <u>régulière</u>, c'est-a-dire résolue par rapport à la dérivée de l'ordre le plus élevé. Il n'en est pas nécessairement ainsi dans la pratique. Il s'introduit alors de nombreuses difficultés, et en particulier l'absence de solution ou l existence de plusieurs solutions, correspondant à des conditions initiales données. Les théorèmes théoriques que nous donnerons dans la suite, supposerons toujours l'équation résolue par rapport à la dérivée de 1'ordre le plus élevé, donc finalement toujours posée la forme (V,1;1).

On appelle condition de Cauchy pour une solution d'une telle équation, la donnée de la valeur y_0 de la solution f, en un point x_0 de l'intervalle $]a, b[$; on dit, en abrégé, "la condition de Cauchy x_0, y_0 ". Si on a à résoudre une équation différentielle d'ordre? du type (V,1;4), compte tenu de ce qu'elle se ramène à la forme (V,1;1) comme nous venons de l indiquer, une condition de Cauchy revient à la donnée d'une valeur $z_0 = (y_0, y_1, \ldots, y_{p-1})$ de la fonction $z = g(x)$, pour $x = x_0$, c'est-à-dire des valeurs de la fonction f et de ses dérivées d'ordre 1, 2, ... $p-1$, au point x_0 de l'intervalle $]a, b[$.

Nous allons démontrer que, moyennant certaines conditions relatives à la fonction $\vec{L}$, l'équation différentielle admet une solution et une seule, correspondant à une condition de Cauchy donnée.

§§ 2 THÉORÈMES D'EXISTENCE ET D'UNICITÉ

Donnons d'abord quelques définitions :

On dit qu'un intervalle borné $J =]x_0-\alpha, x_0+\beta[$, $\alpha, \beta > 0$, contenu dans $]a, b[$, et une boule fermée B de centre y_0 et de rayon R (fini ou infini, > 0) contenue dans Ω, constituent un intervalle et une boule de sécurité pour $\vec{L}$, relativement à $x_0 \in]a, b[$, $y_0 \in \Omega$, s'il existe un nombre M, fini ou infini, ≥ 0 , tel que, d'une part, $\|\vec{L}(x, y)\|$ soit majoré par M dans le produit $J \times B$, et que, d'autre part, on ait les inégalités $\alpha \leq \frac{R}{M}$, $\beta \leq \frac{R}{M}$ (*). Un système d'un intervalle et d'une boule de sécurité existe toujours. En effet, la fonction $\vec{L}$ étant supposée continue, il est d'abord possible de choisir un intervalle J_1 : $]x_0 - \alpha_1, x_0 + \beta_1[$, et une boule6 de centre y_0 et de rayon R, tels que $\vec{L}$ soit bornée dans $J_1 \times B$.

* Voir ce renvoi page 745.

Soit M la borne de $\|\vec{L}(x,y)\|$ dans $J_1 \times B$. Prenons alors l'intervalle $J : |x_0-\alpha, x_0+\beta|$, α et β étant définis par : $\alpha = \text{Inf.}(\alpha_1, \frac{R}{M})$, $\beta = \text{Inf.}(\beta_1, \frac{R}{M})$. Alors J et B répondent à la question.

On peut naturellement, si c'est nécessaire pour "ne raison quelconque, remplacer le nombre R par un nombre plus petit R'. Mais alors, R' une fois choisi, α' et β' se déterminent de nouveau à partir de R'. Si donc J, B, est un système de sécurité, et si $J'\subset J$, $B'\subset B'$, J', B', n'est pas nécessairement un système de sécurité; mais J', B, en est un.

Nous dirons d'autre part que $\vec{L}$ est <u>localement lipschitzienne en y dans</u> $|a,b| \times \Omega$, si, quels que soient x_0 de $|a,b|$ et y, de Ω, il existe des voisinages $\mathcal{A}$ et $\mathcal{B}$ de ces points, et un nombre $k \geq 0$, tels que l'on ait, pour

$$x \in \mathcal{A}, \quad y_1 \in \mathcal{B}, \quad y_2 \in \mathcal{B},$$

$$\|\vec{L}(x,y_1) - \vec{L}(x,y_2)\| \leq k\,\|\vec{y}_1 - \vec{y}_2\|. \tag{V,2;1}$$

étant donnés x_0 et y_0, on peut déterminer l'intervalle de sécurité J et la boule de sécurité B, comme il est dit antérieurement, mais de façon qu'en outre $\vec{L}$ soit lipschitzienne en y dans $J \times B$, c'est-à-dire vérifie (V,2;1) pour $x \in J$, $y_1 \in B$, $y_2 \in B$.

* Renv~oi de la page précédente.

Exceptionnellement dans ce chapitre, nous nous permettrons éventuellement de prendre $R = +\infty$, de sorte qu'alors $B = F$, tout en continuant à appeler B "ne boule de centre y_0. Dans ce cas, on peut prendre $M = +\infty$; alors α et β sont quelconques, mais finis (J doit être borné). Dans le cas exceptionnel où x_0 est une extrémité inférieure a (resp. supérieure b) de l'intervalle $|a,b|$, un Intervalle de sécurité sera de la forme $[a, a+\beta|$, $\beta > 0$ (resp. $|b-\alpha, b]$, $\alpha > 0$). <u>L'important, c'est qu'un tel intervalle soit un voisinage de x_0 dans $|a,b|$.</u> Nous ne reparlerons plus de ces cas exceptionnels; le lecteur fera de lui-même les adaptations nécessaires.

Théorème 1 (Cauchy) - Supposons donnée une équation différentielle (V,1;1), dans laquelle $\vec{L}$, est continue de $]a,b[\times \Omega$ dans $\vec{F}$ et localement lipschitzienne en y et $\vec{F}$ est complet. Alors, étant donné une condition de Cauchy, x_0, y_0, et un système d'un intervalle de sécurité J et d'une boule de sécurité B relatifs à x_0, y_0 tels que $\vec{L}$ soit lipschitzienne en y dans $J \times B$, l'équation différentielle admet une solution et une seule f, satisfaisant à la condition de Cauchy donnée $f(x_0) = y_0$, définie dans J et telle que $f(J) \subset B$.

Démonstration - Nous remplacerons l'équation différentielle, avec la condition initiale, par une équation intégrale équivalente. Si en effet f est une fonction dérivable, qui est solution de l équation différentielle et satisfait à la condition initiale $f(x_0) = y_0$, alors, comme f et $\vec{L}$ sont continues, $x \longrightarrow \vec{L}(x, f(x))$ est continue d'après le théorème des fonctions composées, et le corollaire 1 du théorème 89 du chapitre II dit que f, primitive de cette fonction, en est une intégrale indéfinie, de sorte que f est une fonction continue, qui est solution de l'équation intégrale :

$$f(x) = y_0 + \int_{x_0}^{x} \vec{L}(\xi, f(\xi))\, d\xi \,. \tag{V,2;2}$$

Réciproquement, si f est une fonction continue, solution de cette équation intégrale, le même théorème 89 du chapitre II dit que le 2ème membre est dérivable et que sa fonction dérivée est $x \longrightarrow \vec{L}(x, f(x))$, donc f est solution de l'équation différentielle; en outre on a bien $f(x_0) = y_0$.

Appelons E l'espace métrique $(B^J)_{cb}$ des applications continues bornées de J dans B. Soit f un élément de E. A partir de f, construisons la fonction g, définie sur J à valeurs dans F, par la formule :

$$g(x) = y_0 + \int_{x_0}^{x} \vec{L}(\xi, f(\xi))\, d\xi \,. \tag{V,2;3}$$

D'après le choix de l'intervalle et de la boule de sécurité, cette fonction prend en réalité ses valeurs non pas dans F , mais dans la boule de sécurité $\mathcal{B}$.

En effet, elle vérifie la majoration :

(V,2;4) $$\|\overrightarrow{g(x) - y_0}\| \leqslant |x-x_0| \operatorname{Sup}_{\xi\in J} \|\vec{L}(\xi, f(\xi))\| \leqslant$$
$$\leqslant |x-x_0| \operatorname{Sup}_{\substack{\xi\in J\\ \eta\in\mathcal{B}}} \|\vec{L}(\xi,\eta)\|$$
$$\leqslant \frac{R}{M} M = R .$$

Ainsi nous faisons correspondre à tout élément f de E un autre élément g de E , et par conséquent nous définissons une application Φ de E dans E , en posant $g = \Phi(f)$.

L'équation intégrale que nous voulons résoudre est alors équivalente à l'équation :

(V,2;5) $$f = \Phi(f) .$$

Pour montrer que cette équation admet une solution et une seule, nous allons tout simplement montrer que nous sommes dans les conditions d'application du théorème du point fixe (Théorème 46 du chapitre II).

Tout d'abord E est un espace métrique complet. En effet $\mathcal{B}$ est fermée dans F supposé complet, donc elle est complète d'après le théorème 43 du chapitre II, et alors $(\mathcal{B}^J)_{c\ell}$ est complet d'après le corollaire 2 du théorème 65 du chapitre II.

Cherchons maintenant si Φ est une contraction dé E .

Pour $u \in E$, $v \in E$ on a :

(V,2;6) $$\|\vec{L}(\xi, u(\xi)) - \vec{L}(\xi, v(\xi))\| \leqslant k \|\overrightarrow{u(\xi) - v(\xi)}\|$$
$$\leqslant k\, d(u, v) , \text{ pour } \xi \in J ;$$

(V,2;7) $$d(\Phi(u), \Phi(v)) \leqslant \operatorname{Sup.}(\alpha, \beta)\, k\, d(u,v)$$

puisque $|x - x_0| \leq \text{Sup.}(\alpha, \beta)$. C'est donc bien une relation de Lipschitz, mais dans laquelle le coefficient de Lipschitz n'est pas nécessairement < 1 ; Φ n'est pas nécessairement une contraction. Cependant, la boule B étant choisie, il nous est toujours possible de remplacer l'intervalle J par un intervalle plus petit J_0 , en remplaçant α ou β par des nombres plus petits α_0 , β_0 avec Max $(\alpha_0, \beta_0) < \frac{1}{k}$, de manière que Φ soit une contraction * . Mais ce n'est pas nécessaire et nous allons voir que, quel que soit le choix qui a été fait de J, B , <u>il existe toujours une itérée de</u> Φ qui est une contraction. Partons en effet de deux éléments quelconques u et v de E , et construisons leurs images par les itérées successives de Φ.

On a : $u_0 = u$, $u_1 = \Phi(u_0)$,..., $u_n = \Phi(u_{n-1})$,... , et $v_0 = v$, $v_1 = \Phi(v_0)$,..., $v_n = \Phi(v_{n-1})$,...; alors on a les majorations :

(V,2;8) $$\| \overrightarrow{u_1(x) - v_1(x)} \| \leq |x - x_0| \, k \, d(u_0, v_0) .$$

* On pourrait parfaitement se contenter de cela, en modifiant un peu l'énoncé : au lieu de dire que, <u>quel que soit</u> le système de sécurité J, B , tel que L soit lipschitzienne dans $J \times B$ il existe une solution et une seule, il suffisait d'énoncer ; <u>il existe</u> un système de sécurité $J_0 \times B$ tel qu'il existe une solution et une seule......

S'il ne s'agit que de la théorie des équations différentielles,un tel énoncé modifié est entièrement satisfaisant, car, de toute manière, le théorème 2 permettra de récupérer la solution dans tout son domaine d'existence.

Mais dans des théories voisines (équations intégrales), où il est indispensable d'avoir la solution globale d'un seul coup et non un morceau de solution, on ne peut pas se permettre de restreindre ainsi l'énoncé. D'autre part, il est toujours utile d'avoir un intervalle J <u>aussi grand que possible</u> dans lequel le procédé des approximations <u>successives</u> donne la solution.

On a alors ensuite :

(V,2;8) $$\overrightarrow{u_2(x)-v_2(x)} = \int_{x_0}^{x} \left(\vec{L}(\xi, u_1(\xi)) - \vec{L}(\xi, v_1(\xi))\right) d\xi \text{ , donc}$$

(V,2;9) $$\left\| \overrightarrow{u_2(x)-v_2(x)} \right\| \leq \int_{|x_0,x|} k \left\| \overrightarrow{u_1(\xi)-v_1(\xi)} \right\| d\xi$$

$$\leq k^2 d(u_0, v_0) \int_{|x_0,x|} |\xi - x_0| \, d\xi = \frac{(x-x_0)^2}{2} k^2 d(u_0, v_0).$$

On devine facilement la formule générale.

Supposons que l'on ait démontré la majoration :

(V,2;10) $$\left\| \overrightarrow{u_{n-1}(x)-v_{n-1}(x)} \right\| \leq \frac{|x-x_0|^{n-1}}{(n-1)!} k^{n-1} d(u_0, v_0) .$$

Montrons que la même majoration est valable pour $u_n - v_n$. On a en effet :

(V,2;11) $$\overrightarrow{u_n(x)-v_n(x)} = \int_{x_0}^{x} \left(\vec{L}(\xi, u_{n-1}(\xi)) - \vec{L}(\xi, v_{n-1}(\xi))\right) d\xi \text{ , d'où}$$

$$\left\| \overrightarrow{u_n(x)-v_n(x)} \right\| \leq \int_{|x_0,x|} k \left\| \overrightarrow{u_{n-1}(\xi)-v_{n-1}(\xi)} \right\| d\xi$$

$$\leq k^n d(u_0, v_0) \int_{|x_0,x|} \frac{|\xi - x_0|^{n-1}}{(n-1)!} d\xi = \frac{|x-x_0|^n}{n!} k^n d(u_0, v_0).$$

Ainsi l'itérée Φ^n de Φ vérifie

(V,2;12) $$d\left(\Phi^n(u), \Phi^n(v)\right) \leq \frac{(\operatorname{Sup}(\alpha, \beta))^n}{n!} k^n d(u, v)$$

Or, pour n assez grand, on a nécessairement : $\frac{(\operatorname{Sup}(\alpha,\beta))^n}{n!} k^n < 1$, ce qui prouve bien que l'itérée correspondante Φ^n de Φ est une contraction de E . Il résulte alors de la remarque qui suit le théorème 46 du chapitre II que la méthode des approximations successives est valable, et qu'il existe un point fixe, et un seul, pour Φ , ce qui démontre l existence et l unicité de la solution de l'équation différentielle. Les majorations que nous venons de faire montrent en outre que, si l'on détermine la solution-) par la méthode des approximations successives, on a la représentation de la solution par une série :

$$(V,2;13)\quad \begin{cases} f = f_0 + (f_1 - f_0) + (f_2 - f_1) + \cdots + (f_n - f_{n-1}) + \cdots \\ f_n(x) = y_0 + \int_{x_0}^{x} \vec{L}(\xi, f_{n-1}(\xi))\, d\xi , \end{cases}$$

avec la majoration :

$$(V,2;14)\quad \left\| \overrightarrow{f_n - f_{n-1}} \right\| \leqslant \frac{(\operatorname{Sup}(\alpha,\beta))^{n-1}}{(n-1)!} k^{n-1} \left\| \overrightarrow{f_1 - f_0} \right\|$$

terme général d une série très rapidement convergente, comme celle qui représente la fonction exponentielle.

Remarques 1°/ Il est possible de montrer, si F est de dimension finie, que, sous la seule hypothèse de continuité de $\vec{L}$ sans condition de Lipschitz, l'équation différentielle possède toujours au moins une solution locale satisfaisant à la condition de Cauchy donnée. Autrement dit, il y a toujours existence, et c'est seulement le résultat de l'unicité qui est en défaut. Si F est de dimension infinie, et si $\vec{L}$ est seulement continue, il n y a peut-être ni existence ni unicité de la solution. Un exemple très simple montre, si $\vec{L}$ ne vérifie pas de condition de Lipschitz, que l équation différentielle peut avoir plusieurs solutions pour la même condition de Cauchy. Considérons par exemple l'équation différentielle scalaire ($F = \mathbb{R}$) :

$$(V,2;15)\quad y' = 3y^{\frac{2}{3}}$$

Il est clair qu'il n'existe pas de condition de Lipschitz au voisinage de $y = 0$, car $\dfrac{|y^{\frac{2}{3}} - 0|}{|y - 0|} = \dfrac{1}{|y|^{1/3}}$ ne reste pas borné quand $y \neq 0$ tend vers 0.

La "solution générale" de l'équation est

$$(V,2;16)\quad y = (x - c)^3 \quad , c \text{ constante; et } y = 0 .$$

Si nous cherchons les solutions de l'équation différentielle, correspondant a la condition initiale $y = 0$ pour $x = 0$, on voit qu'il en existe une infinité, et en particulier les quatre suivantes :

$$(V,2;16) \quad a)\ y = 0\ ; \quad b)\ y = \begin{cases} 0 \text{ pour } x \leq 0 \\ x^3 \text{ pour } x > 0 \end{cases}\ ; \quad c)\ y = \begin{cases} x^3 \text{ pour } x < 0 \\ 0 \text{ pour } x \geq 0 \end{cases}\ ; \quad d)\ y = x^3.$$

2°/ On peut voir aussi facilement qu'il est indispensable de se limiter à un Intervalle de sécurité J et à une boule de sécurité $\mathbf{B}$. Ce système de sécurité joue un triple rôle :

a) l'intervalle $\mathbf{J}$ doit être borné($|a,b|$ ne l'est pas nécessairement, et peut être $\mathbb{R}$ tout entière), pour la majoration (V,2;12).

b) Il faut avoir un espace $\mathbf{E}$ dont l'application Φ ne fasse pas sortir, de façon à pouvoir appliquer le théorème du point fixe. Mais, s'il se trouve que $\mathbf{L}$ est définie sur $|a,b| \times \mathbf{F}$ (autrement dit, si $\Omega = \mathbf{F}$), on peut, de ce point de vue. prendre $R = M = +\infty$, $E = (F^J)_{cb}$, où $\mathbf{J}$ est n'importe quel intervalle borné de $|a,b|$.

c) La condition de Lipschitz est supposée seulement locale, et on est obligé de choisir J et $\mathbf{B}$, pour qu'elle soit vérifiée dans $\mathbf{J} \times \mathbf{B}$. Cette nécessité disparaît évidemment, si l'on suppose que $\mathbf{L}$ satisfait à une condition de Lipschitz globale, c'est-à-dire que l'on a la relation (V,2;1), quels que soient x dans $|a,b|$, y_1 et y_2 dans Ω.

Il en résulte donc que, <u>si $\vec{L}$ est continue sur $|a,b| \times \mathbf{F}$ et vérifie une même condition de Lipschitz (V,2;1) dans $|a,b| \times \mathbf{F}$, le système de sécurité n'est plus nécessaire, et la solution de l'équation différentielle existe et est unique sur $|a,b|$ tout entier; la suite des approximations successives converge partout, et uniformement sur toute partie bornee de $|a,b|$.</u>

Nous allons donner un exemple dans lequel? est définie sur $\mathbb{R} \times F$, et partout de classe C^∞, et où cependant la solution de l'équation différentielle n'existe que dans un petit intervalle ayant pour centre le point initial x_0 ; bien entendu, la condition de Lipschitz est seulement vérifiée localement.

Considérons à cet effet, l'équation différentielle scalaire :

(V,2;17) $$y' = - y^2 .$$

La régularité du second membre aurait permis d'espérer a priori une solution définie sur tout l'axe réel ; or , l'équation se résout par les formules :

(V,2;18) $$y = \frac{1}{x-c} \; ; \; c \text{ constante; et } y = 0$$

Pour la constante c donnée, on voit que la solution présente la singularité imprévisible $x = c$. Si donc on a choisi la condition initiale $f(0) = y_0 = -\frac{1}{c}$, $c > 0$, on peut être sûr que l'intervalle de sécurité ne va pas jusqu'au point c . Cherchons ici J et B .

Prenons, par exemple, pour B , la boul; de centre y_0 et de rayon R , on trouve alors $M = \left(\frac{1}{c} + R\right)$, d'où $\beta = \frac{R}{\left(\frac{1}{c}+R\right)^2}$.

Nous avons intérêt à choisir R de façon que β soit aussi grand que possible; mais la quantité trouvée au deuxiime membre de l'égalité précédente est sûrement $< c$, car on a toujours : $R < c\left(\frac{1}{c} + R\right)^2$, inégalité qui est en effet équivalente à : $R < \frac{1}{c} + 2R + cR^2$, vérifiée puisque $R < 2R$.

Extension de la méthode à la résolution de certaines équations intégrales

Soient toujours F un espace affine normé complet, Ω un ouvert de F , $|a,b|$ un intervalle de $\mathbb{R}$, et soit $\vec{L}$ une application continue $:(x,\xi,y) \longrightarrow \vec{L}(x,\xi,y)$ de $|a,b| \times |a,b| \times \Omega$ dans $\vec{F}$.

Soit d'autre part $x \longrightarrow h(x)$ une fonction continue sur $|a,b|$ à valeurs dans? ; et soit x_0 un point de $|a,b|$.

Considérons l'équation intégrale

$$f(x) = h(x) + \int^{x} \vec{L}(x, \xi, f(\xi))\, d\xi \quad , \tag{V,2;19}$$

où f est "ne fonction inconnue, continue sur $|a,b|$, à valeurs dans Ω . On dit que c'est "ne équation intégrale, parce que l'inconnue f figure sous le signe d'intégration. C'est "ne généralisation de l'équation (V,2;2), qui était équivalente à "ne équation différentielle avec conditions Initiales; on a ici $h(x)$ a" lieu de y_0 , et $\vec{L}$ dépend dexet de ξ , autrement dit x aussi figure sous le signe d'intégration; l'équation ne se ramène pas à "ne équation différentielle. Ici on cherche en général "ne solution f définie dans $|a,b|$ tout entier.

Théorème 1 bis - Supposons qu'il existe un point y_0 de F, un nombre (fini ou Infini) $R \geqslant 0$, un nombre M (fini ou Infini $\geqslant 0$, un nombre $a \geqslant 0$, tels que la boule $B = B(y_0 ; R)$ de centre y_0 et de rayon R soit dans Ω , et que :

1°/ pour $x \in |a,b|$, $\xi \bullet |a,b|$, y dans la boule B , on ait :

$$\|\vec{L}(x, \xi, y)\| \leqslant M \ ; \tag{V,2;20}$$

2°/ pour $x \in |a,b|$, $\xi \in |a,b|$, y_1 et y_2 dans la même boule B , on ait la condition de Lipschitz :

$$\|\vec{L}(x,\xi,y_1) - \vec{L}(x,\xi,y_2)\| \leqslant k \|\overrightarrow{y_1 - y_2}\| \ ; \tag{V,2;20bis}$$

3°/ **pour tout** x de $|a, b|$, **on ait** :

(V,2 ;20 ter) $$\left\| \overrightarrow{h(x) - y_0} \right\| + |x - x_0| M \leq R$$

Alors l'équation intégrale (V,2;19) admet une solution et une seule, prenant ses valeurs dans B , et elle est donnée par la méthode des approximations successives, uniformément convergente sur tout intervalle borné $|a', b'|$ de $|a, b|$.

La démonstration est identique à celle du théorème 1, avec l'espace $E = \left(B^{|a', b'|}\right)_{c, b}$, et l'application $\Phi : f \to g = \Phi(f)$, de E dans E , définie par :

(V,2 ;20 quarto) $$g(x) = h(x) + \int_{x_0}^{x} \vec{L}(x, \xi, f(\xi))\, d\xi .$$

Remarque - Supposons qu'on puisse prendre $R = +\infty$, $M = +\infty$, et que, pour tout intervalle $|a', b'|$ borné de $|a, b|$, existe un nombre $k = k(a', b')$, mais non pour $|a, b|$ lui-même. Alors la conclusion subsiste.

Prolongement des solutions locales d'une équation différentielle

L'intervalle J ,choisi d'après les méthodes précédentes, est un intervalle où le théorème du point fixe et la méthode des approximations successives réussissent à coup sûr. Il n'est pas nécessairement le plus grand intervalle où existe une solution de l'équation différentielle.

Théorème 2 - Avec les hypothèses du théorème 1, si, dans un sous-intervalle $|a_1, b_1|$ de $|a, b|$, deux solutions de l'équation différentielle prennent la même valeur en un point C , ces deux solutions coïncident dans tout l'intervalle $|a_1, b_1|$.

Démonstration. - Soit J, B , un système de sécurité relatif à c . Comme f_1 et f_2 sont continues, on peut trouver un intervalle plus petit $J' \subset J$, tel que $f_1(J') \subset B$, $f_2(J') \subset B$. Alors J', B est encore un système de sécurité. Comme dans J', il existe une seule solution ayant en c une valeur donnée et prenant ses valeurs dans B (théorème 1), f_1 et f_2 coïncident dans J'. Donc f_1 et f_2 ne peuvent coïncider en un point de $|a_1, b_1|$, sans coïncider dans tout un voisinage de ce point; l'ensemble $\mathscr{E}$ des x tels que $f_1(x) = f_2(x)$ est donc un ouvert non vide de $|a_1, b_1|$. Mais il est aussi fermé, puisque c'est l'image réciproque de $\{\vec{O}\} \in \vec{F}$ par l'application continue $\overrightarrow{f_1 - f_2}$ de $|a_1, b_1|$ dans $\vec{F}$. Comme $|a_1, b_1|$ est connexe. on a nécessairement $\mathscr{E} = |a_1, b_1|$, ce qui démontre le théorème.

<u>Corollaire - Etant donné l'équation différentielle (V,1;1) vérifiant les conditions du théorème 1, et la condition de Cauchy x_0, y_0, il existe un intervalle maximum</u> (a, , b_0 |, $a \leqslant a_0 \leqslant x_0 \leqslant b_0 \leqslant b$, <u>dans lequel existe une solution de l'équation différentielle, satisfaisant à la condition de Cauchy x_0, y_0 Cette solution est unique dans cet intervalle; elle n'est pas prolongeable jusqu'au point</u> a, , <u>sauf peut-être si</u> $a_0 = a$; <u>elle n est pas prolongeable jusqu'au point</u> b_0, <u>sauf peut-être si</u> $b_0 = b$ *.

<u>Démonstration</u> - En effet considérons tous les points γ tels qu'il existe une solution de l'équation différentielle correspondant à la condition de Cauchy x_0, y_0, dans l'intervalle $[x_0, \gamma]$. D'après le théorème, si γ_1 et γ_2 sont deux de ces points, $\gamma_1 \geqslant \gamma_2$, alors la solution définie dans $[x_0, \gamma_1]$ est un prolongement de la solution définie dans $[x_0, \gamma_2]$. L'ensemble de tous ces γ a une borne supérieure b_0 dans $\overline{\mathbb{R}}$. Il existe alors une solution dans l'intervalle $[x_0, b_0[$.

* Autrement dit, $]a_0, b_0[$ est un intervalle ouvert de $\overline{\mathbb{R}}$, sauf peut-être si $a_0 = a$ ou $b_0 = b$.

Montrons que cette solution n'est pas prolongeable jusqu'à b_0, si $b_0 < b$, autrement dit que $f(x)$ n'a pas de limite quand $x < b_0$ tend vers b_0. Si en effet il en a une, appelons-la $f(b_0)$. Alors $f'(x)$ a la limite $\vec{L}(b_0, f(b_0))$. Le théorème 14 du chapitre III montre alors que f a en b_0 une dérivée à gauche, égale à $\vec{L}(b_0, f(b_0))$. Mais alors il existe, dans un voisinage de b_0, une solution de l'équation, correspondant à la condition de Cauchy $b_0, f(b_0)$; comme f vérifie l'équation, à gauche de b_0, l'unicité nous affirme que cette solution est f elle-même, à gauche de b_0 ; donc f sera prolongeable en une solution à droite de b_0, ce qui contredit la propriété de borne supérieure de b_0.

On raisonne de même pour trouver l'intervalle maximum $]a_0, x_0]$, et montrer que f n'est pas prolongeable jusqu'à a_0, si $a_0 \neq a$.

Si nous reprenons l'exemple donné plus haut à la formule (V,2;17), on voit que, quand on part de la condition initiale : $x_0 = 0$, $y_0 = -\frac{1}{c}$, l'intervalle maximum d'existence de la solution, est l'intervalle $]-\infty, c[$. Cette solution ne peut pas se prolonger jusqu'au point c lui-même.

Pour avoir des solutions globales, définies sur $|a, b|$ lui-même on devra faire sur $\vec{L}$ des hypothèses plus restrictives. C'est ce que nous avons vu à la remarque 2° page 11 ($\vec{L}$ est définie sur $|a,b| \times F$, et satisfait à une même condition de Lipschitz globale en y sur $|a,b| \times F$). Nous verrons un cas plus général au théorème 4, en utilisant d'abord une majoration a priori des solutions (théorème 3).

Majoration a priori des solutions d'une équation différentielle

Théorème 3 - Considérons une équation différentielle scalaire.

(V,2;21) $$z' = M(x, z),$$

M réelle ≥ 0, définie et continue sur $[x_0, b_0] \times \mathbb{R}_+^*$

* $\mathbb{R}_+$ est toujours l'ensemble des nombres réels ≥ 0.

Soit f une i-onction dérivable sur $[x_0, b_0]$, à valeurs dans l'espace affine normé F, vérifiant l'inéquation différentielle stricte.

(V,2;22) $$\|\overrightarrow{f'(x)}\| < M(x, \|\overrightarrow{f(x)-O}\|), \text{ou } \|\vec{y'}\| < M(x, \|\overrightarrow{y-O}\|),$$

O étant une origine choisie dans F. Si alors on a $f(x_0)=y_0$, et si g est une solution $\geqslant 0$ de (V,2;21), correspondant aux conditions initiales $x_0, \|\overrightarrow{y_0 - O}\|$, définie dans $[x_0, b_0]$, alors on a, pour tout x de $[x_0, b_0]$, la majoration :

(V,2;23) $$\|\overrightarrow{f(x)-O}\| \leqslant g(x)$$

avec $<$ dès que $x > x_0$.

Si on remplace l'intervalle $[x_0, b_0]$ par un intervalle $[a_0, x_0]$, on doit remplacer g par une solution $h \geqslant 0$ de l'équation $Z' = -M(x, z)$, correspondant aux mêmes conditions initiales $x_0, \|\overrightarrow{y_0 - O}\|$

Avant de donner la démonstration, disons pourquoi ce théorème s'appelle "majoration a priori des solutions d'une équation différentielle". Soit une équation différentielle (V,1;1), avec $\Omega = F$, et supposons que $\vec{L}$ admette la majoration :

(V,2;23bis) $$\|\vec{L}(x,y)\| < M(x, \|\overrightarrow{y-O}\|)$$

Alors, s'il existe une solution f de (V,1;1), définie dans $[x_0, b_0]$, correspondant à la condition initiale x_0, y_0 (nous n'affirmons pas qu'il en existe, nous disons : s'il en existe). elle admet sûrement, a priori, la majoration (V,2;23). En majorant le 2ème membre $\vec{L}$ d'une équation différentielle, on majore les solutions.

Démonstration - Soit $x > x_0$. La démonstration est très analogue à celle du lemme de la formule des accroissements finis (théorème 13 du chapitre III).

Naturellement g est croissante. Appelons A l'ensemble des points ξ de l'intervalle $[x_0, x]$ où l'on a l'inégalité (V,2;23) (avec ξ à la place de x). Cet ensemble n'est pas vicie, car $x_0 \in A$. Soit c sa borne supérieure. A est fermé, car une limite de points ξ vérifiant l'inégalité large (V,2;23) la vérifie encore. Donc A contient c .

Montrons que l'hypothèse $c < x$ est contradictoire. Si en effet il en est ainsi, on doit avoir, pour $\xi > c$, l'inégalité $\|\overline{f(\xi) - O}\| > g(\xi)$; en faisant tendre $\xi > c$ vers c , on en déduit $\|\overrightarrow{f(c)-O}\| \geq g(c)$, et comme on a déjà $\leq$ puisque $c \in A$, on aura nécessairement

$$\|\overrightarrow{f(c)-O}\| = g(c) . \tag{V,2;24}$$

Dans ce cas, on a :

$$\|\vec{f}'(c)\| < M\left(c, \|\overrightarrow{f(c) - O}\|\right) \tag{V,2;25}$$

$$= M(c, g(c)) = g'(c)$$

Ainsi $\|\vec{f}'(c)\| < g'(c)$; si λ est un nombre strictement compris entre les deux, et si γ est réel, assez petit en module, on a :

$$\begin{cases} \|\overrightarrow{f(c+\gamma) - f(c)}\| < \lambda\gamma , \\ |g(c+\gamma) - g(c)| > \lambda\gamma , \end{cases} \tag{V,2;26}$$

donc

$$\|\overrightarrow{f(c+\gamma) - f(c)}\| < |g(c+\gamma) - g(c)| . \tag{V,2;27}$$

En combinant (V,2;24) et (V,2;27), on a, pour $\gamma > 0$:

$$\|\overrightarrow{f(c+\gamma) - O}\| \leq \|\overline{f(c) - O}\| + \|\overrightarrow{f(c+\gamma) - f(c)}\| \tag{V,2;28}$$

$< g(c)+(g(c+\gamma)-g(c)) = g(c+\gamma)$; ce qui contredit la propriété de borne supérieure de c. On a donc nécessairement $c = x$, ce qui démontre (V,2;23) avec l'inégalité large $\leq$.

Reste à montrer l'inégalité stricte < dès que $x > x_0$. Supposons qu'en un point $c > x_0$, on ait l'égalité, c'est-a-dire (V,2;24). Nous prendrons cette fois $\gamma < 0$.

En combinant encore (V,2;24) et (V,2;30), on a :

$$\text{(V,2;32)} \quad \|\overrightarrow{f(c-\gamma)-O}\| \geq \|\overrightarrow{f(c)-O}\| - \|\overrightarrow{f(c-\gamma)-f(c)}\|$$

$$> g(c)-(g(c)-g(c-\gamma)) = g(c-\gamma) .$$

Ainsi, pour $\gamma > 0$, on aurait $\|\overrightarrow{f(c-\gamma)-O}\| > g(c-\gamma)$, ce qui contredit l'inégalité large (V,2;23) déja obtenue; donc (V,2;23) est bien une inégalité stricte dès que $x > x_0$. Ainsi le théorème est démontré pour l'intervalle $[x_0, b_0]$, $b_0 > x_0$, et se démontre de manière analogue pour $[a_0, x_0]$.

Remarque 1°/ Si on suppose une inégalité large (V,2;22) avec $\leq$, il n'est pas certain qu'on ait (V,2;24) avec $\leq$, sans hypothèsessupplémentaires . Prenons par exemple l'équation (V,2;15) pour M , avec $F = \mathbb{R}$, O étant le nombre 0 . On en déduirait que, si f et g sont 2 solutions de cette équation. correspondant à la même condition initiale 0 , 0 , on a nécessairement $f \leq g$ pour $x \geq 0$, donc aussi $g \leq f$ et par suite $f = g$, autrement dit on en déduirait l'unicité de la solution pour (V,2;15) , qui n'est pas réalisé comme nous l'avons vu.

2°/ Ce théorème redonne évidemment le théorème des **accroissements** finis (théorème 13 du chapitre III), et le lemme qui sert à le démontrer. Prenons, par exemple, le lemme, qui est plus général que le thécrème.

La fonction $x \longrightarrow g(x)-g(0)+\varepsilon x + \varepsilon$ de ce lemme est solution de l'équation différentielle $z' = g'(x)+\varepsilon$, et on a $\| \overrightarrow{f'(x)} \| < g'(x) + \varepsilon$. L'inégalité (V,2;23) donne donc $\| \overrightarrow{f(1) - f(0)} \| < g(1) - g(0) + 2\varepsilon$, d'où (III,5;1) puisque ε est arbitraire.

Toutefois nous avons démontre le lemme et le théorème 13 du chapitre III, sans supposer l'existence de la dérivée de f aux extrémités de l'intervalle. Cr ici nous devons supposer , non seulement que $\vec{f}'(x_0)$ existe, mais qu'il vérifie lui aussi l'inégalité stricte $\|\vec{f}'(x_0)\| < M(x_0, \|\overrightarrow{y_0 - O}\|)$; faute de quoi le même exemple (V,2;15) montrerait aisément que la conclusion du théorème serait en défaut.

<u>Une condition d'existence de solutions globales sur $]a, b[$</u> .

<u>Théorème 4</u> - <u>Considérons l'équation différentielle</u> (V,1;1),<u>et supposons que $\vec{L}$ ait les deux proprietes suivantes</u> :

1°/ $\vec{L}$ <u>est définie sur</u> $]a, b[\times F$ <u>(autrement dit</u> $\Omega = F$), et <u>on a la majoration</u> :

$$\| \vec{L}(x, y) \| \leq \mu \| \overrightarrow{y - O} \| + \nu \quad , \tag{V,2;32}$$

où μ et ν sont des constantes ≥ 0 , O une origine choisie dans F * .

2°/ Quel que soit le nombre $\rho > 0$, il existe un nombre $k=k(\rho)$, tel que, lorsque x varie dans $]a,b[$ et que y_1 et y_2 varient dans la boule $B(O;\rho)$, on ait la condition de Lipschitz (V,2;1).

Alors, si F est complet, pour toute condition initiale x_0 , y_0 , l'équation différentielle admet une solution et une seule, définie sur tout l'intervalle $]a,b[$.

Démonstration - Cherchons en effet à déterminer la longueur d'un intervalle de sécurité possible autour du point x_0 . Choisissons arbitrairement et une fois pour toutes un nombre fini $R > 0$, et prenons pour **B** la boule de centre y_0 et de rayon R. Dans cette boule, la quantité $\|\vec{L}(x,y)\|$ est majorée par $M = \mu(\|\overrightarrow{y_0 - O}\| + R) + \nu$.

L'intervalle de sécurité est donc déterminé par $J=[x_0-\alpha, x_0+\beta]$,

(V,2;32[bis]) $$\operatorname{Sup.}(\alpha,\beta) \leq \frac{R}{\mu(\|\overrightarrow{y_0 - O}\| + R)+\nu} .$$

Par ailleurs, les hypothèses relatives à la condition de Lipschitz prouvent que, dans l'ensemble de sécurité **J**x **B** , $\vec{L}$ est lipschitzienne en y . (La constante4 est celle qui correspondrait au rayon $\rho = \|\overrightarrow{y_0 - O}\| + R$).

* Cette origine est sans importance. Si on la remplace par une autre O' , on a la même inégalité, avec des constantes $\mu' = \mu$, $\nu' = \mu\,\|\overrightarrow{O' - O}\| + \nu$. On peut, si l'on veut, prendre $O = y_0$.

De cela nous tirons le résultat fondamental : <u>Pour toutes les conditions initiales y_0 pour lesquelles $\|\overrightarrow{y_0 - O}\|$ reste borné par un nombre fixe, la longueur de l'intervalle de sécurité reste bornée inférieurement par un nombre > 0 fixe.</u> Partons alors du point x_0, l'intervalle de sécurité nous mène, vers la droite, jusqu'à un point x_1

$x_1 = x_0 + \dfrac{R}{\mu(\|\overrightarrow{y_0 - O}\| + R) + \nu}$. A partir de là, et de la condition initiale $f(x_1)$, nous pouvons recommencer, et un nouvel intervalle de sécurité nous mène jusqu'à un point x_2. Partant de ce point et avec la nouvelle condition initiale $f(x_2)$, un nouvel intervalle de sécurité nous mène jusqu'au point x_3, et ainsi de suite. Ainsi nous avons une suite d'intervalles $[x_0, x_1], [x_1, x_2], [x_2, x_3], \ldots, [x_{n-1}, x_n], \ldots$ et, dès que nous avons formé l'intervalle $[x_{n-1}, x_n]$, nous avons par là même défini une solution de l'équation différentielle dans l'intervalle $[x_0, x_n]$. Nous voulons démontrer qu'en fait, après un nombre fini de construc-tions d'intervalles de sécurité par la méthode précédente, nous atteignons nécessairement tout point b' de $]a, b[$, donné à l'avance, à distance finie.

Il nous suffit, pour le savoir, de montrer que les longueurs des intervalles $[x_{n-1}, x_n]$ sont bornées inférieurement par un nombre > 0 fixe. Or, d'après ce que nous avons vu précédemment, cela sera sûrement vrai, si l'on sait à l'avance que les conditions initiales successives $f(x_0), f(x_1), \ldots, f(x_n), \ldots$ restent dans un ensemble borné, pour $x < b'$. Or, par hypothèse, on a la majoration (V,2;32). On est donc dans les conditions du théorème 3, avec l'équation

$$z' = M_\varepsilon(x, z) = \mu z + \nu + \varepsilon, \quad \varepsilon > 0 . \tag{V,2;33}$$

La solution de cette dernière équation correspondant aux conditions initiales x_0, $\|\overrightarrow{y_0 - O}\|$, s'obtient élémentairement; c'est

$$g_\varepsilon(x) = \left(\|\overrightarrow{y_0 - O}\| + \frac{\nu + \varepsilon}{\mu}\right) e^{\mu(x - x_0)} - \frac{\nu + \varepsilon}{\mu} \tag{V,2;34}$$

On a donc, quel que soit $\varepsilon > 0$, l'inégalité a priori. dans tout intervalle d'origine x_0 où la solution de l'équation considérée existe :

(V,2;35) $$\|\overrightarrow{f(x) - O}\| \leq g_\varepsilon(x) ;$$

ceci étant vrai pour tout $\varepsilon > 0$, on a d'ailleurs aussi :

(V,2;36) $$\|\overrightarrow{f(x) - O}\| \leq \left(\|\overrightarrow{y_0 - O}\| + \frac{\nu}{\mu}\right)e^{\mu(x-x_0)} - \frac{\nu}{\mu} .$$

En particulier, pour tout $x_n < b'$, on a :

(V,2;37) $$\|\overrightarrow{f(x_n) - O}\| \leq \left(\|\overrightarrow{y_0 - O}\| + \frac{\nu}{\mu}\right)e^{\mu(b'-x_0)} - \frac{\nu}{\mu} .$$

Cela prouve bien que les $f(x_n)$ en question restent dans une partie bornée de F, et cela montre bien que les intervalles $[x_n, x_{n+1}]$, pour $x_n < b'$, sont de longueur bornée inférieurement par un nombre > 0 fixe, donc que b' est sûrement atteint pour une valeur finie de n. On en déduit bien que la solution existe dans $[x_0, b']$, pour tout b' fini $\leq b'$, donc dans $[x_0, b|$.

De même elle existe dans $|a, x_0]$, donc dans $|a, b|$.

<u>Remarques</u> 1°/ Si l'on remplace l'inégalité (V,2;19) par une moins bonne, du type

(V,2;38) $$\|\vec{L}(x, y)\| \leq \mu \|\overrightarrow{y - O}\|^{1+\alpha} + \nu , \quad \alpha > 0 ,$$

la conclusion du théorème 4 cesserait d'être vraie. Ainsi l'équation scalaire

(V,2;39) $$y' = -\beta\, Y^{\frac{\beta+1}{\beta}} , \beta > 0 ,$$

satisfait à cette condition, sur l'intervalle $[0, +\infty[$ de $\mathbb{R}$, avec $\alpha = \frac{1}{\beta}$; cependant, avec la condition initiale $0, \frac{1}{c^\beta}$ $(c > 0)$ sa solution est :

$$f(x) = \left(\frac{1}{c-x}\right)^{\beta}, \tag{V,2;40}$$

définie dans $[0, c[$, mais non prolongeable jusqu'au point c lui-même.

2°/ Le théorème 3 est valable dans des conditions un peu plus générales. Supposons, par exemple, $b = +\infty$. Il suffit que, dans tout intervalle $|a, b'|$, b' fini, les conditions du théorème 4 soient satisfaites, pour que la solution existe dans $|a, b'|$, pour tout b', donc aussi dans $|a + \infty[$.

De même, pour a et b finis, si $\vec{L}$ est définie seulement dans un produit $|a, b| \times \Omega$, Ω étant la boule de centre O et de rayon R_0, et si les autres conditions du théorème 4 sont vérifiée:, on peut toujours affirmer que la solution existe dans $|a, b|$ si la majoration a priori (V,2;36) oblige $f(x)$ à rester à une distance $\leq R' < R_0$ de $\mathbf{0}$, c'est-a-dire si, en posant $l = \operatorname{Sup}((b - x_0), (x_0 - a))$, on a

$$\left(\|\overrightarrow{y_0 - O}\| + \frac{\nu}{\mu}\right) e^{\mu l} - \frac{\nu}{\mu} < R_0 . \tag{V,2;41}$$

Application à la mécanique

Considfrons comme exemple typique un problème de dynamique du point matériel.

Un point matériel est supposé être soumis dans l'espace à trois dimensions à une force $\vec{F}$, qui dépend de la position du point matériel, de la vitesse de ce point, et du temps, à savoir : $\vec{F}(t, M, \vec{V})$. L'équation fondamentale de la dynamique $\vec{F} = m\vec{\gamma}$ se traduit alors par l'équation différentielle du second ordre :

$$m \frac{\overrightarrow{d^2 M}}{dt^2} = \vec{F}\left(t, M, \frac{\overrightarrow{dM}}{dt}\right). \tag{V,2;41}$$

En remplaçant M par f ou y, et en posant $\vec{f}' = \vec{g}$ ou $\vec{y}' = \vec{z}$, ceci est ramené au système différentiel du premier ordre •

$$\begin{cases} \vec{y}' = \vec{z} \\ \vec{z}' = \vec{F}(t, y, \vec{z}) \end{cases}$$

Si alors la force $\vec{F}$ est continue, localement lipschitzienne en $(M, \vec{v})$, et satisfait à la majoration suivante :

(V,2;43) $$\|\vec{F}(t, M, \vec{v})\| \leq \alpha \, \|\overrightarrow{M - O}\| + \beta \|\vec{v}\| + \gamma \quad ,$$

(la norme étant par exemple la norme naturelle de l'espace euclidien à trois dimensions). Alors on peut appliquer le théorème précédent, et voir que la trajectoire d'un point. correspondant à une condition initiale $M = M_0, \vec{v} = \vec{v}_0$, pour $t = t_0$, peut se poursuivre jusqu'à $t = +\infty$. Si au contraire la force ne satisfaisait pas à une majoration (V,2;43), il n en serait pas nécessairement ainsi, et il pourrait arriver qu il existe un temps limite $t = t_1$, tel que la trajectoire soit définie dans l'intervalle ouvert $]t_0, t_1[$, et ne puisse pas se prolonger jusqu'à t_1. On se trouverait dans une circonstance analogue à celle qui se produit dans l'équation différentielle (V,2;17).

Continuité de la solution en fonction d'un paramètre

Supposons que l équation différentielle dépende d'un paramètre λ, parcourant un espace topologique A . $\vec{L}$ est alors une fonction supposée continue sur $|a, b| \times \Omega \times \Lambda$. Nous supposerons en outre qu'elle est toujours localement lipschitzienne par rapport à y, c'est-a-dire que, quel que soit le point x_0, y_0, λ de $|a,b| \times \Omega \times \Lambda$, il existe un voisinage $J \times B \times \mathcal{V}$ de ce point et une constante-k , tels que l'on ait, pour tout x de J , tout y_1 et tout y_2 de B , et tout λ de $\mathcal{V}$, l'inégalité de Lipschitz :

(V,2;44) $$\|\vec{L}(x, y_1, \lambda) - \vec{L}(x, y_2, \lambda)\| \leq k \, \|\overrightarrow{y_1 - y_2}\| \, .$$

* La variable s'appelle t au lieu de x .

On se propose de voir si la solution de l'équation différentielle, correspondant à une condition initiale $y = y_0(\lambda)$ pour $x = x_0(\lambda)$, dépendant elle-même continuement du paramètre λ, est une fonction continue du paramètre λ. Nous désignerons par f_λ la solution de l'équation différentielle correspondant à la condition initiale donnée, pour la valeur λ du paramètre. En réalité f est une fonction de x et de λ, et la quantité f_λ est la fonction partielle correspondante, pour la valeur fixée λ du paramètre.

L'équation différentielle et la condition initiale s'écrivent donc :

$$(V,2;45)\quad \begin{cases} (\vec{f}_\lambda)'(x) = \dfrac{\partial \vec{f}}{\partial x}(x,\lambda) = \vec{L}(x, f(x,\lambda),\lambda) \\ \\ f(x_0(\lambda),\lambda) = y_0(\lambda)\,. \end{cases}$$

Théorème 5 - Si $\vec{L}$ est une fonction continue sur $|a,b| \times \Omega \times \Lambda$, localement lipschitzienne par rapport à y, si $\lambda \longrightarrow x_0(\lambda)$, $\lambda \longrightarrow y_0(\lambda)$ sont des fonctions continues sur Λ à valeurs dans $|a,b|$ et Ω respectivement, et si F est complet, alors, quelle que soit la valeur λ_0 du paramètre, on peut trouver un intervalle $J = |x_0(\lambda_0)-\alpha, x_0(\lambda_0)+\beta|$, $\alpha>0$, $\beta>0$, une boule B de centre $y_0(\lambda_0)$ et de rayon R, et un voisinage $\mathcal{V}$ de λ_0 dans Λ, tels que, pour tout $\lambda \in \mathcal{V}$, il existe une solution et une seule de l'équation différentielle, correspondant à la condition initiale $x_0(\lambda), y_0(\lambda)$, définie dans J et prenant ses valeurs ~~dans~~ B ; et qu'en outre cette solution soit une fonction continue du paramètre; lus précisément, $(x,\lambda) \longrightarrow f(x,\lambda)$ est continue de $J \times \mathcal{V}$ dans B, et l'application partielle $\lambda \longrightarrow f_\lambda$ de $\mathcal{V}$ dans l'espace métrique $(B^J)_{cb}$ est continue.

Démonstration - Nous allons nous ramener au théorème 46 bis du chapitre II, qui donne la continuité, par rapport au paramètre, du point fixe d'une contraction.

Partons d'abord d'un intervalle $J_1 = |x_0(\lambda_0)-\alpha_1, x_0(\lambda_0)+\beta_1|$, d'une boule $B = B(y_0(\lambda_0);\ R\ )$, et d'un voisinage $\mathcal{V}_1$ de λ_0, tels que :

a) L vérifie une condition de Lipschitz (V,2;44), avec un certain coefficient k, dans $J_1 \times B \times \mathcal{V}_1$;

b) $\|\vec{L}(x, y, \lambda)\|$ admette une borne supérieure finie M dans $J_1 \times B \times \mathcal{V}_1$.

Un tel choix est possible, puisque? est continue et localement lipschitzienne en y.

Déterminons alors l'intervalle $J = |x_0(\lambda_0)-\alpha, x_0(\lambda_0)+\beta|$, de manière que $\alpha \leq \alpha_1$, $\beta \leq \beta_1$, et

(V,2;45 bis) $$\alpha + \beta \leq \frac{R}{2M} \quad \text{et} \quad < \frac{1}{k}.$$

Déterminons enfin un voisinage $\mathcal{V}$ de λ_0, $\mathcal{V} \subset \mathcal{V}_1$, de manière que, pour $\lambda \in \mathcal{V}$, on ait :

(V,2;46) $$\begin{cases} x_0(\lambda_0) - \alpha \leq x_0(\lambda) \leq x_0(\lambda_0) + \beta\,, \\ \|\overrightarrow{y_0(\lambda) - y_0(\lambda_0)}\| \leq \dfrac{R}{2}\,; \end{cases}$$

ceci est possible, en vertu de la continuité supposée des conditions initiales par rapport au paramètre.

Prenons alors toujours l'espace $E = (\mathring{B}^J)_{cb}$, qui est complet. Appelons Φ_λ l'application de E dans $(F^J)_{cb}$ définie par :

(V,2;47) $$\Phi_\lambda(f) = g\,, \quad g(x) = y_0(\lambda) + \int_{x_0(\lambda)}^{x} \vec{L}(\xi, f(\xi), \lambda)\, d\xi$$

Montrons d'abord que g prend ses valeurs dans B ; ainsi J, B, constitueront un système de sécurité valable pour tous les λ de $\mathcal{V}$, et Φ_λ sera, pour tout λ de $\mathcal{V}$, une application de E dans lui-même. On a en effet :

(V,2;48) $$\left\|\overrightarrow{g(x)-y_0(\lambda_0)}\right\| \leqslant \left\|\overrightarrow{y_0(\lambda)-y_0(\lambda_0)}\right\| + |x-x_0(\lambda)|\, M$$
$$\leqslant \frac{R}{2} + (\alpha+\beta) M \leqslant \frac{R}{2} + \frac{R}{2} = R \ ,$$

d'après (V,2;45 bis et 46); donc g prend bien ses valeurs dans B , et Φ_λ est toujours une application de E dans E .

Φ_λ est toujours une contraction, avec un coefficient de Lipschitz < 1 indépendant de λ , car on a :

(V,2;49) $$d(\Phi_\lambda(u)\,,\,\Phi_\lambda(v)) = \operatorname*{Sup}_{x\in J}\left\|\int_{x_0(\lambda)}^{x}\left(\vec{L}(\xi,u(\xi),\lambda)-\vec{L}(\xi,v(\xi),\lambda)\right)d\xi\right\|$$
$$\leqslant (\alpha+\beta)\, k\, d(u,v) \ ,$$

et on a $(\alpha+\beta)\,k < 1$ d'après (V,2;45 bis).

Enfin, pour tout f fixé dans E l'application partielle $\lambda \longrightarrow \Phi_\lambda(f)$ est continue sur $\mathcal{V}$.

En effet, si λ tend vers λ_1 , on a :

(V,2;50) $$d(\Phi_\lambda(f)\,,\Phi_{\lambda_1}(f)) \leqslant \left\|\overrightarrow{y_0(\lambda)-y_0(\lambda_1)}\right\|$$
$$+ \operatorname*{Sup}_{x\in J}\left\|\int_{x_0(\lambda)}^{x}\vec{L}(\xi,f(\xi),\lambda)\,d\xi - \int_{x_0(\lambda_1)}^{x}\vec{L}(\xi,f(\xi),\lambda_1)\,d\xi\right\|$$

Or, lorsque λ tend vers λ_1, le premier terme du 2ème membre tend vers 0 , en vertu de la continuité de la condition initiale $y_0(\lambda)$ par rapport à λ .

Et la 2ème expression est majorée par

(V,2;51) $$\left\|\int_{x_0(\lambda)}^{x_0(\lambda_1)}\vec{L}(\xi,f(\xi),\lambda)\,d\xi\right\|$$
$$+ \operatorname*{Sup}_{x\in J}\left\|\int_{x_0(\lambda_1)}^{x}\left(\vec{L}(\xi,f(\xi),\lambda)-\vec{L}(\xi,f(\xi),\lambda_1)\right)d\xi\right\|$$

Le premier terme de (V,2;51) tend vers 0 quand λ tend vers λ_1, car il est majoré par $|x_0(\lambda_1) - x_0(\lambda)|$ M , et on a supposé la continuité de x_0 (A) par rapport à λ. Le 2ème terme de (V,2;51) est majoré par

(V,2;52) $$\int_J \left\| \vec{L}(\xi, f(\xi), \lambda) - \vec{L}(\xi, f(\xi), \lambda_1) \right\| d\xi \ ;$$

Comme $(\xi, f(\xi))$ décrit un compact K de $J \times \Omega$ quand ξ décrit le-compact J , il résulte du théorème 66 du chapitre IV que $\| L(\xi, f(\xi), \lambda) - L(\xi, f(\xi), \lambda_1) \|$ tend vers 0 quand λ tend vers λ_1, uniformément pour $(\xi, f(\xi)) \in K$, c'est-à-dire pour $\xi \in J$, donc (V,2;52) tend vers 0 .

Nous nous trouvons donc dans les conditions d'application du théorème 46 bis du chapitre II. Le point fixe de Φ_λ est la solution f_λ de l'équation différentielle, définie dans J , à valeurs dans B ,et correspondant aux conditions initiales $x_0(\lambda), y_0(\lambda)$. Le théorème 46 bis nous dit que $\lambda \longrightarrow f_\lambda$ est continue de $\mathcal{V}$ dans $E = (B^J)_{cb}$. C'est d'ailleurs équivalent à la continuité de $(x, \lambda) \to f(x, \lambda)$ de J Y $\mathcal{V}$ dans B . On a en effet :

(V,2;53) $$\left\| \overrightarrow{f(x,\lambda) - f(x_1, \lambda_1)} \right\| \leq \left\| \overrightarrow{f(x,\lambda) - f(x, \lambda_1)} \right\| + \left\| \overrightarrow{f(x, \lambda_1) - f(x_1, \lambda_1)} \right\| .$$

$$\leq d(f_\lambda, f_{\lambda_1}) + \left\| \overrightarrow{f_{\lambda_1}(x) - f_{\lambda_1}(x_1)} \right\| ,$$

le premier terme tend vers 0 puisque nous avons démontré la continuité de A- h , et le 2ème parce que f_{λ_1} est continue a" point x_1, donc $(x, \lambda) \longrightarrow f(x, \lambda)$ est continue sur $J \times \mathcal{V}$. (Inversement, de la continuité de $(x, \lambda) \longrightarrow f(x, \lambda)$, on déduirait la continuité de $\lambda \longrightarrow f_\lambda$ d'après le théorème 66 du chapitre IV, puisque J est un compact). Ainsi le théorème 5 est démontré.

Corollaire - Soient $\vec{L}_n$ une suite de fonctions continues sur $]a,b[\times \Omega$, à valeurs dans $\vec{F}$, convergeant, pour n infini, vers une limite $\vec{L}$, localement uniformément; $(x_0)_n$ une suite de points de $]a,b[$ convergeant vers le point x_0 de $]a,b[$, $(y_0)_n$ une suite de points de Ω convergeant vers le point y_0 de Ω.

On suppose que tout point de $]a,b[\times \Omega$ possède un voisinage dans lequel les $\vec{L}_n$ sont lipschitziennes en y, avec une constante de Lipschitz commune indépendante de n. Alors il existe un intervalle J, voisinage de x_0 dans $]a,b[$, une boule B de centre y_0, et un entier p tels que, pour $n \geqslant p$, on ait $(x_0)_n \in J$, $(y_0)_n \in B$, et que l'équation $\vec{y}' = \vec{L}_n(x,y)$ ait une solution et une seule f_n dans J, prenant ses valeurs dans B, correspondant à la condition initiale $(x_0)_n$, $(y_0)_n$, et que lorsque n tend vers l'infini, f_n converge, uniformément dans J, vers une limite f, unique solution de l'equation différentielle $\vec{y}' = \vec{L}(x,y)$, correspondant a la condition initiale x_0, y_0.

Démonstration - Appelons Λ le sous-espace topologique $\{0,1,2,\ldots,n,\ldots;+\infty\}$ de $\mathbb{R}$. Posons $\vec{L}_\infty = \vec{L}$, $(x_0)_\infty = x_0$, $(y_0)_\infty = y_0$.

Alors nous sommes ramenés au théorème 5 (La convergence uniforme locale des $\vec{L}_n$ vers $\vec{L}$ signifie que $(x,y,\lambda) \longrightarrow \vec{L}_\lambda(x,y)$ est continue de $]a,b[\times \Omega \times \Lambda$ dans $\vec{F}$; car, si x_ν converge vers x, y_ν vers y, λ_ν vers λ, on a

$$\|\vec{L}(x_\nu,y_\nu,\lambda_\nu)-\vec{L}(x,y,\lambda)\| \leqslant \|\vec{L}(x_\nu,y_\nu,\lambda_\nu)-\vec{L}(x_\nu,y_\nu,\lambda)\|$$

$+\|\vec{L}(x_\nu,y_\nu,\lambda)-\vec{L}(x,y,\lambda)\|$, qui tend vers 0 pour ν infini), qui donne toutes les conclusions.

Naturellement le théorème précédent est insuffisant, comme le théorème 1 lui-même, car il ne donne la continuité de la solution de l'équation différentielle que dans un intervalle J trop petit. On peut alors perfectionner le théorème sous la forme suivante :

Théorème 6 - Si $\vec{L}$ est "ne application continue de $|a,b| \times \Omega \times \Lambda$ dans $\vec{F}$, si d'autre part on se trouve, pour toute valeur λ du paramètre λ dans les conditions du théorème 3, les constantes $\mu, \nu, k(\rho)$, étant indépendantes de λ, alors, pour toute valeur du paramètre λ, il existe "ne solution et "ne seule de l'équation différentielle correspondant à la condition initiale $y = y_0(\lambda)$ pour $x = x_0(\lambda)$, et définie dans tout l'intervalle $|a,b|$; de plus, si $\lambda \longrightarrow x_0(\lambda)$ et $\lambda \longrightarrow y_0(\lambda)$ sont des applications continues de Λ dans $|a,b|$ et F respectivement, la solution dépend continûment du paramètre λ sur tout intervalle compact $[a',b']$ contenu dans $|a,b|$, autrement dit l'application $\lambda \longrightarrow f_\lambda$ de Λ dans $E = (F^{[a',b']})_{c\ell}$ est continue.

Démonstration - Tout d'abord l'existence de la solution dans tout l intervalle $|a,b|$, quel que soit λ, dépend tout simplement du théorème 3. Soit alors $\varepsilon >$ o donné. En vertu de la continuité de la condition initiale, si λ_0 est un point de Λ, il existe un voisinage $\mathcal{V}_1$ de λ_0 tel que, pour λ dans $\mathcal{V}_1$ les quantités $\|y(\lambda) - 0\|$ et $x_0(\lambda)$ restent bornées. Dans ces conditions la majoration (V,2;36) de la 1 n montre que celle-ci reste bornée dans tout l'intervalle $[a',b']$, indépendamment du paramètre λ lorsque celui-ci parcourt $\mathcal{V}_1$.

Alors $\vec{L}$ vérifie "ne même condition de Lipschitz (V,2;1) dans l'ensemble borné des valeurs de la solution; soit k la constante de Lipschitz correspondante.

On a d'autre part :

(V,2;54)
$$\overrightarrow{f'_\lambda(x) - f'_{\lambda_0}(x)} = \vec{L}(x, f_\lambda(x), \lambda) - \vec{L}(x, f_{\lambda_0}(x), \lambda_0)$$
$$= \left[\vec{L}(x, f_\lambda(x), \lambda) - \vec{L}(x, f_{\lambda_0}(x), \lambda)\right]$$
$$+ \left[\vec{L}(x, f_{\lambda_0}(x), \lambda) - \vec{L}(x, f_{\lambda_0}(x), \lambda_0)\right]$$

Lorsque λ tend vers λ_0 , le 2ème terme tend vers $\vec{0}$, uniformément pour x dans l'intervalle compact $[a', b']$ de $]a, b[$ (théoréme 66 du chapitre IV ; voir démonstration du théorème 5, pour (V,2;52)); on peut donc déterminer un voisinage $\mathcal{V}'$ de λ_0 , $\mathcal{V}' \subset \mathcal{V}_1$, de manière qu'il soit majoré en norme par $\delta > 0$, nombre donné à l'avance, pour $\lambda \in \mathcal{V}'$. pour des raisons qui apparaîtront plus loin, nous prendrons $\delta = \dfrac{\varepsilon}{(1+\frac{1}{k})e^{k(b'-a')}}$.

A cause de la condition de Lipschitz (V,2;1) , on a :

(V,2;55)
$$\left\| \vec{L}(x, f_\lambda(x), \lambda) - \vec{L}(x, f_{\lambda_0}(x), \lambda) \right\| \leq k \left\| \overrightarrow{f_\lambda(x) - f_{\lambda_0}(x)} \right\|$$

Donc la fonction $\vec{Y} = \overrightarrow{f_\lambda - f_{\lambda_0}}$ vérifie <u>l'inéquation</u> différentielle :

(V,2;56)
$$\|\vec{Y}'\| \leq k \|\vec{Y}\| + \delta .$$

La valeur que prend Y au point $x_0(\lambda)$ est

(V,2;57)
$$\overrightarrow{f_\lambda(x_0(\lambda)) - f_{\lambda_0}(x_0(\lambda))}$$
$$= \left(\overrightarrow{y_0(\lambda) - y_0(\lambda_0)}\right) + \left(\overrightarrow{f_{\lambda_0}(x_0(\lambda_0)) - f_{\lambda_0}(x_0(\lambda))}\right).$$

La première parer-thèse tend vers 0 quand λ tend vers λ_0 (à cause de la continuité de la condition initiale $y_0(\lambda)$ par rapport à λ). La 2ème tend vers 0 parce que $x_0(\lambda)$ tend vers $x_0(\lambda_0)$ (a cause de la continuité de $x_0(\lambda)$ par rapport à λ) et que f_{λ_0} est continue au point $x_0(\lambda_0)$.

Donc le 2ème membre de (V,2;57) tend vers $\vec{0}$, et par suite on peut trouver un voisinage $\mathcal{V}''$ de λ_0 , $\mathcal{V}'' \subset \mathcal{V}_1$, de manière qu'il soit majoré en norme par δ , pour $\lambda \in \mathcal{V}''$.

Alors, pour $\lambda \in \mathcal{V} = \mathcal{V}' \cap \mathcal{V}''$, nous pouvons appliquer à $\vec{Y}$ la majoration (V,2;36), puisqu'elle vérifie (V,2;56), avec une valeur initiale majorée en norme par δ pour $x = x_0(\lambda)$.

On a donc , pour $\lambda \in \mathcal{V}$ et $x \in [a', b']$:

(V,2;58)
$$\left\| \overrightarrow{f_\lambda(x) - f_{\lambda_0}(x)} \right\| = \left\| \overrightarrow{Y(x)} \right\|$$

$$\leqslant \left(\delta + \frac{\delta}{k}\right) e^{k|x - x_0(\lambda)|} - \frac{\delta}{k}$$

$$\leqslant \delta\left(1 + \frac{1}{k}\right) e^{k(b'-a')} = \varepsilon \ ,$$

et le théorème est démontré.

<u>Remarque</u> - Nous verrons ultérieurement (théorème 15 bis) un autre théorème global, ainsi qu'un théorème de dérivabilité par rapport au paramètre.

Dérivées d'ordre supérieur de la solution d'une équation différentielle

<u>Théorème 8</u> - Si L <u>est une application de classe C^m de $]a,b[\times \Omega$ dans F , toute solution de l'équation différentielle (V,1;1) est de classe C^{m+1}</u>.

<u>Démonstration</u> - Tout d'abord $\vec{f}$ est nécessairement de classe C^1, dès que $\vec{L}$ est continue, comme nous l'avons dit dès le début. Faisons alors une récurrence sur m, supposons démontré que $\vec{f}$ est de classe C^m dès que $\vec{L}$ est de classe C^{m-1}. Et supposons que $\vec{L}$ soit de classe C^m. Alors en tout cas elle est de classe C^{m-1} et $\vec{f}$ est de classe C^m, d'après l'hypothèse de récurrence. Mais alors, d'après le théorème des fonctions composées (théorème 19 du chapitre III) la fonction
$\vec{f}' : x \longrightarrow \vec{L}(x, \vec{f}(x))$ est de classe C^m ; et cela prouve bien que $\vec{f}$ est de classe C^{m+1}, et le théorème est démontré.

Naturellement les dérivées successives s'obtiennent par application du théorème des fonctions composées: par exemple :

$$\overrightarrow{f''(x)} = \frac{\partial \vec{L}}{\partial x}(x, f(x)) + \frac{\partial L}{\partial y}(x, f(x)) \cdot \overrightarrow{f'(x)} \tag{V,2;70}$$

$$= \frac{\partial \vec{L}}{\partial x}(x, f(x)) + \frac{\partial L}{\partial y}(x, f(x)) \cdot \vec{L}(x, f(x))$$

Ce qu'il y a de remarquable, c'est que, pour le point initial x_0, la seule connaissance de y_0 permet de calculer les dérivées successives, sans avoir besoin de résoudre l'équation différentielle.

Ainsi on aura :

$$\begin{aligned} f(x_0) &= y_0 \\ \vec{f}'(x_0) &= \vec{L}(x_0, y_0) \\ \vec{f}''(x_0) &= \frac{\partial \vec{L}}{\partial x}(x_0, y_0) + \frac{\partial L}{\partial y}(x_0, y_0) \cdot \vec{L}(x_0, y_0), \text{ etc,...} \end{aligned} \tag{V,2;71}$$

Cette remarque est la source de nombreuses méthodes de résolution, valables dans les cas où on peut d'avance affirmer que la solution, non seulement a des dérivées successives de tous les ordres, mais encore est représentée, au voisinage de x_0, par son développement de Taylor (méthode des majorantes de Cauchy, que nous verrons ultérieurement).

Corollaire - Si $\vec{L}$ est de classe C^∞, toute solution de l'équation différentielle (V,1;1) est de classe C^∞.

Intégrales premières d'une équation différentielle

Etant donnée une équation différentielle (V,1;1), on appelle intégrale première une fonction scalaire H sur $]a,b[\times \Omega : x, y \dashrightarrow H(x, y)$, non constante, mais qui devienne une constante lorsqu'on remplace y par n'importe quelle solution de l'équation différentielle : quelle que soit la solution f, H $(x, f(x))$ est une constante.

Supposons, pour fixer les idées, que F soit de dimension finie n et muni d'un référentiel. Soit f la solution correspondant à la condition initiale x_0, y_0. Fixons une fois pour toutes x_0, mais permettons-nous de faire varier y_0 ; alors f est une fonction de x et de y_0. L'équation $y=f(x;y_0)$ peut en général être résolue par rapport à y_0, au moins localement, sous la forme $y_0 = h(x,y)$; en effet, $h(x,y)$ est tout simplement la valeur, au point x_0, de la solution de l'équation différentielle qui, au point x, prend la valeur y. Alors chacune des composantes $H_1, H_2, \ldots, H_n$, de la fonction h ainsi trouvée est évidemment une intégrale première.

En effet, si f est une solution de l'équation différentielle, la valeur y_0, qu elle prend au point x_0, est évidemment une constante, à savoir $f(x_0)$, (on a $h(x, f(x)) = f(x_0)$), et par conséquent aussi chacune des composantes de cette valeur. Par ailleurs, ces intégrales premières sont évidemment n fonctions scalaires indépendantes de $x, y_1, y_2, \ldots, y_n$; en effet, elles peuvent prendre des valeurs arbitrairement données à l avance, puisqu au point x_0 on peut fixer arbitrairement la valeur initiale y_0. Il n'y a d'ailleurs pas d'autre intégrale première indépendante de celles-là, puisque, si l'on connait les valeurs des H_i, on connaît y_0, et la solution de l'équation est connue.

La méthode précédente est seulement locale, elle prouve, au voisinage d un point (c, γ) de $]a,b[\times \Omega$, l'existence de n intégrales premières indépendantes; mais elle ne prouve pas l'existence globale d intégrales premières, c'est-à-dire de fonctions H qui soient définies sur $]a,b[\times\Omega$ tout entier. Cette existence est d'ailleurs tout une autre affaire, et, même dans le cas d équations différentielles très simples, il n'existe pas en général, d'intégrales premières définies dans tout le domaine d existence de la fonction $\vec{F}$. C'est là une question difficile de l'étude des équations différentielles.

Donnons une application particulièrement importante à la mécanique.

L'équation différentielle d'un problème de mécanique est de la forme :

$$(V,2;72) \qquad \vec{q}'' = \vec{F}(t, q, \vec{q}'),$$

où q représente un point de $\mathbb{R}^n$, c'est-à-dire un nombre fini des scalaires $q_1, q_2, \ldots, q_n$. Si l'on pose $\vec{q}' = \vec{r}$, ce système devient un système différentiel du premier ordre :

(V,2;73)
$$\begin{cases} \vec{q}' = \vec{r} \\ \vec{r}' = \vec{F}(t, q, \vec{r}) \end{cases}$$

Une intégrale première est alors une fonction H de t, de q, et de $\vec{r} = \vec{q}'$, qui soit constante sur une trajectoire du problème.

Or on a vu, dans certains cas, une telle intégrale première, c'est l'énergie, somme de 1 énergie potentielle et de l'énergie cinétique.

Les théorèmes élémentaires de la mécanique : théorème du centre de gravité, théorème du moment cinétique, etc.... permettent fréquemment d'obtenir d'autres intégrales premières indépendantes de l'énergie.

L'intérêt des intégrales premières est que la connaissance d'une intégrale première permet de faciliter la recherche de la solution générale de l'équation différentielle.

Supposons, par exemple, que nous ayons à résoudre une équation différentielle d'ordre p par rapport à une fonction réelle, définie sous la forme :

(V,2;74)
$$y^{(p)} = L\left(x, y, y', \ldots, y^{(p-1)}\right).$$

et supposons que nous en possédions une intégrale première $H(x, y, y', \ldots y^{(p-1)})$. Cela signifie que, pour toute solution de l'équation différentielle, on a :

(V,2;75)
$$H(x, y, y', \ldots, y^{(p-1)}) = C;$$

C est une constante. Si alors cette équation (V,2;75) peut être résolue, par rapport a $y^{(p-1)}$, sous la forme

(V,2;76)
$$y^{(p-1)} = g(x, y, y', \ldots, y^{(p-2)}, C),$$

alors (V,2;74) est remplacée par (V,2;76), qui est une équation différentielle d'ordre $p-1$ au lieu de p, mais dépendant d'une constante arbitraire. La connaissance d'une nouvelle intégrale première, indépendante de la précédente, permettrait alors de ramener l'équation à une équation différentielle d'ordre $p-2$, etc..... * .

Equation différentielle définie par un champ de vecteurs

Soit **V** une **variété** de dimension n, de classe C^m, dans un espace affine, E de dimension N sur le corps des **réels.** <u>Un champ de vecteurs sur V</u> est une application $\vec{X} : x \longrightarrow \vec{X}(x)$ de V dans l'espace vectoriel $\vec{E}$, telle que, pour tout x de V, $\vec{X}(x)$ soit dans l'espace vectoriel $T(x;V)$ tangent en x à V. On dit que le champ de vecteurs est de classe C^p $(p \leq m)$ si l'application $\vec{X}$ de V dans $\vec{E}$ est de classe C^p (définition page 321) Ce champ de vecteurs définit une équation différentielle :

$$\frac{dx}{dt} = \vec{X}(x(t)) \quad , \text{ ou } \quad \vec{x}' = \vec{X}(x) \tag{V,2;77}$$

Une solution de cette équation est une fonction $f : t \longrightarrow f(t)$, définie sur un intervalle de l'axe des temps $\mathbb{R}$, à valeurs dans V, telle que, pour tout t, la vitesse $\vec{f}'(t)$ soit le vecteur $\vec{X}(f(t))$ tangent à V au point $f(t)$. Cette équation n'est pas directement du type (V,1;1), même si le champ est continu, car la fonction $\vec{X}$ est définie seulement sur V, et non sur E ou un ouvert de **E** . Par contre, on remarque que X dépend de x, mais non de t. Ce sont de telles équations qu'on résout quand on cherche les lignes asymptotiques, les lignes de courbure, etc..... d'une surface d'un espace affine euclidien à 3 dimensions.

in dit que le champ est localement lipschitzien, si, quel que soit a de **V** , il existe un voisinage $\mathcal{V}$ de a dans E et un nombre $k \geq 0$, tels que, quels que soient x_1 et x_2 dans $\mathcal{V} \cap$ **V** , on ait :

* Les choses sont rarement aussi simples. L'intégrale première H est souvent seulement locale, on ne peut pas nécessairement la résoudre en $y^{(p-1)}$;(V,2;75) n'est pas toujours <u>équivalente</u> à (V,2;74), etc.......

$$(V,2;78) \qquad \|\vec{X}(x_1) - \vec{X}(x_2)\| \leq k \|\overrightarrow{x_1 - x_2}\| \quad *$$

Théorème 8 bis - Soit (V,2;77) une équation différentielle définie par un champ $\vec{X}$ localement lipschitzien sur une variété V de classe C^m de E, $m \geq 2$. Si $t_0 \in \mathbb{R}$, $x_0 \in V$ est une condition de Cauchy, il existe un intervalle ouvert de $\mathbb{R}$: $]t_0-\alpha, t_0+\beta[$, $\alpha > 0$, $\beta > 0$, où existe une solution et une seule de l'équation différentielle, correspondant à cette condition de Cauchy, tel en outre que la solution ne soit pas prolongeable jusqu'à $t_0 - \alpha$ ni jusqu'à $t_0 + \beta$; dans un intervalle contenant t_0, il ne peut pas exister plus d'une solution prenant une valeur donnée en t_0. Si la variété V est compacte, ou si elle est fermée et si on a une majoration

$$(V,2;78^{bis}) \qquad \|X(x)\| \leq \mu \|x - O\| + \nu ,$$

μ et ν constantes ≥ 0, O point fixé de E, on a

$\alpha = \beta = +\infty$. Si le champ?? est de classe C^p ($p \leq m$), la solution est une fonction de classe C^{p+1} .

Démonstration - Soit $\Phi : \mathcal{O} \longrightarrow \Phi(\mathcal{O})$, une carte d'un voisinage de x_0 sur V . On suppose donc que $\mathcal{O}$ est un ouvert de $\mathbb{R}^n$, et que $x_0 \in \Phi(\mathcal{O})$. On peut supposer $\mathcal{O}$ assez petit pour qu'il existe un prolongement Θ de classe C^m de Φ^{-1}, défini dans un ouvert Ω de E (théorème 33 du chapitre III).

On peut alors transporter le champ de vecteurs en un champ de vecteurs sur $\mathcal{O}$; à chaque point u de $\mathcal{O}$, on associera le vecteur $\vec{U}(u)$ dont l'image par $\Phi'(u)$ soit le vecteur $\vec{X}(\Phi(u))$; c'est possible, puisque $\Phi'(u)$ est une bijection de $\mathbb{R}^n$ sur l'espace vectoriel tangent à V au point $\Phi(u)$ (corollaire 1 du théorème 33 quarto du chapitre III). En outre, d'après le corollaire 2 de ce même théorème, on peut écrire :

* Cela suppose donc, a priori, E normé. Mais comme toutes les normes sur E sont équivalentes (théorème 13 du chapitre II), et qu'un champ localement lipschitzien pour une norme l'est pour toute norme équivalente (avec une constante k différente), le fait pour un champ d'être localement lipschitzien est finalement indépendant de toute norme.

$$(V.2\,;79)\quad \begin{cases} \vec{U}(u) = \Theta'(\Phi(u)) \cdot \vec{X}(\Phi(u)) \text{ ou } \vec{U}(\Theta(x)) = \Theta'(x) \cdot \vec{X}(x)\,; \\[2ex] \vec{X}(\Phi(u)) = \Phi'(u) \cdot \vec{U}(u) \end{cases}$$

Montrons que le champ U ainsi défini sur l'ouvert $\mathcal{O}$ de $\mathbb{R}^n$ est localement lipschitzien.

Soit $u_0 \in \mathcal{O}$, $\Phi(u_0) = x_0 \in V$. Comme $m \geq 2$, Qn peut trouver un voisinage $\mathcal{O}_0$ de u_0 dans $\mathcal{O}$, et un voisinage Ω_0 de x_0 dans Ω, avec $\Phi(\mathcal{O}_0) \subset \Omega_0$, tels que Φ, Φ', soient bornées dans $\mathcal{O}_0$, et $\Theta, \Theta', \Theta''$ bornées dans Ω_0, et que $\vec{X}$ soit lipschitzien et borné dans $\Omega_0 \cap V$; soit k une constante de Lipschitz et soit M une borne des normes de toutes ces fonctions.

On a la majoration, pour u_1 et u_2 dans $\mathcal{O}_0$:

$$(V.2\,;80)\quad \left\| \vec{U}(u_1) - \vec{U}(u_2) \right\| = \left\| \Theta'(\Phi(u_1)) \cdot \vec{X}(\Phi(u_1)) - \Theta'(\Phi(u_2)) \cdot \vec{X}(\Phi(u_2)) \right\|$$

$$\leq \left\| \Theta'(\Phi(u_1)) \cdot \left(\vec{X}(\Phi(u_1)) - \vec{X}(\Phi(u_2)) \right) \right\|$$

$+\|(\Theta'(\Phi(u_1)) - \Theta'(\Phi(u_2)))$●✠☎♎▤☹⌘))‖

$$\leq Mk \left\| \overline{\Phi(u_1) - \Phi(u_2)} \right\| + \left\| \Theta'(\Phi(u_1)) - \Theta'(\Phi(u_2)) \right\| M$$

La formule des accroissements finis montre que ceci est majoré par

(V,2;81)
$$M^2 k \|\vec{u}_1 - \vec{u}_2\| + M^2 \|\overrightarrow{\Phi(u_1) - \Phi(u_2)}\| ^{*}$$
$$\leq (M^2 k + M^3) \|\vec{u}_1 - \vec{u}_2\| ,$$

donc $\vec{U}$ est bien localement lipschitzien. Mais alors l'équation différentielle $\frac{du}{dt} = \vec{U}(u)$ rentre bien dans les conditions du théorème 1, car $\mathcal{O}$ est un <u>ouvert</u> de $\mathbb{R}^n$.

Donc, étant donné la condition initiale t_0, $u_0 = \Theta(x_0)$ il existe un intervalle $]t_0-\alpha_0, t_0+\beta_0[$ dans lequel existe une solution et une seule de cette équation différentielle, correspondant à cette condition initiale. Si nous appelons $\varphi : t \longrightarrow \varphi(t)$, une telle solution, alors $\Phi \circ \varphi : t \longrightarrow \Phi(\varphi(t))$ est bien solution de (V,2;77), car :

(V,2;82)
$$\begin{cases} \dfrac{\vec{d}}{dt}\big(\Phi(\varphi(t))\big) = \Phi'(\varphi(t)) . \overrightarrow{\varphi'(t)} = \Phi'(\varphi(t)) . \vec{U}(\varphi(t)) \\ \\ = \vec{X}(\Phi(\varphi(t))) \text{ d'après (V,2;79).} \end{cases}$$

Inversement une solutions= $f(t)$ de (V,2;77) provient nécessairement d'une telle solution au moins pour les valeurs de t assez voisines de t_0 pour que $f(t) \in \Phi(\mathcal{O}_0)$, car

(V,2;82 bis)
$$\begin{cases} \dfrac{\vec{d}}{dt}\big(\Theta(f(t))\big) = \Theta'(f(t)) \cdot \overrightarrow{f'(t)} \\ = \Theta'(f(t)) \cdot \vec{X}(f(t)) = \vec{U}(\Theta(f(t))) , \end{cases}$$

* On écrit ici $\|\Theta'(a)-\Theta'(b)\| \leq \|\overrightarrow{a-b}\| \sup_{\xi\in[a,b]} \|\Theta''(\xi)\|$; c'est ici qu'intervient l'hypothèse $m \geq 2$, V doit être au moins de classe C^2 .

ou $\varphi'(t) = \vec{U}(\varphi(t))$, si on pose $\varphi(t) = \Theta(f(t))$

Cela montre bien l'existence de solutions locales de (V,2;77). En outre, deux solutions de (V,2;77), qui prennent en t_0 la même valeur x_0, coïncident nécessairement dans tout un voisinage de x_0, et alors la méthode du théorème 2 et de son corollaire donnent les mêmes résultats pour l'équation (V,2;77) que pour (V,1;1).

Supposons V compacte. Soit $]t_0-\alpha, t_0+\beta[$ l'intervalle maximum de définition d'une solution f ; supposons β fini; on sait que, lorsque $t < t_0 + \beta$ tend vers $t_0+\beta$, $f(t)$ ne peut pas avoir de limite, sans quoi $\vec{X}(f(t))$ en aurait une aussi, et la solution serait prolongeable jusqu'à $t_0+\beta$ lui-même; nous avons vu que c'était impossible. Mais, comme **V** est compacte, on peut trouver une suite $t_1, t_2, \ldots t_n, \ldots$ tendant vers $t_0+\beta$ pour n infini, telle que $f(t)$ ait une limite dans V. Il en est de même si V est Seulement fermée, et si on a la majoration (V,2;78 bis); en effet, dans ce cas, on a l'inéquation différentielle $\left\|\frac{df}{dt}\right\| = \|\vec{X}(f(t))\| \leq \mu \|\overrightarrow{f(t)-O}\| + \nu$, donc la fonction f vérifie la majoration (V,2;36) (où x est remplacée par t), donc $f(t)$ est bornée, c'est-à-dire reste dans une boule fermée, pour $t < t_0+\beta$; elle reste donc dans l'intersection de cette boule fermée et de V, supposée fermée, donc dans un compact de V; on peut donc bien encore trouver une suite t_n tendant vers $t_0+\beta$, telle que $f(t_n)$ ait une limite dans V. Alors le corollaire du théorème 5 (avec $\vec{L}_n = \vec{L}$) nous dit que l'on peut trouver un intervalle $[t_0+\beta-\varepsilon, t_0+\beta+\varepsilon]$, $\varepsilon > 0$, tel que les solutions de (V,2;77), correspondant aux conditions initiales t_n, $f(t_n)$, soient toutes définies dans cet intervalle, pour n assez grand; mais ces solutions sont toutes f elle-même. Cela prouve que la solution f est prolongeable jusqu'à $t_0+\beta+\varepsilon > t_0+\beta$, ce qui est contradictoire. Donc $\beta = +\infty$, et de même $\alpha = +\infty$.

Enfin si X est de classe C^p, $p \leq m$, la démonstration donnée au théorème 8, par récurrence sur m, montre que les solutions sont de classe C^{p+1} .

Corollaire - Soit $\vec{X}$ un champ de vecteurs localement lipschitzien sur un ouvert Ω d'un espace affine E. Soit V une variété fermée de classe C^m, $m \geq 2$ *, dans Ω ; et supposons qu'en chaque point x de V, $\vec{X}(x)$ soit tangent en x à V. Si une solution $f : t \to f(t)$ de l'équation différentielle $\frac{dx}{dt} = \vec{X}(x(t))$ à valeurs dans E, définie sur un intervalle $\mathbb{R}_1$ de $\mathbb{R}$, est telle que $f(t_0)$ soit dans V, alors $f(t)$ est dans V pour tout t (toute courbe intégrale qui a un point dans V est tout entière dans V).

Démonstration - Dans un voisinage de t_0, il existe une solution de l'équation différentielle, à valeurs dans V, d'après le théorème; or il en existe une seule dans E, d'après le théorème 1; donc la solution unique dans E est dans V pour t assez voisin de t_0. Si alors nous appelons $\mathcal{B}$ l'ensemble des t de $\mathbb{R}_1$ pour lesquels $f(t) \in V$, il est ouvert, puisqu'il ne peut contenir un point sans contenir tout un voisinage de ce point; il est fermé, car c'est l'image réciproque de V, supposée fermée, par l'application continue f ; donc c'est $\mathbb{R}_1$ tout entier, puisque $\mathbb{R}_1$ est connexe.

§ 3 ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES

Soient $]a, b[$ un intervalle réel, et $\vec{F}$ un espace vectoriel normé. On appelle équation différentielle linéaire, une équation de la forme :

(V,3;1) $$\vec{y}' = A(x) \cdot \vec{y} ,$$

où, pour tout x de $]a, b[$, $A(x)$ est une application linéaire continue de $\vec{F}$ dans $\vec{F}$, c'est-a-dire un élément de $\mathcal{L}(\vec{F}; \vec{F})$: alors $A : x \to A(x)$, est une application continue de $]a, b[$ dans $\mathcal{L}(\vec{F}; \vec{F})$.

* Pour ce corollaire, on peut voir que $m \geq 1$ est suffisant.

Supposons, par exemple, que $\vec{F}$ soit 1 espace $\mathbb{R}^n$; alors une fonction $\vec{f}$ définie sur $|a, b|$, à valeurs dans $\vec{F}$, est équivalente à un système de n fonctions scalaires : $f_1, f_2, \dots, f_n$. $A(x)$ est une matrice à n lignes et n colonnes. Désignons par $A_{i,j}(x)$ les coefficients de cette matrice. L'équation différentielle peut s'écrire sous la forme matricielle :

$$(V,3;2) \qquad \begin{pmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_n \end{pmatrix} = \begin{pmatrix} A_{1,1}(x) & A_{1,2}(x) \dots & A_{1,n}(x) \\ A_{2,1}(x) & A_{2,2}(x) \dots & A_{2,n}(x) \\ \dots & \dots & \dots \\ A_{n,1}(x) & A_{n,2}(x) \dots & A_{n,n}(x) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

ou encore sous la forme :

$$(V,3;3) \qquad y'_i = \sum_{j=1}^{n} A_{i,j}(x)\, y_j \quad ; \quad i = 1, 2, \dots, n .$$

Supposons maintenant que nous ayons à résoudre une équation différentielle d'ordre p, mise sous la forme :

$$(V,3;4) \qquad \vec{y}^{(p)} = A_0(x) \cdot \vec{y} + A_1(x) \cdot \vec{y}' + \dots + A_{p-1}(x) \cdot \vec{y}^{(p-1)},$$

où les A_i sont des fonctions continues sur $|a, b|$ à valeurs dans $\mathcal{L}(\vec{F}; \vec{F})$.

En posant $y = z_0$, $\vec{y}' = \vec{z}_1, \dots, \vec{y}^{(p-1)} = \vec{z}_{p-1}$, on est ramené au système linéaire

$$(V,3;5) \qquad \begin{cases} \vec{z}'_0 = \vec{z}_1 \\ \vec{z}'_1 = \vec{z}_2 \\ \dots\dots\dots \\ \vec{z}'_{p-2} = \vec{z}_{p-1} \\ \vec{z}'_{p-1} = A_0(x) \cdot \vec{z}_0 + A_1(x) \cdot \vec{z}_1 + \dots + A_{p-1}(x) \cdot \vec{z}_{p-1} \end{cases}$$

C'est tout simplement un système de n équations différentielles du premier ordre, analogue au système (V,3;2), qui peut s'écrire sous la forme matricielle (avec une matrice d'éléments de $\mathcal{L}(\vec{F};\vec{F})$: c'est une matrice habituelle si $\vec{F}=\mathbb{R}$ ou $\mathbb{C}$).

$$\begin{pmatrix} \vec{z}'_0 \\ \vec{z}'_1 \\ \vdots \\ \vec{z}'_{p-2} \\ \vec{z}'_{p-1} \end{pmatrix} = \begin{pmatrix} 0 & I & 0 & \cdots & 0 & 0 \\ 0 & 0 & I & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & I^* \\ A_0(x) & A_1(x) & A_2(x) & \cdots & A_{p-2}(x) & A_{p-1}(x) \end{pmatrix} \begin{pmatrix} \vec{z}_0 \\ \vec{z}_1 \\ \vdots \\ \vec{z}_{p-2} \\ \vec{z}_{p-1} \end{pmatrix}$$

Théorème 9 - Si $\vec{F}$ est complet, une équation différentielle linéaire du type (V,3;1) admet une solution et une seule, définie dans tout l'intervalle $]a,b[$, correspondant à une condition de Cauchy x_0, y_0 donnée; en outre, si A est une fonction de classe C^m de $]a,b[$ dans $\mathcal{L}(\vec{F};\vec{F})$, la solution est une fonction de classe C^{m+1} de $]a,b[$ dans $\vec{F}$.

Démonstration - A étant une application continue de $]a,b[$ dans $\mathcal{L}(\vec{F};\vec{F})$, elle est bornée sur tout intervalle compact $[a',b']$ de $]a,b[$. Soit M la borne supérieure de sa norme. Alors la fonction $\vec{L}$, définie ici par $L(x,\vec{y})=A(x)\cdot\vec{y}$, vérifie d'une part la majoration :

$$\|\vec{L}(x,\vec{y})\| \leq M\,\|\vec{y}\| \tag{V,3;7}$$

et d'autre part, la condition de Lipschitz par rapport à $\vec{y}$:

$$\|\vec{L}(x,\vec{y}_1)-\vec{L}(x,\vec{y}_2)\| \leq M\,\|\vec{y}_1-\vec{y}_2\| \tag{V,3;8}$$

Il en résulte que, dans l'intervalle $[a',b']$, nous pouvons appliquer le théorème 4, qui démontre l'existence et

* I est l'application identique de $\vec{F}$ dans $\vec{F}$, $I\in\mathcal{L}(\vec{F};\vec{F})$

l'unicité de la solution (D'ailleurs ce qui a été dit à la remarque 2° page 11 montre directement que le système de sécurité n'est plus nécessaire, et que la méthode des approximations successives est uniformément convergente dans $[a', b']$). Comme ceci est vrai pour tout $[a', b']$, c'est vrai pour $]a, b[$ lui-même. La classe C^{m+1} de la solution se déduit alors du théorème 8.

Théorème 10 - L'ensemble $\mathcal{E}$ des solutions de (V,3;1) est un sous-espace vectoriel de $\vec{F}^{]a,b[}$. L'application qui, à chaque solution $\vec{f}$, fait correspondre sa valeur $\vec{f}(x_0)$ au point x_0, est une bijection linéaire de $\mathcal{E}$ sur $\vec{F}$.

Démonstration - A cause de la linéarité de l'équation, la somme de 2 solutions est une solution, ainsi que le produit d'une solution par une constante, donc l'ensemble $\mathcal{E}$ des solutions est bien un espace vectoriel.

L'application qui, à $\vec{f} \in \mathcal{E}$, fait correspondre $\vec{f}(x_0) = \vec{y}_0$, est bien linéaire; le théorème 9 d'existence et d'unicité dit que c'est une bijection.

Corollaire 1 - Pour que k solutions soient des éléments indépendants dans l'espace vectoriel $\mathcal{E}$, il faut et il suffit que les valeurs de ces solutions en un point x_0 de $]a, b[$ soient k vecteurs indépendants de $\vec{F}$. Une solution, nulle en x_0, est identiquement nulle.

Cela résulte du caractère injectif de l'application $\vec{f} \longrightarrow \vec{f}(x_0)$ de $\mathcal{E}$ dans $\vec{F}$.

Corollaire 2 - Si $\vec{F}$ est de dimension finie n, l'ensemble $\mathcal{E}$ des solutions est aussi de dimension n.

Conséquence évidente du théorème.

Si, dans ce cas, nous choisissons dans $\vec{F}$ une base, $\vec{e}_1, \vec{e}_2, \ldots \vec{e}_n$, (et ce sera automatique si $\vec{F}$ est donné comme 1 espace $\mathbb{R}^n$), à chacun des vecteurs $\vec{e}_i$, considéré comme condition initiale au point x_0 de $|a,b|$ correspond une solution.

On a ainsi n solutions indépendantes, $\vec{\eta}_0, \vec{\eta}_1, \ldots, \vec{\eta}_n$ qui forment une base de l'espace vectoriel $\mathcal{E}$ des solutions. Si alors on se donne une condition initiale $\vec{y}_0 = \sum_{i=1}^{n} u_i \vec{e}_i$ il lui correspond l'unique solution $\vec{f} = \sum_{i=1}^{n} u_i \vec{\eta}_i$.

La fonction $\vec{\eta}_i$ se développe, suivant la base de $\vec{F}$, sous la forme :

$$\vec{\eta}_i(x) = \sum_{k=1}^{n} \eta_{i,k}(x)\, \vec{e}_k \quad , \tag{V,3;9}$$

où les $\eta_{i,k}$ sont des fonctions scalaires; et la solution correspondant aux conditions initiales ci-dessus a la forme

$$\vec{f}(x) = \sum_{i,k=1}^{n} u_i\, \eta_{i,k}(x)\, \vec{e}_k \ . \tag{V,3;10}$$

Les $\eta_i(x)$ et $\eta_{i,j}$ (30) que nous venons de déterminer sont relatifs au choix du point initial x_0 . Il est utile parfois de les noter $\eta_i(x, x_0)$ et $\eta_{i,k}(x, x_0)$.

Si on appelle f_k la composante de $\vec{f}$ suivant $\vec{e_k}$, les fonctions f_k satisfont au système différentiel (V,3;2 ou 3), et (V,3;10) s'écrit :

$$f_k(x) = \sum_{i=1}^{n} \eta_{i,k}(x, x_0)\, f_k(x_0). \tag{V,3;10 bis}$$

Considérons le cas particulier d'une équation différentielle scalaire * d'ordre? du type (V,3;4). En faisant le changement de fonction déjà indiqué, on est ramené à une seule équation différentielle, relative à l'espace $\vec{F} = \mathrm{R}^p$. La nouvelle inconnue est $\vec{g} = (f, f', f'', \ldots, f^{(p-1)})$ fonction définie sur $|a,b|$ à valeurs dans R^p .

* Les $A_q(x)$, $q = 0, 1, 2, \ldots p-1$, sont donc des fonctions continues scalaires. Le corps des scalaires sera généralement, le corps des complexes $\mathbb{C}$.

La condition initiale est l'équivalent de la donnée, au point x_0, du vecteur $\vec{y}_0$ de R^p, donc des scalaires $y_0^{(q)}, q = 0, 1, 2, \ldots, p-1$, valeurs de la fonction et de ses dérivées d'ordre $1, 2, \ldots, p-1$; c'est dans le sens de l'indépendance des vecteurs de R^p qu'on parlera de l'indépendance d'un système de k conditions initiales.

Z

Par contre, si l'on considère k solutions de l'équation, le mot : "dependantes ou Indépendantes" est a priori ambigu. S'agit-il de l'indépendance des fonctions $\vec{g}$, ou de 1 indépendance des fonctions f ?

Il est facile de voir que ces deux notions d'indépendance sont Identiques, de même que les notions correspondantes de dépend-. Si en effet entre-ii solutions $\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_k$, il existe une relation linéaire à coefficients constants non tous nuls, la même relation existe évidemment entre les f_i. Mais, Inversement, si une telle relation existe entre les fonctions $f_1, f_2, \ldots, f_k$, elle existe aussi entre leurs dérivées premières, entre leurs dérivées secondes, ... entre leurs dérivées d'ordre $p-1$, c'est-b-dire entre les fonctions $\vec{g}_i$ correspondantes.

L'ensemble des solutions f de l'équation (V,3;4) est donc bien encore un sous-espace vectoriel de dimension n de $\mathscr{C}(|a,b|)$. En relation avec ce qui a été dit auparavant, nous serons amenés à appeler η_i, $i = 0, 1, 2, \ldots, p-1$, * la solution de l'équation (ii,3;4) correspondant à la condition initiale :

(V,3;11) $$\eta_i^{(q)}(x_0) = 0 \quad \text{pour } q \neq i, \; \eta_i^{(i)}(x_0) = 1, \text{ ou}$$

(V,3;12) $$\eta_i^{(q)}(x_0) = \delta_i^q ,$$

où δ_i^q est le symbole de Kronecker. SI alors une solution est définie par ses conditions initiales :

(V,3;13) $$u^{(q)} = y_0^{(q)} = f^{(q)}(x_0)$$

elle est donnée par la formule :

* On est amené à prendre les Indices $0, 1, 2, \ldots, p-1$, pour numéroter les éléments de base de R^p.

(V,3;14) $$f(x)=\sum_{i=0}^{p-1} y_0^{(i)}\eta_i(x) \text{ ou } f(x)=\sum_{i=0}^{p-1}\eta_i(x,x_0)f^{(i)}(x_0).$$

Rappelons, au sujet de ces équations différentielles scalaires d'ordre p, des résultats vus en mathématiques spéciales.

Si $p = 1$, l'équation se résout immédiatement par quadrature; en effet l'équation s'écrit alors :

(V,3;15) $$\frac{y'}{y} = A(x);$$

la solution est donnée par :

(V,3;16) $$y = Ce^{\int A(x)dx} \quad ; \quad f(x) = y_0\, e^{\int_{x_0}^{x} A(\xi)d\xi}$$

Au contraire, dès que l équation différentielle est d'ordre ≥ 2, on ne peut pas exprimer, d'une façon générale, sa solution par des intégrales.

On cherchera donc, par la connaissance de solutions particulières, à abaisser le degré de l'équation. Si, pour l'équation (V,3:4), on possède une solution particulière Y, alors on fera le changement de fonctions $y = Yu$. On vérifie immédiatement que u est solution d'une équation différentielle, dans laquelle le coefficient de u est absent.

Autrement dit, en posant $u' = v$, on a à résoudre, pour v, une équation différentielle d'ordre $p - 1$, d'où l'on déduira ensuite& par une quadrature.

Lorsqu'il s'agit d'un système différentiel, ou d'une équation différentielle correspondant à un espace F de dimension ≥ 2, alors, même si l'équation est du premier ordre, on ne sait évidemment pas la résoudre explicitement par des quadratures, puisqu'une équation unique scalaire du second ordre se ramène à un système de deux équations du premier ordre.

Résolvante d'une équation différentielle linéaire

Supposons $\vec{F}$ complet, soit x_1, x_2, deux points de l'intervalle $|a,b|$. D'après le théorème 9, il existe une solution unique de l équation (V,3;1), prenant une valeur donnée :

$\vec{y}_1$ au point x_1 . Cette solution prend alors une valeur déterminée au point x_2 , soit $\vec{y}_2$. On détermine ainsi une application bien déterminée de $\vec{F}$ dans F , celle qui fait correspondre à la valeur de la solution au point x_1 sa valeur au point x_2 . Cette application est manifestement linéaire, car, à la somme de deux valeurs initiales au point x_1 , correspond, comme solution, la somme des solutions correspondantes, et par conséquent, comme valeur au point x_2, la somme des valeurs correspondantes : et de même pour le produit par un scalaire. Si d'ailleurs nous appelons π_x la bijection de $\mathscr{E}$, espace des solutions, sur $\vec{F}$, qui fait correspondre à chaque solution $\vec{f}$, sa valeur au point-s : $\pi_x \vec{f} = \vec{f}(x)$ (théorème 10), l'application considérée maintenant est simplement $\pi_{4} \circ \pi_{x_1}^{-1}$.

Cette application linéaire de $\vec{F}$ dans $\vec{F}$ est continue. Si, en effet, on considère une suite de valeurs initiales $(\vec{y}_1)_\nu$, 3 = 0,1,2,..., correspondant au même point x_1, et convergeant, pour V infini, vers une valeur initiale limite $\vec{y}_1$ dans $\vec{F}$, le théorème 6, de continuité de la solution par rapport aux conditions initiales, montre que la suite des solutions $\vec{f}_\nu$, correspondant aux valeurs initiales $(\vec{y}_1)_\nu$, converge vers la solution $\vec{f}$ correspondant à la condition initiale $\vec{y}_1$, et uniformément sur tout compact de l'intervalle $]a,b[$; en particulier au point x_2, ce qui démontre bien la continuité de $\pi_{x_2} \circ \pi_{x_1}^{-1}$.

L'application précédente $\pi_{x_2} \circ \pi_{x_1}^{-1}$ est donc une application linéaire continue de $\vec{F}$ dans $\vec{F}$, c'est donc un élément de $\mathcal{L}(\vec{F};\vec{F})$. Cet élément se note $R(x_2, x_1)$ * . Cet opérateur est entièrement caractérisé par le fait que, pour toute solution $\vec{f}$ de 1 équation différentielle, on a la formule :

$$\vec{f}(x_2) = R(x_2, x_1) \cdot \vec{f}(x_1) \; ; \text{ ou } R(x_2, x_1) = \pi_{x_2} \circ \pi_{x_1}^{-1} . \tag{V,3;17}$$

$R(x_2,x_1)$ est appelé l'opérateur résolvant (ou la résolvante) de l'équation différentielle, relativement aux deux points x_1 et x_2 de $]a,b[$. Ce qu'on appelle généralement la résolvante de l'équation différentielle, c'est la fonction $R : (x_2,x_1) \to R(x_2,x_1)$; c'est une application de

* On écrit $R(x_2,x_1)$ plutôt que $R(x_1,x_2)$, pour la commodité des formules ultérieures, notamment (V,3;18).

$]a,b[\times]a,b[$ dans $\mathcal{L}(\vec{F};\vec{F})$. Si on a choisi une base de $\vec{F}$, avec les notations de la page 786, la résolvante est représentée par la matrice transposée de celle des η_{ij} : $R_{j,\ell}(x_2,x_1) = \eta_{\ell,j}(x_2,x_1)$.

<u>Théorème 11</u> - <u>La résolvante vérifie les propriétés fondamentales suivantes</u> :

$$\text{(V,3;18)} \qquad \begin{cases} R(x_3,x_2) \circ R(x_2,x_1) = R(x_3,x_1) \\ R(x_1,x_2) = \left(R(x_2,x_1)\right)^{-1} \\ R(x,x) = I. \end{cases}$$

<u>Démonstration</u> - Cela résulte immédiatement de la définition $R(x_2,x_1) = \pi_{x_2} \circ \pi_{x_1}^{-1}$ (ou, de manière plus imagée, de la première formule (V,3;17), qui exprime que $R(x_2,x_1)\cdot\vec{y}_1$ donne la valeur en x_2 de la solution prenant la valeur $\vec{y}_1$ en x_1).

<u>Théorème 12</u> - <u>L'application</u> R, <u>de</u> $]a,b[\times]a,b[$ <u>dans</u> $\mathcal{L}(\vec{F};\vec{F})$ <u>est la seule application qui soit partiellement dérivable par rapport à la première variable, et vérifie, d'une part l'équation différentielle</u>

$$\text{(V,3;19)} \qquad \frac{\partial R}{\partial x}(x,\xi) = A(x) \circ R(x,\xi),$$

<u>et d'autre part, la condition initiale</u> :

$$\text{(V,3;20)} \qquad R(x,x) = 1$$

Démonstration - Si nous fixons en effet ξ dans l'intervalle $]a, b[$, et considérons alors R comme fonction de x seul, c'est une fonction, non pas à valeurs dans l'espace vectoriel normé $\vec{F}$, mais dans l'espace vectoriel normé $\vec{G} = \mathcal{L}(\vec{F}; \vec{F})$; cet espace est complet, puisque $\vec{F}$ est complet (théorème 50 du chapitre II).

Considérons alors, pour une fonction définie sur $|a, b|$ à valeurs dans $\vec{G}$, l'équation différentielle :

(V,3;20) $$\vec{Y}' = A(x) \circ \vec{Y}$$

Vérifions que c'est bien la une équation différentielle linéaire au sens de (V,3;1).

Pour x fixé, $A(x)$ est un élément de $\mathcal{L}(\vec{F}; \vec{F})$; mais alors $\vec{Y} \longrightarrow A(x) \circ \vec{Y}$, est bien une application linéaire continue $B(x)$ de $\vec{G} = \mathcal{L}(\vec{F}; \vec{F})$ dans lui-même; $B(x) \in \mathcal{L}(\vec{G}; \vec{G})$; (V,3;20) s'écrit $Y' = B(x) \cdot \vec{Y}$. De plus, en vertu du théorème 54 du chapitre II, la norme $\|B(x)\|$ de $B(x)$ dans $\mathcal{L}(\vec{G}; \vec{G})$, soit $\displaystyle\sup_{\substack{u \in \mathcal{L}(\vec{F}; \vec{F}) \\ \|u\| \leq 1}} \|A(x \circ u)\|$, est $\leq \|A(x)\|$, norme de **A** (x) dans $\mathcal{L}(\vec{F}; \vec{F})$.

Pour démontrer que nous avons bien affaire à une équation différentielle du type (V,3:1), nous devons démontrer que **B**: $x \longrightarrow B(x)$, est une application continue de $|a, b|$ dans $\mathcal{L}(\vec{G}; \vec{G})$; autrement dit, nous devons montrer que. lorsque x tend vers x_0, $\|B(x) - \mathbf{B}(x_0)\|$ tend vers **0**. Or cette quantité, d'après la définition même de la norme, est majorée par $\|A(x) - A(x_0)\|$, norme dans $\mathcal{L}(\vec{F}; \vec{F})$, **elle** tend donc bien vers 0, d'après l'hypothèse suivant laquelle **A** est une application continue de x dans $\mathcal{L}(\vec{F}; \vec{F})$.

Nous pourrons donc appliquer le théorème 9, et affirmer qu'il existe une solution et une seule de cette équation différentielle. qui satisfait à la condition initiale $Y(\xi) = \mathbf{1}$, opérateur identique de $\vec{F}$ dans $\vec{F}$, élément particulier de $\vec{G} = \mathcal{L}(\vec{F}; \vec{F})$.

Appelons donc $S(x;\xi)$ la solution de cette équation différentielle.

Alors, étant donné un vecteur $\vec{y}_1$ de $\vec{F}$, considérons la fonction $\vec{f}$, définie sur $|a,b|$ à valeurs dans $\vec{F}$:

$$\vec{f}(x) = S(x;\xi)\cdot\vec{y}_1 \quad * \tag{V,3;21}$$

C'est une fonction dérivable et sa dérivée, d'après le corollaire 1 du théorème 11 du chapitre III (permutabilité de la dérivation et des applications linéaires continues), est

$$\vec{f}'(x) = \frac{\partial S}{\partial x}(x,\xi)\cdot\vec{y}_1 \,. \tag{V,3;22}$$

Mais comme $\vec{Y}$ est solution de l'équation différentielle (V,3:20), on a encore :

$$\vec{f}'(x) = (A(x)\circ S(x,\xi))\cdot\vec{y}_1 = A(x)\cdot\vec{f}(x)\,. \tag{V,3;23}$$

Cela prouve que $\vec{f}$ est une solution de l'équation différentielle (V,3;1); mais comme on a la condition initiale $\vec{f}(\xi) = S(\xi,\xi)\cdot\vec{y}_1 = \vec{y}_1$, $\vec{f}$ est l'unique solution de l'équation différentielle (V,3;1) correspondant à la condition initiale $\vec{f}(\xi) = \vec{y}_1$; comme alors sa valeur au point x est donnée par (V,3;21), et que ce résultat est vrai quel que soit le vecteur $\vec{y}_1$ choisi au départ, on a nécessairement $S(x,\xi) = R(x;\xi)$, ce qui démontre le théorème.

<u>Théorème 13</u> - <u>Si, dans l'équation (V,3;1), A est une fonction de classe C^m, alors la résolvante R est une fonction de classe C^{m+1} des variables x et ξ, autrement dit une application de classe C^{m+1} de $|a,b|\times|a,b|$ dans $\mathcal{L}(\vec{F};\vec{F})$.</u>

* $\vec{f}$ dépend aussi du choix de $\vec{y}_1$, on devrait écrire $\vec{f}_{\vec{y}_1}(x)$.

Démonstration - Le fait que R admette des dérivées successives de tous les ordres, par rapport à la variable x, résulte de ce qu'elle vérifie une équation différentielle (V,3;19), dont le coefficient est indéfiniment dérivable, et du corollaire du théorème 8.

Par ailleurs l'équation différentielle à laquelle satisfait R est indépendante du paramètre ξ, mais c'est sa condition initiale qui dépend continuement du parametre ξ dans le sens indique au théorème 6. Il résulte par conséquent de ce théorème que R est uns application continue de $]a,b[\times]a,b[$ dans $\mathcal{L}(\vec{F},\vec{F})$. Par ailleurs, sa dérivée première par rapport à x, donnée par la formule (V,3;19), est elle ~~as~~ si une application continue de $]a,b[$ y $]a,b[$ dans $\mathcal{L}(\vec{F},\vec{F})$.

Pour montrer que R admet une dérivée partielle par rapport à ξ, il faudrait utiliser un théorème de dérivabilité par rapport au paramètre ξ.

Mais nous allons donner ici une démonstration directe.

Nous utiliserons la deuxième égalité (V,3;18), c'est-a-dire $R(x,\xi)=(R(\xi,x))^{-1}$ (*). On voit que la dérivée partielle en ξ, c'est-a-dire la dérivée partielle par rapport à la deuxième variable de R n'est autre que la dérivée partielle, par rapport à la Première variable, de R^{-1}., D'~~autre part~~, nous savons que l'application $R \longrightarrow R^{-1}$ est dérivable, et nous connaissons la formule de dérivée de l'inverse d'un opérateur, formule (III,8;32); nous voyons donc bien que R est dérivable par rapport à ξ, et nous obtenons :

(V,3;24) $$\frac{\partial R}{\partial \xi}(x,\xi) = \frac{\partial}{\partial \xi}(R(\xi,x))^{-1} = -(R(\xi,x))^{-1}\circ\frac{\partial R}{\partial \xi}(\xi,x)\circ(R(\xi,x))^{-1}$$

En remplaçant alors le terme du milieu à partir de l'équation différentielle (V,3;19), on obtient la formule remarquable :

(V,3;24bis) $$\frac{\partial R}{\partial \xi}(x,\xi) = -R(x,\xi)\circ(A(\xi)\circ R(\xi,x))\circ(R(\xi,x))^{-1} \quad \text{ou}$$

(V,3;25) $$\frac{\partial R}{\partial \xi}(x,\xi) = -R(x,\xi)\circ A(\xi)\,.$$

Elle est analogue à la formule (V,3;19), mais avec le signe $-$, et la composition avec l'opérateur A se fait à droite et non pas à gauche. Il en résulte donc bien que R

admet une dérivée partielle par rapport à ξ, continue par rapport à l ensemble des variables x,ξ, et par conséquent, du théorème 15 du chapitre III, on déduit que R est une application de classe C^1 de $|a,b| \times |a,b|$ dans $\mathcal{L}(\vec{F};\vec{F})$

Il suffit maintenant de faire une récurrence sur m. Supposons démontré que, si **A** est de classe C^{m-1} de $|a,b|$ dans $\mathcal{L}(\vec{F};\vec{F})$, R est de classe C^m de $|a,b| \times |a,b|$ dans $\mathcal{L}(\vec{F};\vec{F})$; et supposons **A** de classe C^m, $m \geq 1$.

Alors elle est en tout cas de classe C^{m-1}, donc **R** est en tout cas de classe C^m. Mals alors $\frac{\partial R}{\partial x}$ et $\frac{\partial R}{\partial \xi}$ donnés par (V,3;19) et (V,3;25), sont de classe C^m, d'après le théorème **18** du chapitre III; donc R est de classe C^{m+1}, et le théorème est démontré.

Equation linéaire avec second membre

$\vec{F}$ étant toujours un espace vectoriel normé, on appelle '<équation linéaire avec second membre" une équation de la forme :

(V,3;26) $$\vec{y}' = A(x) \cdot \vec{y} + \vec{B}(x) ,$$

où A est une application continue de $|a,b|$ dans $\mathcal{L}(\vec{F},\vec{F})$,

et où $\vec{B}$ est une application continue de $|a,b|$ dans?. Cette expression : '<linéaire avec second membre" est évidemment impropre, on devrait plutôt dire : une équation différentielle affine; en effet, pour tout x, 1 application $\vec{y} \rightarrow A(x) \cdot \vec{y} + \vec{B}(x)$ est une application affine de $\vec{F}$ dans $\vec{F}$. La fonctions s'appelle le second membre de l'équation différentielle.

On appelle équation linéaire associée ou linéaire sans second membre associée, ou homogène associée, l'équation (V,3;1) correspondant à la même fonction **A** * .

* L'étude des équations différentielles affines devrait se faire dans les espaces affines. Mais, pratiquement, $\vec{F}$ est bien toujours vectoriel, et pas seulement affine.

Théorème 14 - On obtient la solution générale de l'équation linéaire avec second membre (V,3;26), en ajoutant, à une solution particulière, la solution générale de l'équation sans second membre associée.

Démonstration évidente.

Naturellement ceci ne prouve pas l'existence de solutions puisqu'on doit supposer l'existence d'une solution particulière; néanmoins celle-ci résulte immédiatement du théorème 4, si $\vec{F}$ est complet : pour toute condition initiale $x_0, \vec{y}_0$, il existe une solution et une seule, définie dans tout l'intervalle $]a, b[$.

Théorème 15 - Pour la condition initiale donnée $x_0, \vec{y}_0$, l'unique solution de (V,3;26) ($\vec{F}$ complet) est donnée par la formule :

$$\vec{f}(x) = R(x, x_0) \cdot \vec{y}_0 + \int_{x_0}^{x} R(x, \xi) \cdot \vec{B}(\xi)\, d\xi\,, \tag{V,3;27}$$

dans laquelle R est la résolvante de l'équation homogène associée.

Cette formule montre que, si l'on sait résoudre l'équation différentielle sans second membre associée, on sait résoudre, par des quadratures, l'équation avec second membre, quel qu'en soit le second membre.

Démonstration - Nous utiliserons la méthode dite des constantes variables. Si on avait à résoudre l'équation homogène, on trouverait la solution générale :

$$\vec{f}(x) = R(x, x_0) \cdot \vec{C}\,, \tag{V,3;28}$$

dans laquelle $\vec{C}$ est une constante, à savoir la valeur initiale $\vec{y}_0$. Nous ferons donc comme si cette constante était variable, autrement dit nous ferons le changement de fonctions :

$$\vec{f}(x) - R(x, x_0) \cdot \vec{C}(x) \tag{V,3;29}$$

On a bien le droit de faire ce changement de fonctions; autrement dit, on sait d'avance que toute fonction $\vec{f}$ peut se mettre, d'une manière et d'une seule, sous la forme (V,3;29), car il suffit de poser

$$\vec{C}(x) = (R(x,x_0))^{-1}\cdot\vec{f}(x) = R(x_0,x)\cdot\vec{f}(x)\ . \tag{V,3;30}$$

D'autre part, $\vec{C}$ est dérivable, si et seulement si $\vec{f}$ est dérivable. On a, d'après le théorème 12 du chapitre III, et compte tenu de ce que l'application bilinéaire canonique de $\mathcal{L}(\vec{F};\vec{F})\times\vec{F}$ dans $\vec{F}$ est continue, la formule de dérivation :

$$\begin{aligned}\vec{f}'(x) &= \frac{\partial R}{\partial x}(x,x_0)\cdot\vec{C}(x) + R(x,x_0)\cdot\vec{C}'(x)\\ &= (A(x)\circ R(x,x_0))\cdot\vec{C}(x) + R(x,x_0)\cdot\vec{C}'(x).\\ &= A(x)\cdot\vec{f}(x) + R(x,x_0)\cdot\vec{C}'(x).\end{aligned} \tag{V,3;31}$$

Elle donne, si on la transporte dans l'équation (V,3;26), et compte tenu de la disparition du terme en $\vec{C}$:

$$R(x,x_0)\cdot\vec{C}'(x) = \vec{B}(x)\text{, ou} \tag{V,3;32}$$

$$\vec{C}'(x) = R(x_0,x)\cdot\vec{B}(x) \tag{V,3;33}$$

Ainsi on obtient la solution générale de l'équation en prenant pour C n'importe quelle primitive de la fonction $x \longrightarrow R(x_0,x)\cdot\vec{B}(x)$.

Compte tenu de notre condition initiale, $\vec{C}$ doit prendre la valeur initiale

$$\vec{C}(x_0) = R(x_0,x_0) = \vec{f}(x_0) = \vec{y}_0\ , \tag{V,3;34}$$

et par conséquent l'unique solution possible est donnée par

$$\vec{C}(x) = \vec{y}_0 + \int_{x_0}^{x} R(x_0,\xi)\cdot\vec{B}(\xi)\,d\xi\ , \tag{V,3;35}$$

ce qui donne pour $\vec{f}$ la formule :

$$(V,3;36)\qquad \vec{f}(x) = R(x,x_0)\cdot\vec{y}_0 + R(x,x_0)\cdot\int_{x_0}^{x} R(x_0,\xi)\vec{B}(\xi)\,d\xi.$$

Mais, pour x et x_0 donnés, $R(x,x_0)$ est un opérateur linéaire continu de $\vec{F}$ dans $\vec{F}$. On peut donc le faire rentrer sous le signe d'intégration (theoreme b du chapitre IV), et l'on obtient finalement la formule :

$$(V,3;37)\qquad \vec{f}(x) = R(x,x_0)\cdot\vec{y}_0 + \int_{x_0}^{x}\big(R(x,x_0)\circ R(x_0,\xi)\big)\cdot\vec{B}(\xi)\,d\xi$$

qui, compte tenu de (V,3;18), n'est autre que la formule (V,3;27) annoncée

Remarque - Les 2 termes de (V,3;27) ont des significations importantes. Le premier est la solution de l'équation sans second membre correspondant à la condition initiale donnée $\vec{y}_0$. Le second est la solution de l'équation donnée, avec second membre, mais avec la condition initiale $\vec{0}$.

On peut alors retrouver la formule (V,3;27) par une vérification directe. On sait en effet déjà que $R(x,x_0)\cdot\vec{y}_0$ est la solution de l'équation homogène, avec condition initiale $\vec{y}_0$. Considérons alors le terme intégral. Il s'annule bien pour $x = x_0$. Il est dérivable en x ; et sa dérivée en x s'obtient comme celle de la fonction (IV,9;41), avec la méthode de la page 725.

$$(V,3;37\,bis)\qquad \frac{d}{dx}\left(\int_{x_0}^{x} R(x,\xi)\cdot\vec{B}(\xi)\,d\xi\right) = R(x,x)\cdot\vec{B}(x)$$

$$+\int_{x_0}^{x}\frac{\partial R}{\partial x}(x,\xi)\cdot\vec{B}(\xi)\,d\xi$$

$$= \vec{B}(x) + \int_{x_0}^{x} A(x)\cdot\big(R(x,\xi)\cdot\vec{B}(\xi)\big)\,d\xi$$

$$= \vec{B}(x) + A(x)\cdot\int_{x_0}^{x} R(x,\xi)\cdot\vec{B}(\xi)\,d\xi\ ,$$

ce qui prouve que le terme intégral est bien solution de l'équation différentielle (V,3;26) avec la condition initiale $\vec{0}$. Donc le second membre de (V,3;27) est bien une solution de (V,3;26), avec la condition initiale $\vec{y_0}$, et nous savons qu'il n'y a qu'une solution, ce qui redémontre bien le théorème 15.

Cas d'une équation différentielle scalaire d'ordre m avec second membre

Cette équation se présente sous la forme :

$$y^{(p)} = A_0(x)\, y + A_1(x)\, y' + \cdots + A_{p-1}(x)\, y^{(p-1)} + B(X)\,. \tag{V,3;38}$$

En prenant toujours, comme nouvelle fonction, la fonction $z=(z_0, z_1, \ldots, z_{p-1}) = (y, y', y'', \ldots, y^{(p-1)})$ déjà prise auparavant, on peut mettre l'équation avec second membre sous la forme matricielle :

$$\begin{pmatrix} z'_0 \\ z'_1 \\ \vdots \\ z'_{p-2} \\ z'_{p-1} \end{pmatrix} = \begin{vmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ A_0(x) & A_1(x) & A_2(x) & & A_{p-2}(x) & A_{p-1}(x) \end{vmatrix} + \begin{vmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ B(x) \end{vmatrix} . \tag{V,3;39}$$

Compte tenu de la valeur de la résolvante donnée page avec les notations de la page on voit que la solution de cette équation correspondant à la condition initiale $(y_0, y'_0, y''_0, \ldots y_0^{(p-1)})$ est donnée par une formule, dont nous ne retiendrons que ce qui donne $z_0 = y$:

$$f(x) = \sum_{i=0}^{p-1} \eta_i(x, x_0)\, f(x_0) + \int_{x_0}^{x} \eta_{p-1}(x, \xi)\, B(\xi)\, d\xi\,. \tag{V,3;40}$$

Si, en particulier, on cherche la solution de l'équation avec second membre B , correspondant aux conditions initiales $y_0 = 0,\ y'_0 = 0, \ldots, y_0^{(p-1)}(0) = 0$, elle est donnée par la formule :

$$f(x) = \int_{x_0}^{x} \eta_{p-1}(x,\xi)\, B(\xi)\, d\xi . \tag{V,3;41}$$

Ici encore on peut faire une vérification analogue à celle de la formule (V,3;37 bis). On a :

$$\frac{d}{dx}\int_{x_0}^{x} \eta_{p-1}(x,\xi)\, B(\xi)\, d\xi \tag{V,3;41 bis}$$

$$= \eta_{p-1}(x,x)\, B(x) + \int_{x_0}^{x} \eta'_{p-1}(x,\xi)\, B(\xi)\, d\xi$$

Le premier terme est nul. par définition même de η_{p-1} .
De proche en proche, on a ainsi :

$$\left(\frac{d}{dx}\right)^{k} \int_{x_0}^{x} \eta_{p-1}(x,\xi)\, B(\xi)\, d\xi \tag{V,3;41 ter}$$

$$= \int_{x_0}^{x} \eta^{(k)}_{p-1}(x,\xi)\, B(\xi)\, d\xi, \text{ pour } k \leq p-1 \quad ;\text{et}$$

$$\left(\frac{d}{dx}\right)^{p} \int_{x_0}^{x} \eta_{p-1}(x,\xi)\, B(\xi)\, d\xi \tag{V,3; 41 quarto}$$

$$= \int_{x_0}^{x} \eta^{(p)}_{p-1}(x,\xi)\, B(\xi)\, d\xi + \eta^{(p-1)}_{p-1}(x,x)\, B(x)$$

$$= \int_{x_0}^{x} \eta^{(p)}_{p-1}(x,\xi)\, B(\xi)\, d\xi + B(x) \ .$$

On en déduit :

$$\left(\left(\frac{d}{dx}\right)^{p} - \sum_{i=0}^{p-1} A_i(x)\left(\frac{d}{dx}\right)^{i}\right) \int_{x_0}^{x} \eta_{p-1}(x,\xi)\, B(\xi)\, d\xi \tag{V,3;41 quinto}$$

$$= \int_{x_0}^{x} \left[\left(\left(\frac{d}{dx}\right)^{p} - \sum_{i=0}^{p-1} A_i(x)\left(\frac{d}{dx}\right)^{i}\right) \eta_{p-1}(x,\xi)\right] B(\xi)\, d\xi + B(x)$$

$$= B(x) \ ,$$

puisque η_{p-1} est solution de l'équation homogène.

L'intégrale écrite est donc bien solution de l'équation avec second membre, avec des conditions initiales nulles au point x_0 .

Par ailleurs, dans un problème d'équation linéaire avec second membre, on peut toujours, quitte a modifier éventuellement le second membre donné, se ramener au cas où toutes les conditions initiales sont nulles * . Nous voyons donc qu'on sait complètement resoudre l'equation différentielle avec second membre, si l'on connait la seule fonction $\eta_{p-1}(x,\xi)$, correspondant aux conditions initiales $y_0(\xi)=0,\ y_0'(\xi)=0 \ldots y_0^{(p-2)}(\xi)=0,\ y_0^{(p-1)}(\xi)=1$; mais ceci, naturellement, pour tout point ξ de $]a,b[$. ** .

Si nous considérons le cas particulier où tous les A_i sont nuls, nous avons l'équation :

(V,3;42) $$y^{(p)} = B(x)$$

La fonction $\eta_{p-1}(x,\xi)$, solution de l'équation homogène $y^{(p)}=0$, correspondant aux conditions initiales $0,0,0,\ldots,1$ au point ξ , est $\dfrac{(x-\xi)^{m-1}}{(m-1)!}$.

* Posons en effet $g(x)=f(x)-\sum_{q=0}^{p-1}\frac{(x-x_0)^q}{q!}f^{(q)}(x_0)$. Elle est nulle en x_0 , ainsi que ses dérivées d'ordre $1,2,\ldots,p-1$; si on lui applique l'opérateur différentiel $\left(\frac{d}{dx}\right)^p-\sum_{i=0}^{p-1}A_i(x)\left(\frac{d}{dx}\right)^i$

on trouve la différence de $B(x)$ et d'une fonction connue résultat de l'opérateur différentiel sur le polynôme $\sum_{q=0}^{p-1}\frac{(x-x_0)^q}{q!}f^{(q)}(x_0)$. Donc g est solution d'une équation avec second membre (différent de B) connu, et conditions initiales nulles en x_0 .

** Ces conditions initiales 0, 0, 0....,1, jouent donc un rôle très particulier, que nous retrouverons souvent ultérieurement. On peut les appeler conditions initiales de Jean Mineur (cinéma et publicité, tél. 00.01).

On voit donc que la solution de (V,3;42), correspondant à des conditions initiales nulles au point $x_0 = c$, est, d'après (V,3;41)

$$f(x) = \int_c^x \frac{(x-\xi)^{m-1}}{(m-1)!} B(\xi)\, d\xi \; ; \tag{V,3;43}$$

ce qui (si l'on remplace l'équation différentielle scalaire (V,3;42) par une équation différentielle relative à des fonctions à valeurs dans l'espace de Banach $\vec{F}$)redémontre le théorème 91 du chapitre IV

Application de la théorie des équations différentielles linéaires à la continuité et à la dérivabilité de la solution d'une équation différentielle dépendant d un paramètre.

Théorème 15 bis - Soit $|a,b|$ un intervalle de $\mathbb{R}$, F un espace affine normé complet. Soit Λ un espace topologique, $\vec{L}:(x,y,\lambda) \to L(x,y,\lambda)$, une application continue de $|a,b| \times \Omega \times \Lambda$ dans $\vec{F}$; et soient $\lambda \to x_0(\lambda), \lambda \to y_0(\lambda)$ des applications continues de Λ dans $|a,b|$ et Ω respectivement.

On suppose que L a une dérivée partielle en y , $\frac{\partial L}{\partial y}$, application continue de $|a,b| \times \Omega \times \Lambda$ dans $\mathcal{L}(\vec{F};\vec{F})$. Si, pour la valeur λ_0 du paramètre, l'équation différentielle a une solution f_{λ_0} définie dans l'intervalle fermé $[a',b']$ de $|a,b|$, correspondant aux valeurs initiales $x_0(\lambda_0), y_0(\lambda_0)$, alors il existe un voisinage $\mathcal{V}$ de λ_0 dans Λ tel que, pour tout λ de $\mathcal{V}$, l'équation différentielle admette une solution et une seule f_λ dans $[a', b']$, correspondant aux conditions initiales $x_0(\lambda)$, $y_0(\lambda)$; en outre, $\lambda \to f_\lambda$ est une fonction continue sur $\mathcal{V}$, à valeurs dans $(\Omega^{[a',b']})_{cb}$.

Si A est un ouvert d'un espace affine G et si $\vec{L}$ est de classe C^m $(m \geq 1)$ sur $]a,b[\times \Omega \times \Lambda$, alors, si $\mathcal{V}$ est ouvert, la solution définit une fonction $f : (x,\lambda) \longrightarrow f_\lambda(x)$, de classe C^m sur $[a', b'] \times \mathcal{V}$, à valeurs dans F.

La dérivée $g(x) = \frac{\partial f}{\partial \lambda}(x, \lambda_0)$ est une fonction sur $[a', b']$ à valeurs dans $\mathcal{L}(\vec{G}; \vec{F})$, qui est la solution de l'équation différentielle linéaire, dite équation aux variations:

$$g'(x) = \frac{\partial L}{\partial y}(x, f_{\lambda_0}(x), \lambda_0) \circ g(x) + \frac{\partial L}{\partial \lambda}(x, f_{\lambda_0}(x), \lambda_0) \ , \tag{V,3;44}$$

correspondant à la valeur initiale z_0 pour $x = x_0(\lambda_0)$, avec :

$$z_0 = y'_0(\lambda_0) - \vec{L}(x_0(\lambda_0), y_0(\lambda_0), \lambda_0)\, x'(\lambda_0) \tag{V,3;45}$$

Avant de donner la démonstration, expliquons ces formules : 1°/ $\frac{\partial L}{\partial \lambda}(x, f_{\lambda_0}(x), \lambda_0)$ est dans $\mathcal{L}(\vec{G}; \vec{F})$; $g(x)$ est dans $\mathcal{L}(\vec{G}; \vec{F})$, $\frac{\partial L}{\partial y}(x, f_{\lambda_0}(x), \lambda_0)$ dans $\mathcal{L}(\vec{F}; \vec{F})$, donc $\frac{\partial L}{\partial y}(x, f_{\lambda_0}(x), \lambda_0) \circ g(x)$ est dans $\mathcal{L}(\vec{G}; \vec{F})$; on trouve bien, pour $g'(x)$, un élément de $\mathcal{L}(\vec{G}; \vec{F})$; 2°/ $y'_0(\lambda_0)$ est dans $\mathcal{L}(\vec{G}; \vec{F})$; $\vec{L}(x_0(\lambda_0), y_0(\lambda_0), \lambda_0)$ est dans $\vec{F}$, $x'_0(\lambda_0)$ dans $\mathcal{L}(\vec{G}; \mathbb{R})$ *, donc $\vec{L}(x_0(\lambda_0), y_0(\lambda_0), \lambda_0)\, x'(\lambda_0)$ est dans $\mathcal{L}(\vec{G}; \vec{F})$, c'est 1 application linéaire continue $\vec{\gamma} \longrightarrow (x'_0(\lambda_0) \cdot \vec{\gamma})\, \vec{L}(x_0(\lambda_0), y_0(\lambda_0), \lambda_0)$ de $\vec{G}$ dans $\vec{F}$.

On voit l'importance de ce théorème et de l'équation aux variations; celle-ci est une équation linéaire en $z = g$. Ainsi, si on a pu résoudre l'équation différentielle pour la valeur λ_0 du paramètre, la résolution de l'équation linéaire aux variations (V,3;44), beaucoup plus aisée que celle d une équation arbitraire, donne la dérivée $\frac{\partial f}{\partial \lambda}(x, \lambda_0)$

* Ce renvoi se trouve page 803 -

c'est-b-dire "approximativement" la solution de l'équation pour les valeurs de λ voisines de λ_0, par le développement de Taylor $f_\lambda(x) = f_{\lambda_0}(x) + g(x)\cdot(\overrightarrow{\lambda - \lambda_0}) + \cdots$
Remarquons enfin que les conditions de l'énoncé entraînent celles du théorème d'existence (théorème 1) : la formule des accroissements finis montre que $\vec{L}$, ayant une dérivée $\frac{\partial \vec{L}}{\partial y}$ continue, est localement lipschitzienne en y.

<u>Démonstration</u> - L'équation différentielle, avec les conditions initiales données, pour la valeur λ du paramètre est équivalente à l'équation

(V,3;46)
$$\overrightarrow{\Psi}(f, \lambda) = \vec{0},$$

où $\overrightarrow{\Psi}$ est une application de $\mathcal{U} \times \Lambda = (\Omega^{[a',b']})_{cb;1} \times$ * ** dans $(\vec{F}^{[a',b']})_{cb;1}$; $\vec{h} = \overrightarrow{\Psi}(f, \lambda)$ est la fonction continue bornée sur $[a', b']$, à valeurs dans $\vec{F}$, donnée par

(V,3;47)
$$\overrightarrow{h(x)} = f(x) - \left(y_0(\lambda) + \int_{x_0(\lambda)}^{x} \vec{L}(\xi, f(\xi), \lambda)\, d\xi\right).$$

* $x_0'(\lambda_0)$ est une forme linéaire sur $\vec{F}$, considéré comme espace vectoriel sur le corps des réels. Autrement dit A - =%(i() ne peut être dérivable que par rapport au corps des réels, et alors f elle-même n'est dérivable que par rapport à ce corps, et tous les $\mathcal{L}$ qui interviennent veulent dire ; linéaire par rapport au corps des réels. Toutefois, si x_0 est indépendant de λ, $x_0'(\lambda) \equiv 0$, et il pourra être question partout de dérivabilité par rapport au corps des complexes.

** Nous entendons par $(\Omega^{[a',b']})_{cb;1}$ le sous-espace de $(F^{[a',b']})_{cb;1}$ formé des fonctions f à valeurs dans $\Omega \subset F$, sans restriction sur les valeurs de la dérivée f'.

Nous allons nous ramener au théorème des fonctions implicites. On sait que $\mathcal{U}$ est un ouvert de l'espace affine $(F^{[a',b']})_{cb;1}$ (théorème 37 du chapitre III);

Λ est un ouvert de l'espace affine normé G, donc $\mathcal{U} \times \Lambda$ est un ouvert de l'espace affine normé $F \times G$.

L'application3 est continue de $\mathcal{U} \times \Lambda$ dans l'espace vectoriel normé $(\vec{F}^{[a',b']})_{cb;1}$, on le voit par des majorations que nous ne détaillerons pas, et qui ressemblent comme des soeurs à celle du théorème 5 ou du calcul des variations (théorème 38 du chapitre III).
Ensuite on va voir si $\vec{\Psi}$ admet une dérivée partielle par rapport à f. Appelons $\overrightarrow{\delta f}$ un accroissement de f, comme nous l'avons fait dans le calcul des variations (voir page353). Alors les méthodes employées au théorème 38 du chapitre III montrent que $\vec{\Psi}$ a une dérivée partielle $\frac{\partial \Psi}{\partial f}(f,\lambda)$, donnée par

$$(V,3;48) \quad \begin{cases} \overrightarrow{\delta\Psi} = \dfrac{\partial \Psi}{\partial f}(f,\lambda)\cdot\overrightarrow{\delta f}\,, & \text{, et} \\ \overrightarrow{\delta\Psi}(x) = \overrightarrow{\delta f}(x) - \displaystyle\int_{x_0(\lambda)}^{x} \frac{\partial L}{\partial y}(\xi, f(\xi),\lambda)\cdot\overrightarrow{\delta f}(\xi)\,d\xi\,, \end{cases}$$

et que la fonction dérivée partielle, $\frac{\partial \Psi}{\partial f}$, est continue de $\mathcal{U} \times \Lambda$ dans $\mathcal{L}\left((\vec{F}^{[a',b']})_{cb;1}\,;\,(\vec{F}^{[a',b']})_{cb;1}\right)$.
Comme F est complet, il en est de même de $(F^{[a',b']})_{cb;1}$ (théorème 113 du chapitre IV). Nous pourrons donc appliquer le théorème des fonctions implicites, théorème 25 du chapitre III, si nous prouvons que $\frac{\partial \Psi}{\partial f}(f_{\lambda_0}, \lambda_0)$ est un élément inversible de $\mathcal{L}\left((\vec{F}^{[a',b']})_{cb;1}\,;\,(\vec{F}^{[a',b']})_{cb;1}\right)$.

Pour l'inverser, nous devons montrer que la formule (V,3;48) permet de calculer δf en fonction de $\delta\Psi$, et que $\overrightarrow{\delta\Psi} \longrightarrow \delta f$ est linéaire continue. Or, $\overrightarrow{\delta\Psi}$

étant donnée, $\overrightarrow{\delta f}$ est tout simplement la solution d'une équation différentielle linéaire avec second membre

$$\overrightarrow{\delta f}'(x) = \frac{\partial L}{\partial y}(x, f_{\lambda_0}(x), \lambda_0) \cdot \overrightarrow{\delta f}(x) + \overrightarrow{\delta \Psi}'(x) \tag{V,3;49}$$

correspondant à la condition initiale $\overrightarrow{\delta f}(x_0(\lambda_0)) = \overrightarrow{\delta \Psi}(x_0(\lambda_0))$.
C'est une équation différentielle de la forme (V,3;1), avec $A(x) = \frac{\partial L}{\partial y}(x, f_{\lambda_0}(x), \lambda_0)$.

Soit R la résolvante de cette équation différentielle. Alors (V,3;48) se résout en $\overrightarrow{\delta f}$ par :

$$\overrightarrow{\delta f}(x) = R(x, x_0(\lambda_0)) \cdot \overrightarrow{\delta \Psi}(x_0(\lambda_0)) + \int_{x_0(\lambda_0)}^{x} R(x, \xi) \cdot \overrightarrow{\delta \Psi}'(\xi)\, d\xi \,, \tag{V,3;50}$$

d'après la formule (V,3;27); cela montre bien que, si $\overrightarrow{\delta \Psi}$ est donné dans $\left(\vec{F}^{[a',b']}\right)_{cb;1}$, elle provient d'un $\overrightarrow{\delta f}$ et un seul du même espace, et que l'application $\overrightarrow{\delta \Psi} \longrightarrow \overrightarrow{\delta f}$ est linéaire continue; donc $\frac{\partial \Psi}{\partial f}(f_{\lambda_0}, \lambda_0)$ est bien inversible, et le théorème des fonctions implicites est applicable. Il nous prouve qu'il existe un voisinage $\mathcal{V}$ de λ_0 dans Λ et un voisinages de f_{λ_0} dans $\left(\Omega^{[a',b']}\right)_{cb;1}$, tels que, pour tout λ de $\mathcal{V}$, l'équation (V,3;46) ait une solution et un seule dans $\mathcal{F}$. Alors, pour $\lambda \in \mathcal{V}$, l'équation différentielle donnée admet une solution et une seule f_λ, définie sur $[a', b']$, à valeurs dans Ω, correspondant à la valeur initiale $x_0(\lambda)$, $y_0(\lambda)$, et telle que $f_\lambda \in \mathcal{F}$.

[La phrase "telle que $f_\lambda \in \mathcal{F}$" est un enrichissement du résultat qui la précède, en ce qui concerne l'existence (non seulement il existe une solution, mais elle est dans $\mathcal{F}$); mais c'est un appauvrissement en ce qui concerne l'unicité (il existe une seule solution dans le voisinage $\mathcal{F}$ de f_{λ_0}, mais peut être plusieurs dans $\vec{F}$ $(F^{[a',b']})_{\mathcal{C}b;1}$). Mais de toute façon le théorème 2 d'unicité prouve l'unicité dans $(\vec{F}^{[a',b']})_{m=1}$.

En outre le théorème des fonctions implicites prouve la continuité de l'application $\lambda \longrightarrow f_\lambda$ de $\mathcal{V}$ dans $(\vec{F}^{[a',b']})_{\mathcal{C}b;1}$.

Si maintenant $\vec{L}$ est de classe C^1 sur $|a,b| \times \Omega \times \Lambda$, Λ étant un ouvert d'un espace affine normé $\vec{G}$, et si x_0 et y_0 sont de classe C^1 sur Λ, Ψ est de classe C^1 de $\mathcal{U} \times \Lambda$ dans $\left((F^{[a',b']})_{\mathcal{C}b;1}\right)$, comme on le voit encore par des méthodes analogues. Alors le théorème 28 du chapitre III dit que $\lambda \longrightarrow f_\lambda$ est de classe C^1 de $\mathcal{V}$ dans $(\Omega^{[a',b']})_{\mathcal{C}b;1}$, si $\mathcal{V}$ est ouvert. En outre la dérivée s'obtient par la règle (III,8;24). On différentie $\Psi(f,\lambda) = 0$, ce qui donne, au point f_{λ_0}, λ_0 :

$$\text{(V,3;51)}\qquad \overrightarrow{\delta f}(x) - \int_{x_0(\lambda_0)}^{x} \frac{\partial L}{\partial y}\left(\xi, f_{\lambda_0}(\xi), \lambda_0\right) \cdot \overrightarrow{\delta f}(\xi)\, d\xi$$

$$- y_0'(\lambda_0) \cdot \overrightarrow{\delta\lambda} - \int_{x_0(\lambda_0)}^{x} \left(\frac{\partial L}{\partial \lambda}\left(\xi, f_{\lambda_0}(\xi), \lambda_0\right) \cdot \overrightarrow{\delta\lambda}\right) d\xi$$

$$+ \left(x_0'(\lambda_0) \cdot \overrightarrow{\delta\lambda}\right) \vec{L}\left(x_0(\lambda_0), y_0(\lambda_0), \lambda_0\right) = \vec{0}$$

On obtient donc $\overrightarrow{\delta f}$ en fonction de $\overrightarrow{\delta\lambda}$ comme suit : δf est la solution de l'équation différentielle

$$(V,3;52) \qquad \overrightarrow{\delta f}'(x) = \frac{\partial L}{\partial y}\left(x, f_{\lambda_0}(x), \lambda_0\right) \cdot \overrightarrow{\delta f}(x) + \frac{\partial L}{\partial \lambda}\left(x, f_{\lambda_0}(x), \lambda_0\right) \cdot \overrightarrow{\delta \lambda},$$

correspondant a la condition initiale

$$(V,3;53) \qquad \overrightarrow{\delta f}(x_0(\lambda_0)) = y_0'(\lambda_0) \cdot \overrightarrow{\delta \lambda} - \left(x_0'(\lambda_0) \cdot \overrightarrow{\delta \lambda}\right) \vec{L}(x_0(\lambda_0), y_0(\lambda_0), \lambda_0),$$

ce qui revient exactement à (V,3;44) et (V,3;45).

Si enfin L est de classe C^m, et x_0 et y_0 de classe C^m sur A , $\lambda \longrightarrow f_\lambda$ est de classe C^m, comme on le voit par récurrence sur m.

§ 4 EQUATIONS DIFFERENTIELLES LINEAIRES A COEFFICIENTS CONSTANTS

On appelle ainsi une équation différentielle de la forme :

$$(V,4;1) \qquad \vec{y}' = \mathbf{A} \cdot \vec{y},$$

où **A** est une application linéaire continue fixe de **F** dans lui-même, $\vec{A} \in \mathcal{L}(\vec{F}; \vec{F})$. Ici $|a,b| = \mathbb{R}$; le corps des scalaires sera toujours le corps des complexes $\mathbb{C}$, pour l'espace vectoriel $\vec{F}$.

La solution de l'équation s'obtient à partir de la notion d'exponentielle d'un opérateur.

Considérons la série :

$$(V,4;2) \qquad \exp. A = \sum_{n=0}^{\infty} \frac{A^m}{m!}$$

Elle est absolument convergente dans $\mathcal{L}(\bar{F}; \bar{F})$ (qui est un espace de Banach si $\vec{F}$ lui-même est un espace de Banach).

En effet, on a la majoration $\| A^m \| \leqslant \| A \|^m$ (voir page 131); donc :

$$\| \exp. A \| \leqslant e^{\|A\|} . \tag{V,4;3}$$

La sommé de cette série représente donc une nouvelle application linéaire continue de $\vec{F}$ dans lui-même, qu'on appelle e^A ou exp A. On a exp $(0) = I$, identité.

L'exponentielle d'un opérateur possède les remarquables propriétés suivantes :

1°/ Si A et B sont deux opérateurs qui commutent, c'est-à-dire vérifient $AB = BA$, alors on a la formule :

$$\exp (A + B) = (\exp. A) \circ (\exp. B)^{*} \tag{V,4;4}$$

Il suffit eh effet de multiplier les deux séries absolument convergentes, représentant9 A et exp.B , ce qui donne :

$$(\exp. A)(\exp B) = \sum_{p,q=0}^{\infty} \frac{A^p}{p!} \frac{B^q}{q!} , \tag{V,4;5}$$

d'où, par regroupement des termes, l'expression cherchée :

$$(\exp A)(\exp B) = \sum_{m=0}^{\infty} \frac{1}{m!} \left(\sum_{p+q=m} \frac{m!}{p!\,q!} A^p B^q \right) = \sum_{m=0}^{\infty} \frac{(A+B)^m}{m!} \tag{V,4;6}$$

$$= \exp. (A + B)^{**}$$

* Nous mettons le symbole de composition $\circ$. Mais rappelons qu'on considére $\mathcal{L}(\vec{F};\vec{F})$ comme une algèbre, et qu'on ne le met en général pas; d'ailleurs, nous ne l'avons pas mis dans A^m.

** Le passage du 2ème au 3ème membre (formule du binôme) suppose essentiellement que A et B commutent.

En particulier si A est un opérateur, et λ et μ deux scalaires complexes, on a :

$$\exp.(\lambda A)\,\exp.(\mu A) = \exp.((\lambda+\mu)A)\,. \tag{V,4;7}$$

En prenant $\lambda = -\mu$, on voit que exp A est inversible et d'inverse $\exp$ (-A).

Par ailleurs remarquons que A et $\exp A$ commutent toujours :

$$A \exp A = (\exp A)\,A = \sum_{m=0}^{\infty} \frac{A^{m+1}}{m!} \tag{V,4;8}$$

2°/ A étant un opérateur, considérons l'application $x \longrightarrow \exp(xA)$ de la droite réelle $\mathbb{R}$ dans $\mathcal{L}(\vec{F};\vec{F})$.

Montrons que cette application est continue et même dérivable. Pour cela, d après le corollaire du théorème 111 du chapitre IV, il nous faut d'abord dériver formellement terme à terme, c'est-h-dire écrire :

$$\frac{d}{dx}(\exp xA) = \sum_{m=1}^{\infty} \frac{x^{m-1}}{(m-1)!} A^{m} = \sum_{m=0}^{\infty} \frac{x^{m}}{m!} A^{m+1} = A \exp(xA)\,. \tag{V,4;8 bis}$$

Après quoi, il nous suffit de vérifier que la série obtenue est localement uniformément convergente sur $\mathbb{R}$.

Or c'est Evident, car, pour $|x| \leqslant k$, on a la majoration :

$$\left\| \frac{x^m}{m!} A^{m+1} \right\| \leqslant \frac{k^m}{m!} \|A\|^{m+1}, \text{ et } \sum_{m=0}^{\infty} \frac{k^m}{m!} \|A\|^{m+1} < +\infty \tag{V,4;9}$$

On voit alors que la fonction précédente satisfait à l'équation différentielle :

$$\frac{d}{dx}(\exp(xA)) = A \exp(xA) \quad \text{ou} \quad Y' = A \circ Y = Y \circ A\,. \tag{V,4;10}$$

Comme par ailleurs, pour $x = 0$, on a $\exp(0A) = I$, on voit que la fonction précédente est solution de l'équation différentielle de la résolvante, (V,3;19), correspondant à $\xi = 0$, on a donc, pour cette résolvante, d'après le théorème 12, la formule :

$$R(x, 0) = \exp(xA) . \tag{V,4;11}$$

Comme l'équation différentielle est à coefficients constants, elle est invariante par translation en x , et par conséquent la résolvante relative au point ξ s'exprime à partir de la résolvante relative à **0** , par la formule :

$$R(x, \xi) = R(x - \xi, 0) = \exp((x - \xi)A) . \tag{V,4;12}$$

On en déduit donc que l'unique solution de l'équation différentielle, correspondant à la condition initiale : x_0 , $\vec{y}_0$, est donnée par :

$$\vec{f}(x) = \exp((x - x_0)A) \cdot \vec{f}(x_0) = \exp((x - x_0)A) \cdot \vec{y}_0 \tag{V,4;13}$$

Une fois résolue l'équation sans second membre, on peut alors résoudre une équation différentielle à coefficients constants avec second membre :

$$y' = A . \vec{y} + B(x) . \tag{V,4;14}$$

La solution correspondant à la condition initiale x_0, $\vec{y}_0$, est :

$$\vec{f}(x) = \exp((x - x_0)A) \cdot \vec{y}_0 + \int_{x_0}^{x} \exp((x - \xi)A) \cdot \vec{B}(\xi)\, d\xi \tag{V,4;15}$$

four résumer les résultats précédents :

<u>Théorème 16</u> - <u>La série (V,4;2) définit un élément exp **A** de $\mathcal{L}(\vec{F}; \vec{F})$. On a (V,4;4) si A et B commutent, et (V,4;7). On a 1 équation différentielle (V,4;10). La résolvante de (V,4;1) est donnée par (V,4;12). L unique solution de l'équation différentielle (V,4;14), correspondant à la condition initiale x_0 , $\vec{y}_0$, est donnée par (V,4;15).</u>

On voit que, dans la pratique, toute la difficulté résidera dans la recherche de l'exponentielle d'un opérateur.

Cas particulier où $\vec{F}$ est de dimension finie n. Construction de l'exponentielle d'un opérateur

Même dans ce cas, la recherche de l'exponentielle d'un opérateur n'est pas une chose aisée.

Supposons, pour fixer les idées, qu'on ait choisi dans F une base, alors l'opérateur A peut être représenté par une matrice M , et nous avons à rechercher l'exponentielle de cette matrice, définie par la série (V,4;2). Or cette série ne se prête pas bien à un calcul, puisqu'il nous faudrait calculer les puissances successives de la matrice M. Si l'on n'arrive pas, d'une manière ou d'une autre, à connaître les valeurs propres de cette matrice, le calcul pratique de l'exponentielle est difficile. Si les valeurs propres de A sont des nombres tous distincts, $\lambda_1, \lambda_2, \dots, \lambda_n$, alors on sait que, moyennant le choix d'une base convenable, qui est celle des vecteurs propres correspondant aux valeurs propres précédentes, A s exprime par la matrice diagonale :

$$(V,4;16) \qquad M = \begin{vmatrix} \lambda_1 & 0 & 0 & \dots\dots & 0 \\ 0 & \lambda_2 & 0 & \dots\dots & 0 \\ 0 & 0 & \lambda_3 & \dots\dots & 0 \\ \dots & \dots & \dots & \dots\dots & \dots \\ 0 & 0 & 0 & \dots\dots & h_n \end{vmatrix}$$

Alors on a, pour les puissances successives de la matrice, la formule :

$$(V,4;17) \qquad M^m = \begin{vmatrix} \lambda_1^m & 0 & 0 & \dots\dots & 0 \\ 0 & \lambda_2^m & 0 & \dots\dots & 0 \\ 0 & 0 & \lambda_3^m & \dots\dots & 0 \\ \dots & \dots & \dots & \dots\dots & \dots \\ 0 & 0 & 0 & \dots\dots & \lambda_n^m \end{vmatrix}$$

et par conséquent on en déduit que l'exponentielle de la matrice est elle-même diagonale et est égale à :

$$(V,4;18) \qquad \exp M = \begin{vmatrix} e^{r_1} & 0 & 0 & \cdots & 0 \\ 0 & e^{r_2} & 0 & \cdots & 0 \\ 0 & 0 & e^{r_3} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & e^{r_n} \end{vmatrix}$$

En particulier, si **A** est une homothétie de rapport λ, $A = \lambda I$, alors on a $\exp. A = e^{\lambda} I$.

Examinons maintenant le cas général où les racines caractéristiques de l'opérateur **A** ne sont pas nécessairement toutes distinctes.

Appelons $r_1, r_2, \ldots, r_\ell$, ces racines caractéristiques, la racine r_j étant multiple d'ordre p_j; $\sum_{j=1}^{\ell} p_j = n$. Alors, d'après la théorie de la décomposition de Jordan d'une application linéaire d'un espace vectoriel de dimension finie dans lui-même, on sait qu'à chacun des r_j on peut affecter, d'une manière-biendéterminée et unique, un sous-espace vectoriel F_j de F, de dimension p_j, contenant tous les vecteurs propres pour la valeur propre r_j, * stable par l'opérateur A, et tel que l'opérateur :

$(A - r_j I)^{p_j}$ soit nul dans ce sous-espace.

D'autre part on sait que $\vec{F}$ est la somme directe des sous-espaces vectoriels $\vec{F}_1, \vec{F}_2, \ldots, \vec{F}_\ell$ **. Donnons nous alors une condition initiale pour l'équation différentielle, soit $x_0, \vec{y}_0$.

* Il existe au moins un tel vecteur propre. Les vecteurs propres pour la valeur propre r_j forment un sous espace vectoriel de $\vec{F}_j$.

** Autrement dit, tout vecteur de $\vec{F}$ s'exprime, d'une manière unique, comme somme d'un vecteur de $\vec{F}_1$, d'un vecteur de F_2, ... , et d'un vecteur de $\vec{F}_\ell$.

Le vecteur $\vec{y}_0$ admet "ne décomposition, et "ne seule, $\vec{y}_0 = \sum_{j=1}^{\ell} \vec{u}_j$, où $\vec{u}_j$ appartient à $\vec{F}_j$. On peut alors écrire la solution correspondant à cette condition initiale, sous la forme :

$$\vec{f}(x) = \sum_{j=1}^{\ell} \exp\left(\left(x - x_0\right) A\right) \cdot \vec{u}_j \ . \tag{V,4;19}$$

que nous calculerons comme suit. Pour simplifier, posons $x - x_0 = t$.

$$\exp(tA) \cdot \vec{u}_j = \exp\left(t(A - r_j I)\right) \cdot \left[\exp\left(t r_j I\right) \cdot \vec{u}_j\right] \tag{V,4;20}$$

Le crochet s'écrit $e^{t r_j} \vec{u}_j$; d'autre part, si l'on cherche le développement de : $\exp t\,(A - r_j I)$, on voit que, compte tenu de ce que nous avons dit plus haut sur les puissances successives de l'opérateur $A - r_j I$, ce développement ne contient q" un nombre fini de termes, et peut s'écrire :

$$\exp t\,(A - r_j I) = \sum_{m=0}^{p_j - 1} \frac{t^m}{m!} (A - r_j I)^m \tag{V,4;21}$$

Finalement la solution s'écrit sous la forme :

$$\vec{f}(x) = \sum_{j=1}^{\ell} \left(\sum_{m=0}^{p_j - 1} \frac{t^m}{m!} (A - r_j I)^m \cdot e^{t r_j} \vec{u}_j \right), \ t = x - x_0 \tag{V,4;22}$$

On voit ainsi que toute solution de l'équation différentielle peut s'écrire sous la forme :

$$\vec{f}(x) = \sum_{j=1}^{\ell} e^{r_j x} \vec{P}_j(x) \ , \tag{V,4;23}$$

où chaque terme $e^{r_j x} \vec{P}_j(x)$ est le produit de l'exponentielle $e^{r_j x}$ par un certain polynôme en x , à coefficients appartenant a" sous-espace vectoriel $\vec{F}_j$ de $\vec{F}$, et de degré $\leq p_j$. Si toutes les valeurs propres sont distinctes, chacun des sous-espaces $\vec{F}_j$ a la dimension 1, et on retrouve la solution générale sous la forme classique, vue antérieurement :

(V,4;24) $$\vec{f}(x) = \sum_{j=0}^{n} \vec{c}_j \, e^{r_j x}$$

Dans la pratique, on commencera par chercher les racines caractéristiques de l'opérateur **A** ; pour cela on considérera simplement l'équation :

(V,4;25) $$\text{dét}\,(A - r\,I) = 0\ ,$$

que l'on pourra écrire facilement par le choix d'une base quelconque dans $\vec{F}$. D'autre part, on notera l'ordre de multiplicité p_j de chaque racine caractéristique r_j. On cherchera alors la solution sous la forme indéterminée (V,4;23), **P** étant à coefficients dans $\vec{F}_j$ et il restera à déterminer chaque polynôme P_j. On s'apercevra alors qu'il est possible de choisir arbitrairement dans F_j le coefficient constant de ce polynôme, et qu'ensuite l'équation différentielle permet de calculer tous les coefficients ultérieurs du polynôme, Ceci est d'ailleurs évident, puisque, si l'on appelle $c_{j,0}$ le coefficient constant de P_j, alors la solution exprimée sous la forme (V,4;23) a pour valeur initiale, pour $x = 0$, le vecteur :

(V,4;26) $$\vec{f}(0) = \sum_{j=1}^{l} \vec{c}_{j,0}$$

Cette valeur initiale est arbitraire, donc les composantes $\vec{c}_{j,0}$ de ce vecteur dans les différents sous-espaces $\vec{F}_j$ sont arbitraires, à partir de quoi la solution est unique. On voit donc qu'on ne peut évidemment pas prendre pour $\vec{P}_j$ un polynôme arbitraire. C'est un polynôme de degré $\leq p_j$, <u>ayant un coefficient constant arbitraire.</u>

Ce coefficient constant parcourt un sous-espace vectoriel de dimension p_j, on voit que ce polynôme dépend en réalité de p_j constantes scalaires arbitraires; finalement la solution de l'équation différentielle dépend bien de $n = \sum_{j=1}^{l} p_j$ constantes scalaires arbitraires,

<u>Remarques</u> - Il peut parfaitement arriver que les racines caractéristiques soient multiples, et que cependant les solutions soient toutes de la forme (V,4;24), c'est-à-dire les polynômes réduits à leurs coefficients constants.

(Comme nous le verrons plus loin, corollaire du théorème 18, cette circonstance ne peut pas se produire pour une équation differentielle scalaire d'ordre?).Cela se produit si et seulement si tout le sous-espace vectoriel $\vec{F}_j$ est sous-espace vectoriel propre de l'opérateur A , pour la valeur propre r_j , c'est-à-dire si l'on a, pour tout vecteur $\vec{c}_j$ de F_j , la relation : $A.\vec{c} = r_j\,\vec{c}$, puisque, pour tout $\vec{c}$ de $\vec{F}_j$, on veut avoir $\frac{d}{dx}(\vec{c}\,e^{r_j x}) = \vec{c}\,r_j e^{r_j x} = (A.\vec{c})e^{r_j x}$;

dans ce cas, l'exponentielle de l'opérateur A , dans F_j , se réduit à l'homothétie de rapport e^{r_j}.

Le cas le plus typique est celui où $\vec{F} = \mathbb{R}^n$, et où l'équation differentielle est tout simplement la suivante :

$$y_i' = r y_i \quad j = 1,2,\ldots n. \tag{V,4;27}$$

Dans ce cas, il existe une seule racine caractéristique r, multiple d'ordre n . Or la solution s'exprime sous la forme :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} e^{rx} . \tag{V,4;28}$$

Nous résumerons les résultats précédents par :

<u>Théorème 17 - Etant donnée l'équation différentielle (V,4;1), à chaque valeur propre r_j , d'ordre k_j , de A , on peut faire correspondre un sous-espace vectoriel $\vec{F}_j$ de $\vec{F}$ de dimension k_j , contenant tous les vecteurs propres associés à la valeur propre r_j , et dans lequel $(A - r_j I)^{k_j}$ est nul.? est somme directe des $\vec{F}_j$. La solution générale de l'équation s'écrit sous la forme (V,4;23), où $\vec{P}_j$ est un polynôme de degré $\leq k_j$, à coefficients dans $\vec{F}_j$. On peut choisir arbitrairement dans $\vec{F}_j$ le terme constant de ce polynôme, il est alors déterminé d'une manière unique.</u>

Cas d'une équation différentielle scalaire d'ordre p à coefficients constans

Ecrivons cette équation sous la forme :

(V,4;29) $$Dy = y^{(p)} - \sum_{i=0}^{p-1} A_i\, y^{(i)} = 0^{*} ,$$

où $A_0, A_1, \ldots A_{p-1}$ sont des constantes; en faisant le changement de fonction $z = (y, y', \ldots, y^{(p-1)})$, z devient une fonction à valeurs dans $\mathbb{R}^p$, et elle satisfait à l'équation différentielle (V,4;1), où **A** est la matrice écrite à (V,3;39), avec les A_i constants. On pourrait alors utiliser la théorie générale précédente pour résoudre l'équation différentielle. Il est cependant plus intéressant dans bien des cas, d'utiliser une méthode autonome que nous allons développer.

Soit P un polynôme, à coefficients complexes, par rapport à une variable abstraite Z :

(V,4;30) $$P(Z) = \sum_{i=0}^{p} C_i\, Z^i$$

Alors on note par $P\left(\frac{d}{dx}\right)$ l'opdrateur différentiel :

(V,4;31) $$P\left(\frac{d}{dx}\right) = \sum_{i=0}^{p} C_i \left(\frac{d}{dx}\right)^i .$$

Si P et Q sont deux polynômes, il leur correspond deux opérateurs différentiels $P\left(\frac{d}{dx}\right)$ et $Q\left(\frac{d}{dx}\right)$. Alors au polynôme somme **R** $=$ **P + Q,** correspond l'opérateur différentiel noté $R\left(\frac{d}{dx}\right) = P\left(\frac{d}{dx}\right) + Q\left(\frac{d}{dx}\right)$, somme des opérateurs, et au polynôme **S** $= PQ$, produit des polynômes, correspond l'opérateur différentiel $S\left(\frac{d}{dx}\right) = P\left(\frac{d}{dx}\right) \circ Q\left(\frac{d}{dx}\right)$, composé des opérateurs différentiels, noté aussi $P\left(\frac{d}{dx}\right) Q\left(\frac{d}{dx}\right)$, comme un produit.

Considérons alors l'opérateur différentiel D donné par (V,4;29). Il peut s'écrire $L\left(\frac{d}{dx}\right)$, où L est le polynôme :

* On pourra, si on veut, poser $-A_i = a_i$, et $a_p = 1$, pour'avoir
$$D = \sum_{i=0}^{p} a_i \left(\frac{d}{dx}\right)^i .$$

$$(V,4;32) \qquad L(Z) = Z^{p} - \sum_{i=0}^{p-1} A_i Z^i = \sum_{i=0}^{p} a_i Z^i .$$

Mais, d'après le théorème de d'Alembert (théorème **30** du chapitre II), ce polynôme peut s'exprimer d'une manière unique comme produit de puissances de polynômes du ler degré

$$(V,4;33) \qquad L(Z) = \prod_{j=1}^{\ell} (Z - r_j)^{p_j} .$$

Les r_j, ou racines caractéristiques de l'opérateur différentiel, sont, par définition, les racines du polynôme L, autrement dit les solutions de l'équation :

$$(V,4;34) \qquad L(z) = 0 \;, \quad z \in \mathbb{C} \;,$$

chacune comptée avec son ordre de multiplicité * *.

Alors l'opérateur différentiel D peut lui-même s'écrire sous la forme :

$$(V,4;35) \qquad D = L\left(\frac{d}{dx}\right) = \prod_{j=1}^{\ell} \left(\frac{d}{dx} - r_j\right)^{p_j} .$$

Cherchons à quelle condition une exponentielle e^{rx} est solution de l'équation différentielle (V,4;29). On a :

* Voir note * page **816**.

** Nous donnons ici une définition des racines caractéristiques de D, qui va nous suffire pour résoudre complètement l'équation différentielle; il faudrait démontrer que ce sont bien les racines caractéristiques r_j, avec les ordres de multiplicité p_j, correspondant à la matrice **A** de (V,4;39), de façon à rélier le cas particulier au cas général. Nous laissons au lecteur le soin de le faire; ce n'est pas directement utile, puisque nous traitons le cas des équations différentielles scalaires d'ordre p de manière autonome. Cela résultera d'ailleurs du corollaire du théorème 18.

$$\frac{d}{dx}\,e^{rx} = r e^{rx} \qquad , \text{ d'où}$$

$$\left(\frac{d}{dx}\right)^k e^{rx} = r^k e^{rx} \qquad , \text{ d'où}$$

$$L\left(\frac{d}{dx}\right)\cdot e^{rx} = L(r)\,e^{rx}.$$

Donc il faut et il suffit que r soit racine caractéristique.

Si toutes les racines caractéristiques sont distinctes, on possède ainsi p solutions de l'équation différentielle, à savoir $e^{r_1 x}, e^{r_2 x}, \ldots, e^{r_p x}$. Nous montrerons plus loin (théorème 18) que ces solutions sont nécessairement indépendantes, et alors la solution générale de l'équation différentielle sera :

(V,4;37) $$f(x) = \sum_{j=1}^{p} c_j\, e^{r_j x}.$$

Par contre, s'il existe des racines caractéristiques multiples, alors la résolution de l'équation n'est pas terminée par la connaissance des solutions exponentielles.

Introduisons alors la fonction :

(V,4;38) $$\frac{x^k}{k!}\,e^{rx} \quad ,$$

et cherchons sur cette fonction, l'effet de l'opérateur différentiel $\frac{d}{dx} - s$. On a la formule :

(V,4;39) $$\left(\frac{d}{dx} - s\right)\cdot \frac{x^k}{k!}\,e^{rx} = (\ A\ /\ \frac{x^k}{k!}\,e^{rx}) + \frac{x^{k-1}}{(k-1)!}\,e^{rx}.$$

On en déduit alors, si $r = s$, la formule :

(V,4;40) $$\left(\frac{d}{dx} - r\right)\cdot \frac{x^k}{k!}\,e^{rx} = \frac{x^{k-1}}{(k-1)!}\,e^{rx} \quad ,$$

et plus généralement la formule :

$$\text{(V,4;41)} \quad \left(\frac{d}{dx} - r\right)^q \cdot \frac{x^k}{k!} e^{rx} = \begin{cases} 0, \text{ si } q \geq k+1 \\ \frac{x^{k-q}}{(k-q)!} e^{rx}, \text{ si } q \leq k \end{cases} .$$

si au contraire A $\neq r$, on voit que $\left(\frac{d}{dx} - s\right)^q \cdot P(x) e^{rx}$ est le produit de e^{rx} par un polynôme qui reste du degré de P .

Supposons donc que r_j soit racine multiple d'ordre p_j du polynôme L. Alors l'opérateur différentiel D peut s'écrire sous la forme :

$$\text{(V,4;42)} \quad \begin{cases} L\left(\frac{d}{dx}\right) = L_j\left(\frac{d}{dx}\right) \circ \left(\frac{d}{dx} - r_j\right)^{p_j} , \\ L_j(Z) = \prod_{\nu \neq j} (Z - r_\nu)^{p_\nu} . \end{cases}$$

On voit que si on le fait opérer sur n'importe quelle fonction $\frac{x^k}{k!} e^{r_j x}$, $k \leq p_j - 1$, on obtient nécessairement 0 ; autrement dit ces fonctions, au nombre de p_j , sont toutes des solutions de l'équation différentielle. On en déduit que celle-ci a au moins les p solutions suivantes :

$$\text{(V,4;43)} \quad \frac{x^k}{k!} e^{r_j x} , \quad k = 0,1,2,\ldots,p_j - 1 , \quad j = 1,2,\ldots l .$$

Si donc nous pouvons démontrer que ce sont des solutions indépendantes, comme on sait que l'espace des solutions à la dimension p , on en déduira que la solution générale de l'équation (V,4;29) est la fonction :

$$\text{(V,4;44)} \quad f(x) = \sum_{j=1}^{l} \left(\sum_{k=0}^{p_j - 1} c_{j,k} \frac{x^k}{k!} e^{r_j x} \right) .$$

Théorème 18 - Les r fonctions (V,4;43) sont indépendantes. Autrement dit, une fonction de la forme (V,4;44) ne peut être identiquement nulle que si tous ses coefficients sont nuls.

Première démonstration - Supposons que ce ne soit pas vrai, montrons que nous aboutissons à une contradiction. Si tous les coefficients ne sont pas nuls, soit c_{j,q_j}, un coefficient $\neq 0$, tel que les $c_{j,k}$, $k > q_j$, soient nuls (c'est le coefficient au terme de degré le plus élevé du polynôme associé à $e^{r_j x}$). Considérons l'opérateur différentiel suivant :

$$(V,4;45) \qquad M_j\left(\frac{d}{dx}\right), \text{ où } M_j(Z) = \frac{L(Z)}{(Z-r_j)^{p_j - q_j}} = \prod_{\nu \neq j} (Z-r_\nu)^{p_\nu} (Z-r_j)^{q_j}$$

Agissant sur chaque fonction $\frac{x^k}{k!} e^{r_\nu x}$, où r_ν est une racine caractéristique $\neq r_j$, et $k \leq p_\nu - 1$, il donne 0, parce que $M_j(Z)$ contient en facteur $(Z - r_\nu)^{p_\nu}$ (raisonnement déjà vu à (V,4;42)).

Agissant sur $\frac{x^k}{k!} e^{r_j x}$, $k \leq q_j - 1$, il donne encore 0, puisque $M_j(Z)$ contient en facteur $(Z-r_j)^{q_j}$. Faisons-le agir sur $\frac{x^{q_j}}{(q_j)!} e^{r_j x}$. D'abord $\left(\frac{d}{dx} - r_j\right)^{q_j}$ le transforme en $e^{r_j x}$; ensuite il reste à faire agir sur $e^{r_j x}$ l'opérateur $L_j\left(\frac{d}{dx}\right)$, où :

$$(V,4;45 \text{ bis}) \qquad L_j(Z) = \frac{L(Z)}{(Z-r_j)^{p_j}} = \prod_{\nu \neq j} (Z - r_\nu)^{p_\nu} \quad \text{(formule (V,4;42))},$$

ce qui donne $L_j(r_j)\, e^{r_j x} \neq 0$.

$$(V,4;46) \qquad M_j\left(\frac{d}{dx}\right) \cdot f(x) = c_{j,q_j}\, L_j(r_j)\, e^{r_j x}.$$

Mais, si la combinaison (V,4;44) est nulle, on obtient 0 en faisant agir sur elle n'importe quel opérateur différentiel, on en déduit qu'on a nécessairement $c_{j,q_j} = 0$, d'où la contradiction.

<u>Deuxième démonstration</u> - La fonction $\vec{f}$, définie par (V,4;44), se prolonge en une fonction de classe C^∞ dans le plan complexe, puisqu'il en est ainsi des fonctions exponentielles et des polynômes. D'autre part, elle peut s'exprimer sous forme d'une série de Taylor, suivant les puissances de z, parce qu'il en est ainsi des exponentielles. Or on sait qu'une série de Taylor, suivant les puissances de z, ne peut représenter 0 <u>pour tout x réel</u>, que si tous ses coefficients de Taylor sont nuls, puisque ceux-ci sont les dérivées successives de la fonction pour $x = 0$. Par conséquent cette fonction supposée nulle pour toutes les valeurs réelles $z = x$, est aussi nulle pour toutes les valeurs <u>complexes</u> de z.

Raisonnons, comme précédemment, par l'absurde.

soit c_{j,q_j} un coefficient non nul dans l'expression de f, tel que :

a) $c_{j,k} = 0$ pour $k > q_j$;

b) pour toute racine caractéristique $r_\nu \neq r_j$, on ait $|r_\nu| \leq |r_j|$. Soit $r_j = |r_j| e^{i\theta_j}$.

Posons $z = \rho e^{-i\theta_j}$, et faisons tendre ρ vers $+\infty$. Examinons ce que devient chaque terme de f.

Appelons alors r_j ~~l'une~~ quelconque des racines caractéristiques. de module maximum intervenant effectivement dans l'expression de f, et soit q_j le plus haut ~~degré~~ $\frac{x^k}{k!} e^{r_j x}$ à coefficient $\neq 0$ dans f. Appelons $|r_j|$ son module, θ son argument, autrement dit : $r_j = |r_j| e^{i\theta_j}$, et posons $z = \rho e^{-i\theta_j}$.

Examinons alors, lorsque z tend vers l'infini avec l'argument $-\theta_j$, ce que deviennent les différents termes de la somme (V,4;44).

$$
\text{(V;4;47)}\quad
\left|\begin{aligned}
&\left| c_{j,q_j} \frac{z^{q_j}}{(q_j)!} e^{r_j z} \right| = \left| c_{j,q_j} \right| \frac{\rho^{q_j}}{(q_j)!} e^{|r_j|\rho} \;; \\
&\left| c_{j,k} \frac{z^k}{k!} e^{r_j z} \right| = \left| c_{j,k} \right| \frac{\rho^k}{k!} e^{|r_j|\rho} \;; \\
&\left| c_{j,k} \frac{z^k}{k!} e^{r_\nu z} \right| \leq \left| c_{j,k} \right| \frac{\rho^k}{k!} e^{|r_\nu|\rho} \;, \quad \text{pour } |r_\nu| < |r_j| \;; \\
&\left| c_{j,k} \frac{z^k}{k!} e^{r_\nu z} \right| = \left| c_{j,k} \right| \frac{\rho^k}{k!} e^{|r_j| \cos(\theta_\nu - \theta_j)\rho}, \text{pour } r_\nu = |r_j| e^{i\theta_\nu}.
\end{aligned}\right.
$$

On voit alors que chaque terme a un module infiniment petit devant le premier pour ρ infini, le deuxième parce que $k < q_j$, le troisième parce que $|r_\nu| < |r_j|$, le quatrième parce que $\theta_\nu \neq \theta_j$. Donc $f(\rho e^{-i\theta_j})$ est équivalent, pour ρ infini, à $c_{j,q_j} \frac{z^{q_j}}{(q_j)!} e^{|r_j|\rho}$; alors $f(z) \equiv 0$ entraîne $c_{j,q_j} = 0$, ce qui contredit notre hypothèse, et démontre le théorème, par l'absurde.

Corollaire . Si l'équation différentielle (V,4;29), où $D = L\left(\frac{d}{dx}\right)$, a pour racines caractéristiques (racines de $L(z)$) $r_1, r_2, \ldots, r_\ell$, multiples d'ordres $p_1, p_2, \ldots, p_\ell$ respectivement, la solution générale de l'équation est donnée par (V,4;44).

Equation différentielle scalaire d'ordre p à coefficients constants avec second membre

Soit a résoudre maintenant l'équation :

$$Dy = y^{(p)} - \sum_{i=0}^{p-1} A_i y^{(i)} = B(x) , \tag{V,4;48}$$

où les A_i sont des constantes, et où B est une fonction continue, à valeurs complexes.

D'après ce que nous avons vu page 798, la solution du problème est entièrement connue dès que l'on connait la fonction $\eta = \eta_{p-1}(x, \xi)$ solution de l'équation différentielle, correspondant aux conditions initiales $0, 0, 0, \ldots, 1$, au point5 . Désignons par η_{p-1} la solution correspondant aux conditions initiales $0, 0, 0, \ldots 1$ au point0 . Alors, par translation, celle qui correspond aux mêmes conditions au point ξ n'est autre que $x \to \eta(x-\xi)$

Plutôt que d'utiliser la méthode indiquée page 800, note (*), il peut être commode d'utiliser celle de la résolvante .
Pour cela cherchons les fonctions η_i, solutions de l'équation différentielle correspondant aux conditions initiales $\eta_i^{(q)}(0) = \delta_i^q$, symbole de Kronecker.
Nous allons montrer que l'on a, avec $k = 1, 2, \ldots, p$:

$$\eta_{p-k} = - A_{p-k+1}\,\eta - A_{p-k+2}\,\eta' - \cdots - A_{p-1}\,\eta^{(k-2)} + \eta^{(k-1)} \tag{V,4;49}$$

Si on écrit l'équation différentielle suivant les notations de la note * page 816, cela revient à :

(V,4;49 bis)
$$\begin{aligned}
\eta_{p-1} &= \eta \\
\eta_{p-2} &= \eta' + a_{p-1}\eta \\
\eta_{p-3} &= \eta'' + a_{p-1}\eta' + a_{p-2}\eta \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
\eta_1 &= \eta^{(p-2)} + a_{p-1}\eta^{(p-3)} + \cdots + a_3\eta' + a_2\eta \\
\eta_0 &= \eta^{(p-1)} + a_{p-1}\eta^{(p-2)} + \cdots + a_2\eta' + a_1\eta .^{(*)}
\end{aligned}$$

C'est évident pour $h = 1$, où cela se réduit à $\eta_{p-1} = \eta$ montrons le par récurrence en supposant que ce soit vrai pour la valeur $h - 1$ et en le montrant pour la valeur h. Tout d'abord nous remarquerons que la fonction, définie par la formule (V,4;49), est bien solution de l'équation différentielle, car toute dérivée d'une solution est encore une solution. Il reste à montrer qu'elle satisfait aux conditions initiales voulues, pour $x = 0$. Or manifestement ses dérivées d'ordre $\leq p - h - 1$ sont nulles, car les dérivées de η d'ordre $\leq p - 2$ sont nulles. La dérivée d'ordre $p-h$ est égale à la dérivée d'ordre $p - 1$ de η, c'est-à-dire 1.

La dérivée d'ordre $p - h + 1$ se calcule par :

(V,4;50)
$$\begin{aligned}
\eta_{p-h}^{(p-h+1)}(0) &= -A_{p-1}\eta^{(p-1)}(0) + \eta^{(p)}(0) \\
&= \left(-A_{p-1}\eta + \eta'\right)^{(p-1)}(0) = -\eta_{p-2}^{(p-1)}(0) ;
\end{aligned}$$

* Prenons, par exemple, l'équation $y''' + ay'' + by' + cy = 0$

On a alors

$$\eta_2 = \eta \ , \ \eta_1 = \eta' + a\eta \ , \ \eta_0 = \eta'' + a\eta' + b\eta$$

elle est donc nulle, à cause de la propriété supposée démontrée pour η_{p-2} . Et ainsi de suite.

Arrivons à la dérivée d'ordre $p-2$. On a :

(V,4;51) $$\eta_{p-h}^{(p-2)}(0) = -A_{p-h+2}\,\eta^{(p-1)}(0) - \cdots - A_{p-1}\,\eta^{(p+h-4)}(0) + \eta^{(p+h-3)}(0) = \eta_{p-h+1}^{(p-1)}(0) \;;$$

et ceci est encore nulle, en raison des propriétés de la fonction $\eta_{p-(h-1)}$, supposées démontrées.
Finalement on obtient, pour la dérivée d'ordre $p-1$, la formule :

(V,4;52) $$\eta_{p-h}^{(p-1)}(0) = -A_{p-h+1}\,\eta^{(p-1)}(0) - A_{p-h+2}\,\eta^{(p)}(0) - \cdots - A_{p-1}\,\eta^{(p+h-3)}(0) + \eta^{(p+h-2)}(0).$$

Modifions cette formule en tenant compte de ce que $\eta^{(h-2)}$ est solution de l'équation différentielle; on obtient :

(V,4;53) $$A_{p-h}\,\eta^{(p-2)}(0) + \cdots + A_1\,\eta^{(h-1)}(0) + A_0\,\eta^{(h-2)}(0) = 0 \;,$$

en vertu des propriétés de la fonction η , et la propriété est bien démontrée. Nous énoncerons donc :

Théorème 19 - La solution de l'équation différentielle (V,4;48), correspondant aux conditions initiales $y_0, y_0', y_0'', \ldots, y_0^{(p-1)}$, s'écrit :

(V,4;54) $$f(x) = \sum_{i=0}^{p-1} y_0^{(i)}\,\eta_i(x-x_0) + \int_{x_0}^{x} \eta(x-\xi)\,B(\xi)\,d\xi \;,$$

où η est la solution de l'équation homogène correspondant aux conditions initiales $0, 0, 0, \ldots, 1$ * , au point 0 , et où les η_i sont données par (V,4;49).

Cette formule montre bien comment la seule connaissance de la fonction η , solution de l équation différentielle, correspond aux conditions initiales $0, 0, \ldots, 1$, à l'origine, permet complètement de résoudre l'équation.

* Pour l ordre 1, $p = 1$ il n'y a qu'une condition initiale : $\eta(0) = 1$.

Exemples - Soit à résoudre l'équation différentielle

(V,4;55) $$y'' + \omega^2 y = B(x),$$

correspondant aux conditions initiales y_0, y'_0 au point x_0.

La solution générale de l'équation homogène est donnée par :

(V,4;56) $$a \cos \omega x + b \sin \omega x .$$

La fonction η, correspondant aux conditions initiales $0, 1$, est donnée par :

(V,4;57) $$\eta(x) = \frac{\sin \omega x}{\omega},$$

et par conséquent la solution f de l'équation cherchée, correspondant aux conditions initiales données, s'obtient par :

(V,4;58) $$\begin{cases} \eta_1(x) = \eta(x) = \dfrac{\sin \omega x}{\omega}, \quad \eta_0(x) = \eta'(x) = \cos \omega x, \\ f(x) = y_0 \cos \omega(x - x_0) + y'_0 \dfrac{\sin \omega(x-x_0)}{\omega} + \displaystyle\int_{x_0}^{x} \dfrac{\sin \omega(x-\xi)}{\omega} B(\xi)\, d\xi \end{cases}$$

Théorème 20 (Heaviside). - La solution $\eta = \eta_{(p-1}$ de (V,4;29), correspondant aux conditions initiales $0, 0, \ldots, 0, 1$, à l'origine, est donnée par (V,4;44) où les $c_{j,k}$ sont les coefficients de la décomposition en éléments simples de la fraction rationnelle $\frac{1}{L(Z)}$ (L défini à (V,4;32)) :

(V,4;59) $$\frac{1}{L(Z)} = \sum_{j=1}^{\ell} \left(\sum_{k=0}^{p_j - 1} \frac{c_{j,k}}{(Z - r_j)^{k+1}} \right).$$

Démonstration 1°/ Considérons un développement du type du 2ème membre de (V,4;59), à coefficients $c_{j,k}$ arbitraires. C'est la décomposition en éléments simples d'une fraction rationnelle $\frac{H(Z)}{L(Z)}$, où H est un polynôme de degré $\leq p$.

Les $c_{j,k}$ utilisés précisément dans (V,4;59) sont ceux pour lesquels $H(Z) \equiv 1$; ils sont caractérisés par le fait que le développement limité du 2ème membre, pour $Z = z$ infiniment grand, suivant les puissances de $\frac{1}{z}$, commence par $\frac{1}{z^p}$, autrement dit a des coefficients de $\frac{1}{z^q}$, $q < p$, nuls, et un coefficient de $\frac{1}{z^p}$ égal à 1.

2°/ Le coefficient de $\frac{1}{z^q}$, dans le développement suivant les puissances de $\frac{1}{z}$, pour z infini, de $\left(\frac{1}{z-r}\right)^{k+1}$, est égal à la dérivée d'ordre $q-1$, pour $q = 0$, de la fonction $x \longrightarrow \frac{x^k}{k!} e^{rx}$.

En effet, on a, d'une part :

$$\left(\frac{1}{z-r}\right)^{k+1} = \frac{(-1)^k}{k!} \left(\frac{d}{dz}\right)^k \left(\frac{1}{z-r}\right) \tag{V,4;60}$$

$$= \sum_{q=k+1}^{\infty} \frac{(q-1)!}{k!(q-k-1)!} \frac{r^{q-k-1}}{z^q} ;$$

et d'autre part :

$$\frac{x^k}{k!} e^{rx} = \sum_{n=0}^{\infty} \frac{r^n x^{k+n}}{k!\,n!} = \sum_{q=k+1}^{\infty} \frac{r^{q-k-1} x^{q-1}}{k!(q-k-1)!} \tag{V,4;61}$$ d'où

$$\left[\left(\frac{d^{q-1}}{dx^{q-1}}\right)\left(\frac{x^k}{k!} e^{rx}\right)\right]_{x=0} = \frac{(q-1)!\, r^{q-k-1}}{k!(q-k-1)!} ; \tag{V,4;62}$$

les deux coefficients considérés sont bien égaux.

3°/ Alors le coefficient de $\frac{1}{z^q}$, dans le développement suivant les puissances de $\frac{1}{z}$, pour z infini, du 2ime membre de (V,4;59), est la dérivée d'ordre $q-1$, pour $x = 0$, de l'expression écrite au 2ème membre de (V,4;44), avec les mêmes $c_{j,k}$. Et si nous écrivons que le développement, suivant les puissances de $\frac{1}{z}$, du 2ème membre de (V,4;59), commence par $\frac{1}{z^p}$, comme il a été dit à 1°/, cela revient à dire que la fonction représentée par (V,4;44) satisfait, pour $x = 0$, aux conditions initiales 0, 0, 0, ..., 1 ; c'est bien $\eta_{p-1} = \eta$, et le théorème est démontré.

Exemple - Considérons l'équation différentielle

$$y'' - 5y' + 6 = 0 . \tag{V,4;63}$$

On a ici

$$L(Z) = Z^2 - 5Z + 6 \tag{V,4;64}$$

Alors

$$\frac{1}{L(Z)} = \frac{1}{Z-3} - \frac{1}{Z-2} \tag{V,4;65}$$

On a donc

$$\eta_1(x) = \eta(x) = e^{3x} - e^{2x} . \tag{V,4;66}$$

La démonstration que nous avons donné du théorème d'Heaviside est une pure vérification; elle ne permet pas vraiment de comprendre pourquoi les choses se passent ainsi. Nous le verrons ultérieurement, au calcul symbolique.

Remarque 1°/ Il y a des cas où l'on a intérêt à ramener une équation différentielle d'ordre p a un systeme de p équations différentielles d'ordre 1, qu'on résoudra par la méthode matricielle: dans d'autres cas au contraire. on aura intérêt à ramener un système d'équations différentielles d'ordre 1, à la résolution d'une seule équation différentielle scalaire d'ordre p .

2°/ Dans le cas particulier où le second membre B est une combinaison linéaire de produits d'exponentielles par des polynômes, du type $B(x) = \sum_{\nu} P_{\nu}(x)\, e^{\lambda_\nu x}$, on a vu, en mathématiques spéciales, comment on peut trouver une solution particulière de l'équation différentielle qui soit du même type, éventuellement avec des polynômes de degré plus élevé si certains des λ_ν sont des racines caractéristiques. Pour trouver cette solution particulière, on pourra d'ailleurs s inspirer des remarques qui ont été faites page 818.

Une fois trouvée cette solution particulière, la solution générale de l'équation s'obtiendra en ajoutant la solution générale de l'équation sans second membre.

Solutions bornés des équations différentielles linéaires à coefficients constants

Théorème 21 - Pour que toute solution d'une équation différentielle (V,4;1), où $\vec{F}$ est de dimension finie (resp. d'une equation différentielle (V,4;29)), reste bornée, lorsque x tend vers $+\infty$, il faut et il suffit que, d'une part, toutes les racines caractéristiques aient une partie réelle ≤ 0, et que, d'autre part pour les racines caractéristiques de partie reelle nulle les polynômes $\vec{P}_j$ de la formule (V,4;23) se réduisent nécessairement tous à leur terme constant (resp. que les racines caractéristiques de partie réelle nulle soient toutes simples).

Remarque - D'après ce que nous avons vu page 815, cette dernière propriété signifie simplement que, si r_j est une racine caractéristique multiple d'ordre n_j et de partie réelle nulle, tout le sous-espace vectoriel $\vec{F}_j$, qui lui est affecte, est sous-espace vectoriel propre correspondant à la valeur propre r_j.

Démonstration - Raisonnons par exemple dans le cas de l'équation (V,4;1), avec $\vec{F}$ de dimensinn finie. La solution générale de l'équation est alors de la forme (V,4;23), où les r_j sont les racines caractéristiques, et les $\vec{P}_j$ certains polynômes, qui ne sont pas arbitraires, mais dont les termes constants sont arbitraires. Il est évident que la condition énoncée est suffisante, nous devons montrer qu'elle est nécessaire. Par ailleurs, rappelons que toute dérivée d'une solution de l'équation différentielle (V,4;1) est aussi une solution de cette equation, et, par conséquent, si toute solution de l'équation reste bornée, quand x tend vers $+\infty$, il en est de même de chacune de ses dérivées. Il nous suffira donc de démontrer que, si une fonction $\vec{f}$ est de la forme (V,4;23), et si elle reste bornée lorsque x tend vers $+\infty$, ainsi que chacune de ses dérivées, alors tous les r_j ont une partie réelle ≤ 0, et, pour les r_j de partie réelle nulle, les polynômes $\vec{P}_j$ correspondants se reduisent à leur terme constant. Soit $\vec{c}_{j,q_j}$ le coefficient de plus haut degré de $\vec{P}_j$, développé suivant les $\frac{x^k}{k!}$. Considérons alors l'opérateur différentiel, $M_j\left(\frac{d}{dx}\right)$, défini à (V,4;45), avec $L(Z)=\prod_{j=1}(Z-r_j)^{n_j}$.

Nous avons montré que cet opérateur différentiel transforme nécessairement la fonction $\vec{f}$ en $\vec{c}_{j,q_j}\, e^{r_j x}$. Cette expression doit donc rester bornée lorsque x tend vers $+\infty$. Il en est donc de même de l'exponentielle; or cela ne peut se produire que si la partie réelle de r_j est ≤ 0.

Considérons maintenant l'autre opérateur différentiel $L_j\left(\frac{d}{dx}\right)$, défini à (V,4;45 bis).

D'après ce que nous avons vu page 818, il transforme la fonction $\vec{f}$ en l expression $\vec{Q}_j(x)e^{r_j x}$, où $\vec{Q}_j$ est un polynôme de même degré que $\vec{P}_j$. Si alors la partie réelle de r_j est nulle, $\|\vec{Q}_j(x)e^{r_j x}\| = \|\vec{Q}_j(x)\|$ ne peut rester borné pour3 tendant vers $+\infty$, que si $\vec{Q}_j$ se réduit a son **terme** constant, c'est-b-dire si $\vec{P}_j$ lui-même se réduit à son **terme** constant. Ainsi le théorème est démontré. Remarquons que, dans ce cas, l'équation différentielle possède une propriété supplémentaire de stabilité tout-a-fait remarquable. Si on modifie peu les conditions initiales de l'équation différentielle, alors on modifie peu la solution de l'équation différentielle, et cela non seulement pour un ensemble de valeurs de x borné, mais pour x allant depuis x_0 jusqu'à $+\infty$. Plus précisément, quel que soit $\varepsilon > 0$, il existe $\eta > 0$ tel que l'inégalité : $\|\vec{y}_0 - \vec{z}_0\| \leq \eta$, entraîne, pour les solutions $\vec{f}$ et $\vec{g}$ de l'équation différentielle correspondant aux conditions initiales respectives $\vec{y}_0, \vec{z}_0$, au point x_0, l'inégalité : $\|\vec{f}(x) - \vec{g}(x)\| \leq \varepsilon$, quel que soit $x \geq x_0$. En effet les fonctions $\vec{\eta}_i$ définies f.(V,3;9) possèdent la propriété d'être toutes bornées pour $x \geq 0$; alors d'après **la valeur de la matrice** résolvante $R(x,x_0)$ donnée page 790 . elle est aussi de norme **bornée** pour $x \geq x_0$. Les fonctions $\vec{f}$ et $\vec{g}$ sont données par la formule (V,3;17), avec $x_2 = x$, $x_1 = x_0$, et les valeurs initiales respectives $\vec{y}_1$ et $\vec{z}_0$.

Alors $\|\vec{f}(x) - \vec{g}(x)\| \leq \|R(x,x_0)\| \, \|\vec{y}_0 - \vec{z}_0\|$ ce qui prouve que l'on a bien la propriété précédente, si, $\varepsilon > 0$ étant donné, on prend $\eta = \dfrac{\varepsilon}{\operatorname{Sup}_{x \geq x_0} \|R(x,x_0)\|}$.

Corollaire - Pour que toute solution de (V,4;1), ou $\vec{F}$ est de dimension finie (resp. toute solution de (V,4;29)) reste bornee pour tout x reel, il faut et il suffit que toutes les racines caractéristiques r_j soient purement imaginaires, et que les polynômes $\vec{P}_j$ de (V,4;23) se réduisent tous nécessairement à leur terme constant (resp. que toutes les racines r_j soient simples).

Il suffit de raisonner successivement pour x tendant vers $+\infty$ (en appliquant le théorème) et pour x tendant vers $-\infty$ (en appliquant le théorème après remplacement des r_j par $-r_j$).

Le théorème et son corollaire ont d'importantes applications a la mécanique (petits mouvements d'un système autour d une position d'équilibre stable).

Tirage Verlag Anton Hain, Meisenheim am Glan, Allemagne
Dépôt légal quatrième trimestre 1967
Numéro d'édition 2202
Hermann, éditeurs des sciences et des arts