maîtrise de mathématiques pures

Collection dirigée par J. Dieudonné (Algèbre) et P. Malliavin (Analyse).

calcul différentiel

A. AVEZ

MASSON

CHEZ LE MÊME ÉDITEUR

Ouvrages de la même collection (Maîtrise de mathématiques pures) : voir page 4 de couverture.

Collection Maîtrise de mathématiques appliquées, sous la direction de Ph. CIARLET et J. LIONS :

INTRODUCTION A L'ANALYSE NUMÉRIQUE MATRICIELLE ET A L'OPTIMISATION.
Livre de cours, par Ph. CIARLET. 1982, 280 pages.
Livre d'exercices, par Ph. CIARLET et J. M. THOMAS. 1982, 144 pages.

PROBABILITÉS ET STATISTIQUE.
Tome 1 : Problèmes à temps fixe.
Livre de cours, par D. DACUNHA-CASTELLE et M. DUFLO. 1982, 204 pages.
Livre d'exercices, par D. DACUNHA-CASTELLE et M. DUFLO. 1982, 200 pages.
Tome 2 : Problèmes à temps mobile.
Livre de cours, par D. DACUNHA-CASTELLE et M. DUFLO. 1983.
Livre d'exercices, par D. DACUNHA-CASTELLE et M. DUFLO. 1983.

INTRODUCTION A L'ANALYSE NUMÉRIQUE DES ÉQUATIONS AUX DÉRIVÉES PARTIELLES.
Livre de cours, par P. A. RAVIART et J. M. THOMAS. 1983, 216 pages.
Livre d'exercices, par P. A. RAVIART et J. M. THOMAS. 1983.

ANALYSE FONCTIONNELLE APPLIQUÉE.
Livre de cours, par H. BRÉZIS. 1982, 240 pages.
Livre d'exercices, par H. BRÉZIS et G. TRONEL, 1983.

Autres ouvrages :

PROBLÈMES D'ANALYSE. Agrégation de mathématiques. Années 1970-1980, par J. VAUTHIER. 1981, 264 pages.

MATHÉMATIQUES APPLIQUÉES ET CALCULATRICES PROGRAMMABLES. Notation algébrique directe, T.I. 57, 58, 59, par L. SOLOMON et M. HOCQUEMILLER. 1982, 264 pages.

MATHÉMATIQUES PAR L'INFORMATIQUE INDIVIDUELLE, par H. LEHNING et D. JAKUBOWICZ. **Tome 1.** Le basic, arithmétique, cryptographie, équations. 1982, 148 pages.

MATHÉMATIQUES PAR L'INFORMATIQUE INDIVIDUELLE, par D. JAKUBOWICZ et H. LEHNING. **Tome 2.** Approximation, sommation. 1982, 128 pages.

COURS DE MATHÉMATIQUES, par J. BASS. **Tome 1.** *Fascicule 2.* Calcul différentiel. Intégrales multiples. Fonctions de variable complexe. 1978, 5e édition, 384 pages.

EQUATIONS DIFFÉRENTIELLES. Maîtrise de mathématiques, par M. ROSEAU. 1976, 156 pages.

LES ÉQUATIONS AUX DÉRIVÉES PARTIELLES. En physique et en mécanique des milieux continus, par S. COLOMBO. 1976, 196 pages.

Collection Maîtrise de mathématiques pures

sous la direction de J. DIEUDONNÉ et P. MALLIAVIN
de l'Institut

A. AVEZ

Professeur à l'Université
Pierre et Marie Curie

CALCUL DIFFÉRENTIEL

MASSON Paris New York Barcelone Milan Mexico Sao Paulo 1983

ISBN : 2-225-79079-5

MASSON S.A.	120 bd Saint-Germain, 75280 Paris Cedex 06
MASSON PUBLISHING USA Inc.	133 East 58th Street, New York, N.Y. 10022
TORAY-MASSON S.A.	Balmes 151, Barcelona 8
MASSON ITALIA EDITORI S.p.A.	Via Giovanni Pascoli 55, 20133 Milano
MASSON EDITORES	Dakota 383, Colonia Napoles, Mexico 18 DF
EDITORA MASSON DO BRASIL Ltda	Rua Dr Cesario Motta Jr, 61, 01221 São Paulo S.P.

INTRODUCTION AU COURS D'ANALYSE

L'ANALYSE MATHÉMATIQUE donne un ensemble de règles gouvernant la manipulation des limites et des infiniment petits : règles de changement de variables, règles d'interversion de limites, règles de dérivation sous le signe intégrale, etc. On ne peut toutefois réduire l'Analyse à cette gymnastique formelle sans perdre de vue ses objets principaux et le sens même de sa démarche.

Dès le XVIII^{e} siècle les séries ont été utilisées pour définir des fonctions nouvelles. Dans un langage moderne, l'Analyse démontre des *théorèmes d'existence* en formulant les problèmes dans des *espaces complets convenables*. Lorsqu'un résultat d'existence est précisé par un théorème d'unicité, alors, et seulement alors, la notion de solution approchée a un sens ; les algorithmes numériques de calcul des solutions approchées proviendront souvent de la démarche antérieure de l'Analyste.

L'évolution des systèmes mécaniques est gouvernée par le *principe du minimum d'action*. Plus généralement l'Analyse permet de définir des fonctions remarquables : celles qui réalisent le minimum de fonctionnelles naturelles. Les propriétés de ces *fonctions extrémales* pourront être déduites alors des équations aux variations de la fonctionnelle associée.

Les lois élémentaires de conservation de la Physique ne permettent pas de décrire un phénomène complexe. Toutefois la formulation *infinitésimale* de ces lois peut conduire à des équations aux dérivées partielles. L'Analyse, en établissant l'existence globale des solutions de ces équations, ainsi que leurs propriétés, apportera un outil pour passer de l'infinitésimal au global.

Le calcul des probabilités sur un nombre fini n d'événements, est souvent équivalent à des problèmes de combinatoire. Lorsque n tend vers l'infini, des lois limites simples apparaissent. Là où l'on ne trouvait que le contingent et l'enchevêtrement d'énumérations fastidieuses, le passage à la limite fera apparaître des fonctions régulières justiciables des méthodes de calcul de l'Analyse.

Ces points de vue seront mis en évidence dans ce cours, destiné à des étudiants de licence ou de maîtrise, et qui comportera quatre volumes de 100 à 200 pages chacun :

— Topologie et Analyse fonctionnelle ;
— Intégration, Probabilités, Analyse de Fourier et Analyse Spectrale ;
— Calcul différentiel ;
— Analyse complexe.

Chaque volume sera écrit de telle sorte qu'il puisse être lu de façon indépendante.

P. MALLIAVIN

INTRODUCTION AU COURS D'ALGÈBRE

> L'Algèbre n'est pas vraiment une discipline indépendante, mais un fondement et un outil pour l'ensemble des mathématiques, et son développement rapide dans les dernières années a été en fait suscité et dirigé par les besoins d'autres disciplines mathématiques.
>
> L. KRONECKER (1861),
> *Math. Werke*, vol. V, p. 387.

L'OPINION DE KRONECKER (l'un des plus illustres algébristes de tous les temps) peut paraître en opposition avec le phénomène bien connu de la prépondérance de plus en plus grande de l'Algèbre dans les mathématiques actuelles, ce qu'on a pu appeler l'« algébrisation » de l'Analyse, de la Géométrie et de la Topologie. En réalité, cette prépondérance est due au fait que les algébristes ont su infléchir leurs recherches sous l'influence des parties des mathématiques où elles pouvaient apporter un appui décisif. Un exemple historique typique est l'évolution de l'Algèbre linéaire et multilinéaire, qui, pour devenir un outil fondamental en Analyse fonctionnelle, a dû commencer par se débarrasser du fatras des calculs de déterminants et de matrices qui l'encombraient inutilement au XIXᵉ siècle. De même, on sait que l'Algèbre commutative est née, d'une part avec les démonstrations, par Dedekind et Weber, des théorèmes fondamentaux de la Théorie des nombres et de la Théorie des courbes algébriques, et de l'autre avec les découvertes de Hilbert sortant la Théorie des invariants des interminables calculs où elle s'enlisait. Et son essor à partir de 1920 est concomitant avec l'essor simultané, à partir de la même époque, de la Géométrie algébrique et de la Géométrie analytique, dont elle forme la base.

C'est donc dans l'esprit de Kronecker qu'est rédigé ce Cours d'Algèbre ; il ne comprend pas une seule définition ni un seul résultat d'Algèbre pure qui n'ait une application dans une autre partie des mathématiques, et on a veillé à ce que les étudiants s'en rendent compte dans toute la mesure du possible. Pour le premier volume, consacré à l'Algèbre linéaire et multilinéaire, cela ne posait pas de problème, car il s'agit là de ce que l'on peut appeler le « pain quotidien » de *tout* mathématicien, qu'il s'occupe d'Arithmétique, d'Analyse fonctionnelle, de Géométrie différentielle, de Topologie algébrique ou de Mécanique quantique.

Les deux autres volumes sont divisés en trois chapitres, dont deux, consacrés respectivement à la Théorie des groupes et à la Théorie des nombres algébriques, sont déjà essentiellement des chapitres d'applications de l'Algèbre. Le troisième, qui traite des parties élémentaires de l'Algèbre commutative, a pour domaines principaux d'applications la Théorie des nombres et la Géométrie algébrique. Le niveau plus élevé de cette dernière n'a pas permis d'en inclure une partie appréciable dans le texte ni dans les exercices ; mais on a essayé de signaler à quoi correspondent « géométriquement » de nombreuses notions purement algébriques de cette théorie, lorsque cela n'exigeait pas l'introduction d'un trop grand nombre de notions nouvelles.

J. DIEUDONNÉ.

TABLE DES MATIÈRES

CONTENTS

A la mémoire de Renée Gallai et de Gaston Roux, à Jean Houlle. Trois personnes qui m'ont appris la plupart des choses utiles que je connais. Deux enseignants comme on n'en voit plus guère.

AVANT-PROPOS

Il y a quelques années la commission des enseignements de l'Université Pierre et Marie-Curie élaborait un programme minimum pour l'unité d'enseignement de Calcul différentiel :

Notion de différentielle. Classe C^r et C^∞.

Symétrie de la différentielle seconde.

Formule de Taylor (à plusieurs variables).

Fonctions implicites. Cas du rang constant.

Equations différentielles : théorème d'existence et dépendance des conditions initiales et de paramètres, dans le cas lipschitzien.

Ce livre couvre ce programme, et un peu plus.

Les appendices sont de deux sortes. Ou bien ce sont des rappels de notions étrangères au calcul différentiel proprement dit, et qu'on a inclus afin de rendre le texte autonome, ou bien ce sont des compléments qu'on peut négliger lors d'une première lecture.

Les références sont notées ainsi : « (2.3. chap. 4) » renvoie à la partie (ou à la formule) 3 du paragraphe 2 du chapitre 4. Si le numéro du chapitre est omis, c'est qu'on renvoie au chapitre en cours.

Les références bibliographiques sont indiquées complètement dans le corps du texte si elles ne sont utilisées qu'une fois. Sinon, on mentionne le nom de l'Auteur (plutôt qu'un numéro !), et la liste alphabétique, qui figure à la fin de ce livre, indique l'ouvrage correspondant. Il se trouve qu'il n'y a pas d'ambiguïté : on n'a retenu qu'un seul ouvrage par Auteur.

Outre que ce livre peut servir de support à l'enseignement du certificat de Calcul différentiel, il est conçu comme une introduction à des ouvrages plus avancés tels que celui de R. Abraham et J. Marsden, auquel il doit beaucoup.

Enfin, l'auteur de ces lignes se doit de remercier Monsieur le Professeur Paul Malliavin, qui a su vaincre sa paresse, et sans qui ce livre n'aurait jamais vu le jour. Il remercie aussi Monsieur le Professeur B. El Mabsout, qui a bien voulu relire le manuscrit et prévenir certaines fautes.

L'auteur adresse aussi ses tendres excuses à sa Femme et à ses Fils pour les avoir frustrés d'un temps dû, normalement, à la vie familiale.

INDEX DES NOTATIONS

$\mathbf{R}^+$: ensemble des réels positifs.
$[a, b]$: ensemble des réels t tels que $a \leqslant t \leqslant b$.
$]a, b[$: ensemble des réels t tels que $a < t < b$.
sup : borne supérieure.
max $\{ a, \dots \}$: le plus grand des nombres $a, \dots$
f^{-1} : application inverse (ou réciproque) de la bijection f.
id_E : application identique de E dans E.
$\overline{A}$: fermeture de l'ensemble A.
$\exp(t)$: exponentielle de t.
ch : cosinus hyperbolique.
sh : sinus hyperbolique.
e.v. : espace vectoriel.
E^* : dual de l'e.v. E.
E^c : complexité de l'e.v. réel E.
$\mathscr{L}(E; F)$: ensemble des applications linéaires continues de l'e.v. E dans l'e.v. F.
End (E) : $\mathscr{L}(E; E)$.
GL (E) : groupe des bijections linéaires continues de l'e.v. E dans E.
SL (n) : groupe unimodulaire de $\mathbf{R}^n$.
$E \oplus F$: somme directe des e.v. E et F.
Ker (f) : noyau de l'application linéaire f.
dét (f) : déterminant de l'endomorphisme linéaire f.
tr (f) : trace de l'endomorphisme linéaire f.
t_f : transposée de l'application linéaire f.
$\| x \|_V$: norme dans l'e.v. normé V.
$B(a, r)$ ou $B_r(a)$: boule ouverte de centre a et de rayon r.
$\langle , \rangle$: produit scalaire ou hermitien.
f^* : adjoint de l'application linéaire f.
S^n : sphère usuelle de dimension n.
$\mathrm{O}(n)$: groupe orthogonal de $\mathbf{R}^n$.
$Df(a)$: différentielle de f en a.
Df : application différentielle de f.
f' : dérivée de la fonction f.
$T_a f$: application tangente de f en a.
grad f : gradient de la fonction f.
$D_r f$: différentielle partielle relativement à la r-ième variable.
C^k : classe C^k.
$C^k(U; F)$: ensemble des applications de classe C^k de U dans F.
$\| f \|_{C^r}$: C^r norme de f.
$\mathrm{Diff}^r(E)$: groupe des difféomorphismes de classe C^r de E dans E.
$\varphi_* X$: image du champ de vecteur X par l'application différentielle φ.
$T_m M$: espace tangent en m à la sous-variété M.

1 NOTION DE DIFFÉRENTIELLE

Ce chapitre est consacré à la définition de la différentielle et à ses propriétés élémentaires qui ne font pas intervenir les notions d'espace complet et d'intégration.

Préliminaires

Dans ce chapitre les espaces vectoriels (en abrégé « e.v. ») sont des espaces normés E construits sur le corps $K = \mathbf{R}$ ou $\mathbf{C}$. L'appendice A (en abrégé « A ») rappelle les notions utilisées ici concernant ces espaces. En particulier la norme de E sera notée $\| \ \|_E$, ou, si le contexte évite les ambiguïtés, $\| \ \|$.

Lorsque plusieurs espaces interviennent dans le même énoncé, il est sous-entendu qu'ils sont construits sur le même corps.

1. Différentielle. Exemples

Une fonction $f : \mathbf{R} \to \mathbf{R}$ est différentiable en $a \in \mathbf{R}$ s'il existe un nombre réel $f'(a)$ tel que $\lim\limits_{h \to 0} [f(a + h) - f(a)]/h = f'(a)$. Cette définition n'a pas de sens pour une application $f : \mathbf{R}^n \to \mathbf{R}^m$, $n > 1$ (comment diviser par un vecteur h de $\mathbf{R}^n$?) ; mais on peut la reformuler de façon qu'elle en ait un : $\lim\limits_{h \to 0} |f(a + h) - f(a) - f'(a).h|/|h| = 0$, qui admet la généralisation suivante :

Définition. 1.1. — Soient U un ouvert non vide d'un e.v. normé E, f une application de U dans un e.v. normé F. On dit que f est différentiable en $a \in U$ s'il existe une application linéaire *continue* $L \in \mathscr{L}(E ; F)$ telle que $\lim\limits_{h \to 0} \| f(a + h) - f(a) - L.h \|_F / \| h \|_E = 0$.

Cela suppose $\| h \|$ assez petit pour que $a + h \in U$.

Donnons nous $k \in E$ arbitrairement et un réel t. Puisque U *est un ouvert*, $a + t.k \in U$ si $|t|$ est assez petit. Changeant h en $t.k$ dans la définition, on voit que si L existe elle est unique et qu'elle est déterminée par $L(k) = \lim\limits_{t \to 0} [f(a + t.k) - f(a)]/t$.

C'est donc pour assurer l'unicité de L qu'on suppose U ouvert.

On appelle L la différentielle de f en a et on la note $Df(a)$; c'est un élément de $\mathscr{L}(E ; F)$.

$Df(a).k$ s'appelle la dérivée de f suivant le vecteur k. On a donc

$$Df(a).k = \frac{\mathrm{d}}{\mathrm{d}t} f(a + t.k) \big|_{t=0} , \qquad (1.2)$$

où le membre de droite est le vecteur dérivé usuel de $t \in \mathbf{R} \rightsquigarrow f(a + t.k) \in F$ pour $t = 0$.

Rappelons qu'une fonction à valeurs réelles d'une variable réelle t, définie sur un voisinage de zéro, est dite $\mathrm{o}(t)$ si $\lim\limits_{t \to 0} \mathrm{o}(t)/t = 0$. On a donc $f(a + h) = f(a) + Df(a).h + \mathrm{o}(\| h \|)$, et $f(a) + Df(a).h$ apparaît comme une approximation de f au premier ordre en h près dans le voisinage de a.

Remarques. 1.3. — *a*) La définition de la différentielle dépend des normes de E et F. Remplaçons-les par des normes équivalentes (voir A. 2.6.) ; U demeure un ouvert car les topologies de E et F sont inchangées. On vérifie sans peine que f demeure différentiable et que sa différentielle en a est encore $Df(a)$. Si, en particulier, E et F sont de dimension finie, la différentiabilité ne dépend pas des normes, car elles sont toutes équivalentes (voir A.2.7.), et l'hypothèse de continuité de $Df(a)$ est superflue, car toute application linéaire est continue (voir A.2.7.).

b) On sait (A.4.1.) que si $E = K$, $\mathscr{L}(E; F)$ s'identifie à F par l'isomorphisme canonique $L \in \mathscr{L}(E; F) \rightsquigarrow L.1 \in F$. Ainsi $Df(a)$ s'identifie à $Df(a).1$, qui n'est autre que le vecteur dérivé usuel $f'(a)$ car

$$Df(a).1 = \frac{\mathrm{d}}{\mathrm{d}t} f(a + t)\Big|_{t=0} = f'(a).$$

c) Si $F = K$, $Df(a)$ est une forme linéaire continue. Si, en outre, la norme de E est déduite d'un produit intérieur $\langle , \rangle$, alors $E^* = \mathscr{L}(E; K)$ s'identifie à E par l'isomorphisme canonique $x \in E \rightsquigarrow \langle x, . \rangle \in E^*$ et $Df(a)$ est l'image d'un vecteur grad $f(a)$, appelé le gradient de f en a et caractérisé par $\langle \operatorname{grad} f(a), v \rangle = Df(a).v$ pour tout v de E. Notons que grad $f(a)$ dépend du produit intérieur choisi, tandis que $Df(a)$ n'en dépend pas.

d) *Matrice jacobienne.* Si $E = K^n$, $F = K^m$, la matrice de l'application linéaire $Df(a)$ dans les bases canoniques est une matrice à m lignes et n colonnes. On l'appelle la matrice jacobienne de f en a. Nous apprendrons à en déterminer les éléments.

e) Une application différentiable en a est continue en a car, $Df(a)$ étant continue, $\lim_{h \to 0} f(a + h) = f(a)$. [Réciproque fausse : $t \in R \rightsquigarrow |t|$ pour $t = 0$.]

f) La notion de différentielle s'étend aux espaces affines, et cela a une grande importance en Physique et en Mécanique.

Soient U un ouvert d'un espace affine E dont l'e.v. associé $\vec{E}$ est normé, f une application de U dans un espace affine F dont l'e.v. associé $\vec{F}$ est normé. On dit que f est différentiable en $a \in U$ s'il existe $L \in \mathscr{L}(\vec{E}; \vec{F})$ telle que

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

Via le choix d'origines *fixes* dans E et F cette notion se ramène à celle relative aux e.v., à laquelle nous nous limiterons car cela simplifie les notations (suppression des flèches).

Définitions. 1.4. — Si $f : U \to F$ est continue en chaque point de l'ouvert U de E, on dit que f est de classe C^0, ou simplement C^0, dans U, et il sera commode de poser $D^0 f = f$.

Si f est différentiable en chaque point de U, on dit que f est différentiable dans U. S'il en est ainsi, l'application Df de l'ouvert U dans l'e.v. normé $\mathscr{L}(E; F)$, définie par $x \rightsquigarrow Df(x)$, s'appelle la différentielle de f, et il sera commode de poser $D^1 f = Df$.

Si Df est continue pour les topologies de U et de $\mathscr{L}(E; F)$ définies par leurs normes, on dit que f est continûment différentiable, ou de classe C^1, ou simplement C^1, dans U.

Si A est une partie, pas nécessairement ouverte, de E, on dit que $f : A \to F$ est différentiable (resp. C^1) dans A si f est la restriction à A d'une application différentiable (resp. C^1) d'un ouvert contenant A dans F.

Exemples. 1.5. — *a*) Soient $E = F = \mathbf{R}^2$, normés par $|(x, y)| = |x| + |y|$, et $f : E \to F$ définie par $f(x, y) = (x + y, x.y)$. Si $a = (a_1, a_2)$ et $h = (h_1, h_2)$ on a $f(a + h) = f(a) + (h_1 + h_2, h_1.a_2 + h_2.a_1) + (o, h_1.h_2)$. Puisque $\|(o, h_1.h_2)\| = |h_1.h_2| \leqslant (|h_1| + |h_2|)^2 = \|h\|^2$, alors $(o, h_1.h_2) = o(\|h\|)$. Ainsi f est différentiable en a et $Df(a).h = (h_1 + h_2, h_1.a_2 + h_2.a_1)$; en sorte que sa matrice jacobienne en a est

$$\begin{pmatrix} 1 & 1 \\ a_2 & a_1 \end{pmatrix}$$

b) D'après A.1.5. l'espace $E = C^1([0, 1])$ des fonctions $u : [0, 1] \to \mathbf{R}$ possédant une dérivée continue est normé par $\|u\|_{C^1} = \|u\|_{C^0} + \|u'\|_{C^0}$. L'espace $F = C^0([0, 1])$ des fonctions continues $v : [0, 1] \to \mathbf{R}$ est normé par $\|v\|_{C^0}$. Soit $f : E \to F$ l'application qui associe à

$u \in E$ la fonction $f(u) : t \in [0, 1] \rightsquigarrow u'(t) + t.u^2(t)$. Avec un abus évident de notation, si $u, h \in E$, on a $f(u + h) = f(u) + L.h + t.h^2$, où $L.h = h' + 2\,t.u.h$. Evidemment L est linéaire, et elle est continue car

$$\| L.h \|_{C^0} \leqslant \| h' \|_{C^0} + 2 \| u \|_{C^0}. \| h \|_{C^0} \leqslant [1 + 2 \| u \|_{C^0}]. \| h \|_{C^1}.$$

D'autre part, $\| t.h^2 \|_{C^0} \leqslant (\| h \|_{C^0})^2 \leqslant (\| h \|_{C^1})^2$ montre que $t.h^2 = o(\| h \|_{C^1})$. Par conséquent f est différentiable sur E et $Df(u).h = h' + 2\,t.u.h$.

2. Règles de calcul

Différentielle d'une application linéaire continue. 2.1.

Une application linéaire continue $f \in \mathscr{L}(E; F)$ est différentiable dans E et sa différentielle $Df(a) = f$ pour tout $a \in E$.

En effet $f(a + h) = f(a) + f.h$ [le $o(\| h \|)$ se réduit à zéro].

Différentielle d'une application affine continue. 2.2.

Plus généralement, soit $f : E \to F$ une application affine continue : $f(x) = L.x + b$, où $L \in \mathscr{L}(E; F)$, $b \in F$. Alors f est différentiable dans E et $Df(a) = L$.

Si, en particulier, f est constante sa différentielle est nulle. Nous verrons que la réciproque est vraie si f est définie sur un ouvert *connexe* U. Elle est inexacte si U n'est pas connexe (prendre $E = F = \mathbf{R}$, U = union disjointe de deux intervalles non vides A et B, $f = 0$ sur A et $f = 1$ sur B).

Applications bilinéaires continues. 2.3.

Soient E_1 E_2 et F trois e.v. normés et $b : E_1 \times E_2 \to F$ une application bilinéaire continue (voir A.3.). Alors b est différentiable dans $E_1 \times E_2$ et la valeur en $h = (h_1, h_2) \in E_1 \times E_2$ de sa différentielle en $a = (a_1, a_2)$ est $Db(a).h = b(h_1, a_2) + b(a_1, h_2)$.

PREUVE. — Puisque $b(a + h) = b(a) + b(h_1, a_2) + b(a_1, h_2) + b(h_1, h_2)$ il suffit de montrer que $b(h_1, h_2) = o(\| h \|)$. Mais cela résulte de

$$\| b(h_1, h_2) \| \leqslant \| b \|. \| h_1 \|. \| h_2 \| \leqslant \| b \|. [\| h_1 \| + \| h_2 \|]^2 \leqslant \| b \|. \| h \|^2. \qquad \square$$

Généralisation (la prouver). — Soient $E_1, \dots, E_n$ et F des e.v. normés et $f \in \mathscr{L}(E_1, \dots, E_n; F)$. Alors f est différentiable en tout point $a = (a_1, \dots, a_n)$ et la valeur en $h = (h_1, \dots, h_n)$ de sa différentiable est

$$Df(a).h = \sum_{k=1}^{n} f(a_1, \dots, a_{k-1}, h_k, a_{k+1}, \dots, a_n).$$

Trace d'une application différentiable. 2.4.

Soit $f : E \to F$ une application différentiable entre e.v. normés. Un sous-e.v. E' de E est normé selon $\| x \|_{E'} = \| x \|_E$ si $x \in E'$. Alors la restriction $f|_{E'}$ de f à E' est différentiable et sa différentielle est la restriction à E' de la différentielle de f. La preuve est évidente.

Cas particulier. — Supposons que E soit un espace de Hilbert et que sa norme soit déduite de son produit intérieur $\langle , \rangle$. Si E' est un sous-espace fermé de E, tout x de E admet une projection orthogonale $p.x$ sur E' (théorème de la projection). Supposons que f soit une fonction numérique ($F = K$). Alors le gradient de $f|_{E'}$ est la projection sur E' du gradient de f.

En effet, pour tout $h \in E'$ on a $\langle p.\operatorname{grad} f(a), h \rangle = \langle \operatorname{grad} f(a), h \rangle = Df(a).h = Df|_{E'}(a).h = \langle \operatorname{grad} f|_{E'}(a), h \rangle$.

Linéarité de la différentielle. 2.5.

Soient U un ouvert non vide d'un e.v. normé E, f et g deux applications de U dans un e.v. normé F. Si $k, k' \in K$, définissons $k.f + k'.g : U \to F$ par $(k.f + k'.g)(x) = k.f(x) + k'.g(x)$ pour tout x de U. On vérifie sans peine que si f et g sont différentiables en $a \in U$ il en est de même de $k.f + k'.g$ et que

$$D(k.f + k'.g)(a) = k.Df(a) + k'.Dg(a).$$

L'ensemble des applications différentiables en a (resp. dans U) est donc une e.v. Il en est de même de l'ensemble des applications C^1 de U dans F; on le note $C^1(U, F)$.

Différentielle d'une application composée. 2.6.

Soient E, F et G trois e.v. normés. Soit f une application d'un ouvert U de E dans F, supposée différentiable en $a \in U$. Soit g une application d'un ouvert de F, contenant $f(U)$, dans G. Supposons g différentiable en $b = f(a)$. Alors $g \circ f$ est différentiable en a et

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a),$$

où le membre de droite est le composé des applications linéaires continues $Dg(b)$ et $Df(a)$.

Preuve. — Posons $y = f(a + h)$ et $b = f(a)$. On a par hypothèse

$$g(y) = g(b) + Dg(b).(y - b) + r(y - b),$$

où $\lim\limits_{y \to b} \| r(y - b) \| / \| y - b \| = 0$. En tenant compte de la linéarité de $Dg(b)$ et de l'hypothèse $y - b = Df(a).h + s(h)$, où $\lim\limits_{h \to 0} \| s(h) \| / \| h \| = 0$, on a

$$(g \circ f)(a + h) - (g \circ f)(a) - Dg(b) \circ Df(a).h = A$$

avec $A = Dg(b).s(h) + r(y - b)$.

Montrons que $A = o(\| h \|)$. D'abord $\| Dg(b).s(h) \| \leqslant \| Dg(b) \| . \| s(h) \|$ montre que $Dg(b).s(h) = o(\| h \|)$. Ensuite, étant donné $e > o$, on peut choisir $\| h \|$ assez petit pour que $\| y - b \| = \| Df(a).h + s(h) \| \leqslant (\| Df(a) \| + e). \| h \|$. Ainsi $\| y - b \| / \| h \| \leqslant \| Df(a) \| + e$. D'autre part la continuité de f en a (voir 1.3.*e*.) entraîne $y \to b$ si $h \to o$. Il en résulte $(y - b) = o(\| h \|)$ et l'on a bien $A = o(\| h \|)$. □

Application tangente. — Supposons $f : U \subset E \to F$ différentiable dans U. L'application tangente $Tf : U \times E \to F \times F$ est définie par $Tf(a, h) = (f(a), Df(a).h)$.

Voici son *interprétation géométrique*. Donnons-nous une courbe différentiable de E, d'origine a, c'est-à-dire une application différentiable $c : \mathbf{R} \to E$ telle que $c(o) = a$. On peut interpréter $c(t)$ comme la position d'un point mobile de E à l'instant t. Soit $c'(o) = h$ son vecteur-vitesse à l'origine. D'après le théorème précédent la courbe $f \circ c : \mathbf{R} \to F$, image par f de la courbe c, est différentiable et son vecteur tangent à l'origine est (voir 1.3.*b*.) :

$$(f \circ c)'(o) = D(f \circ c)(o).1 = Df(c(o)) \circ Dc(o).1 = Df(a).c'(o) = Df(a).h.$$

Ainsi Tf associe au couple formé d'un point d'une courbe et du vecteur-vitesse en ce point le couple formé par l'image du point et du vecteur-vitesse de cette image.

Il est aisé de voir que le théorème précédent s'écrit $T(g \circ f) = Tg \circ Tf$; T est donc un foncteur covariant. Il n'en est pas de même de D, car $D(g \circ f) = (Dg \circ f) \circ Df$ et non pas $Dg \circ Df$.

Application dans une somme directe. 2.7. — Soit $F = F_1 \oplus \cdots \oplus F_n$ une somme directe d'e.v. normés (voir A. 1.6.). Désignons par $p_r : F \to F_r$ la projection $(x_1, \dots, x_n) \mapsto x_r$ et par $i_r : F_r \to F$ l'injection, qui fait correspondre à x_r le vecteur dont toutes les composantes sont nulles à l'exclusion de la r-ième qui est égale à x_r.

Une application f d'un ouvert U d'un e.v. normé E dans F s'écrit $x \mapsto (f_1(x), \dots, f_n(x))$, où $f_r(x)$ est la r-ième composante de $f(x)$. En sorte que

(2.8) $$f_r = p_r \circ f, \qquad f = \sum_{r=1}^{n} i_r \circ f_r.$$

Théorème. — ***f est différentiable en a ∈ U si et seulement si $f_1, \ldots, f_n$ le sont. Dans ces conditions $Df(a) = (Df_1(a), \ldots, Df_n(a))$, c'est-à-dire***

$$Df(a) = \sum_{r=1}^{n} i_r \circ Df_r(a) .$$

Il en résulte que f est C^1 dans U si et seulement si $f_1, \ldots, f_n$ le sont.

PREUVE. — Conséquence immédiate du théorème de différentiation des applications composées et des formules (2.8). □

CONSÉQUENCE. — Si $f : x \in \mathbf{R}^n \rightsquigarrow (f_1(x), \ldots, f_m(x)) \in \mathbf{R}^m$ est différentiable en a, la r-ième ligne de sa matrice jacobienne en a (voir A.1.3.d) est la matrice jacobienne en a de $f_r : \mathbf{R}^n \to \mathbf{R}$.

Formule de Leibniz. 2.9. — Soient E, F_1, F_2 et G des e.v. normés, U un ouvert de E, $f : U \to F_1$ et $g : U \to F_2$ des applications différentiables en $a \in U$ (resp. C^1 dans U), $b : F_1 \times F_2 \to G$ une application bilinéaire continue. Définissons $p = b(f, g) : U \to G$ par $x \rightsquigarrow b(f(x), g(x))$. Alors p est différentiable en a (resp. C^1 dans U) et $Dp(a).h = b(Df(a).h, g(a)) + b(f(a), Dg(a).h)$ pour $h \in E$.

PREUVE. — p est l'application composée $x \rightsquigarrow (f(x), g(x)) \rightsquigarrow b(f(x), g(x))$. Le théorème résulte alors du théorème de différentiation des applications composées, de 2.7. et de 2.3. :

$$\begin{aligned} Dp(a).h &= Db(f(a), g(a)) \circ D(f, g)(a).h = \\ &= Db(f(a), g(a)) \circ (Df(a).h, Dg(a).h) = b(Df(a).h, g(a)) + b(f(a), Dg(a).h) . \end{aligned}$$ □

Cas particuliers. — Si $E = K$, (1.3.b.) montre que $Df(a).h$ s'identifie au produit $h.f'(a)$ du scalaire $h \in K$ par le vecteur dérivé $f'(a)$. De même $Dg(a).h = h.g'(a)$. La formule de Leibniz s'écrit donc, en faisant $h = 1$,

$$p'(a) = b(f'(a), g(a)) + b(f(a), g'(a)) .$$

Si $F_1 = F_2 = G = K$ et $b(u, v) = u.v$, on retrouve ainsi la formule qui donne la dérivée d'un produit $f.g$ de fonctions à valeurs numériques : $(f.g)' = f'.g + f.g'$.

Si $F_1 = F_2 = G$ est l'espace $\mathbf{R}^3$, orienté et doté de son produit scalaire usuel, et si b est le produit vectoriel, on obtient $(f \wedge g)'(a) = f'(a) \wedge g(a) + f(a) \wedge g'(a)$. On obtient une formule analogue pour le produit scalaire.

Application définie sur une somme directe. 2.10.

Différentielles partielles. — Supposons que E soit la somme directe $E_1 \oplus \cdots \oplus E_n$ d'e.v. normés et que U soit un ouvert de E. Une application f de U dans un e.v. normé F s'écrit donc $(x_1, \ldots, x_n) \rightsquigarrow f(x_1, \ldots, x_n)$; c'est une fonction de n variables $x_1 \in E_1, \ldots, x_n \in E_n$.

Donnons-nous $a = (a_1, \ldots, a_n) \in E$. L'application

$$x \in E_r \rightsquigarrow f(a_1, \ldots, a_{r-1}, x_r, a_{r+1}, \ldots, a_n) \in F \tag{2.11}$$

[la r-ième composante a_r est remplacée par x_r] est l'application composée $f \circ I_r$, où $I_r : x_r \rightsquigarrow (a_1, \ldots, a_{r-1}, x_r, a_{r+1}, \ldots, a_n)$ est évidemment une application continue de E_r dans la somme directe $E_1 \oplus \cdots \oplus E_n$. Il en résulte que $f \circ I_r$ est définie dans l'ouvert $I_r^{-1}(U)$ de E_r, qui contient a_r.

Si l'application (2.11), c'est-à-dire $f \circ I_r$, est différentiable en a_r, on appelle sa différentielle en a_r la différentielle partielle de f par rapport à la r-ième variable au point a, et on la note $D_r f(a)$; c'est un élément de $\mathscr{L}(E_r ; F)$.

Si $E_1 = \cdots = E_n = K$, $D_r f(a)$ s'appelle, plus classiquement, la dérivée partielle par rapport à la r-ième variable au point a. C'est la dérivée usuelle $g'(a_n)$ de la fonction numérique à valeurs vectorielles

$$x \rightsquigarrow g(x) = f(a_1, \ldots, a_{r-1}, x, a_{r+1}, \ldots, a_n)$$

au point $x = a_r$. On la note aussi $\dfrac{\partial f}{\partial x_n}(a)$.

Par exemple, si $f : \mathbf{R}^2 \to \mathbf{R}$ est définie par $f(x, y) = \sin(x^2 . y)$, alors $D_1 f(a, b)$ est la dérivée en a de $x \rightsquigarrow \sin(x^2 . b)$; soit $2\, a . \cos(a^2 . b)$. De même $D_2 f(a, b)$ est la dérivée en b de $y \rightsquigarrow \sin(a^2 . y)$; soit $a^2 . \cos(a^2 . b)$.

Revenons au cas général et à ses notations.

Théorème. — ***Si f est différentiable en a, alors chacune des différentielles partielles $D_r f(a)$, $r = 1, \ldots, n$, existe et, si $h = (h_1, \ldots, h_n) \in E$, on a***

$$\boldsymbol{Df(a).h = \sum_{r=1}^{n} D_r f(a).h_r\,.}$$

PREUVE. — Si $i_r : E_r \to E$ est l'injection $x_r \rightsquigarrow (0, \ldots, 0, x_r, 0, \ldots, 0)$ déjà considérée en 2.7., on a $I_r(x_r) = a + i_r(x_r - a_r)$. En sorte que I_r est une application affine, donc $DI_r(a_r) = i_r$ (voir 2.2.). La première partie résulte alors du théorème de différentiation des applications composées et

$$D_r f(a) = D(f \circ I_r)(a_r) = Df(a) \circ i_r\,. \tag{2.12}$$

Mais, si $p_r : E \to E_r$ est la projection sur la r-ième composante, déjà considérée en 2.7., on a évidemment $f = \sum_{r=1}^{n} f \circ i_r \circ p_r$. Le théorème de différentiation des applications composées et (2.12) entraînent

$$Df(a) = \sum_{r=1}^{n} Df(a) \circ i_r \circ p_r = \sum_{r=1}^{n} D_r f(a) \circ p_r\,, \quad \text{donc} \quad Df(a).h = \sum_{r=1}^{n} D_r f(a).h_r\,. \quad \square$$

Remarque. — L'existence des dérivées partielles n'entraîne pas la différentiabilité. Par exemple si l'on prend $f : \mathbf{R}^2 \to \mathbf{R}$ définie par

$$f(x, y) = \begin{cases} x.y.(x^2 + y^2)^{-1/2} & \text{si} \quad x^2 + y^2 \neq 0\,, \\ 0 & \text{si} \quad x = y = 0\,, \end{cases}$$

la fonction est nulle pour $x = 0$ (resp. $y = 0$), donc $D_1 f(o, o) = D_2 f(o, o) = 0$. Si f était différentiable en (o, o), sa différentielle en ce point serait donc nulle, d'après le théorème précédent, et l'on aurait $f(h_1, h_2) = o(\| h \|)$, où $\| h \| = (h_1^2 + h_2^2)^{1/2}$. Pourtant, si $h = (3\,t, 4\,t)$, on a $f(3\,t, 4\,t)/\| h \| = 12/25$, qui ne tend pas vers zéro avec t.

Mieux ! La fonction f peut posséder des dérivées partielles en un point sans y être continue : considérer en (o, o) la fonction

$$f(x, y) = \begin{cases} x.y.(x^2 + y^2)^{-1} & \text{si} \quad x^2 + y^2 \neq 0\,, \\ 0 & \text{si} \quad x = y = 0\,. \end{cases}$$

La raison de cette « pathologie » est que l'existence des dérivées partielles en (o, o) garantit l'existence des tangentes à l'origine des courbes $x \rightsquigarrow f(x, o)$ et $y \rightsquigarrow f(o, y)$, qui sont les intersections de la surface d'équation $z = f(x, y)$ avec les plans $y = 0$ et $x = 0$, mais qu'elle ne dit rien sur le comportement de la surface à l'extérieur de ces plans.

Remarque. — Si f est C^1 dans U, alors les différentielles partielles le sont aussi.

En effet (2.12) montre que $D_r f(a) = Df(a) \circ i_r$ dépend continûment de a.

Nous démontrerons ultérieurement une réciproque (chap. 2).

Calcul de la matrice jacobienne. 2.13.

Si $f : x = (x_1, \ldots, x_n) \in \mathbf{R}^n \rightsquigarrow (f_1(x), \ldots, f_m(x)) \in \mathbf{R}^m$ est différentiable dans un ouvert contenant a, l'élément occupant la i-ième ligne et la r-ième colonne de sa matrice jacobienne en a, est $D_r\, f_i(a)$.

Preuve. — D'après 2.7., la i-ième ligne de la matrice jacobienne en a est la matrice jacobienne en a de $f_i : \mathbf{R}^n \to \mathbf{R}$. D'après 2.10., la matrice jacobienne de f_i en a est $(D_1\, f_i(a), \ldots, D_n\, f_i(a))$.

Voici une autre preuve. Si (e_i) est la base canonique de $\mathbf{R}^n$, considérons la courbe différentiable $c : t \in \mathbf{R} \rightsquigarrow a + t.e_r \in \mathbf{R}^n$ et l'application composée $(f \circ c)(t) = f(a + t.e_r) = (f_1(a + t.e_r), \ldots)$.

Appliquons 1.3.*b.* et le théorème de différentiation des applications composées, puis faisons $t = 0$:

$$Df(a).e_r = \frac{\mathrm{d}}{\mathrm{d}t} f(a + t.e_r)\Big|_{t=0} = (D_r\, f_1(a), \ldots, D_r\, f_m(a)) . \qquad \square$$

Voici une application importante dans la pratique.

Théorème du changement de variables. 2.14. — ***Soient $g_1, \ldots, g_m : \mathbf{R}^n \to \mathbf{R}$ des fonctions différentiables en a et $f : \mathbf{R}^m \to \mathbf{R}$ une fonction différentiable en $b = (g_1(a), \ldots, g_m(a))$. Définissons $F : \mathbf{R}^n \to \mathbf{R}$ par $F(x) = f(g_1(x), \ldots, g_m(x))$. Alors $D_i\, F(a) = \sum_{r=1}^{n} D_r f(b).D_i\, g_r(a)$.***

Preuve. — Si $g : \mathbf{R}^n \to \mathbf{R}^m$ est définie par $g(x) = (g_1(x), \ldots, g_m(x))$, alors $F = f \circ g$. Le théorème de différentiation des applications composées donne $DF(a) = Df(g(a)) \circ Dg(a) = Df(b) \circ Dg(a)$. Il suffit d'écrire que la matrice jacobienne de F en a est égale au produit des matrices jacobiennes de f en b et de g en a, matrices dont les expressions sont données par le théorème précédent. $\square$

Exemple. — On se propose de calculer les dérivées partielles de $F : \mathbf{R}^2 \to \mathbf{R}$ définie par $F(x, y) = f(h(x), s(y))$, où f, h et s sont différentiables. On se ramène au théorème précédent en posant $g_1(x, y) = h(x)$ et $g_2(x, y) = s(y)$. Ainsi

$$F(x, y) = f(g_1(x, y), g_2(x, y)) ,$$

et

$$D_1\, F(x, y) = D_1\, f(u).D_1\, g_1(x, y) + D_2\, f(u).D_1\, g_2(x, y) ,$$
$$D_2\, F(x, y) = D_1\, f(u).D_2\, g_1(x, y) + D_2\, f(u).D_2\, g_2(x, y) ,$$

où l'on a posé pour abréger, $u = (g_1(x, y), g_2(x, y))$.

Mais $D_1\, g_1(x, y) = h'(x)$, $D_1\, g_2(x, y) = 0$, $D_2\, g_1(x, y) = 0$ et $D_2\, g_2(x, y) = s'(y)$.

Ainsi $D_1\, F(x, y) = D_1\, f(u).h'(x)$, $D_2\, F(x, y) = D_2\, f(u).s'(y)$.

2 THÉORÈMES DE LA MOYENNE

Ce chapitre est consacré à la généralisation du théorème classique des accroissements finis concernant les fonctions numériques. Entre autres applications, on démontre le théorème fondamental du calcul intégral.

Préliminaires

Soit f une fonction à valeurs réelles définie et continue sur un segment $[a, b]$ et dérivable dans $]a, b[$. Le théorème classique des accroissements finis dit qu'il existe $c \in]a, b[$ tel que $f(b) - f(a) = (b - a).f'(c)$. En termes géométriques, cela signifie qu'il existe un point $(c, f(c))$ du graphe de f où la tangente est parallèle au segment joignant les extrémités $(a, f(a))$ et $(b, f(b))$ de ce graphe.

Si, maintenant, sous les mêmes hypothèses, f est une fonction à valeurs dans $\mathbf{R}^n$, $n \geqslant 2$, la propriété géométrique précédente peut être en défaut : penser à un graphe en « tire-bouchon », où la tangente en un point ne peut être parallèle à un segment joignant deux points du graphe. Exemple : $f : [0, \pi/2] \to \mathbf{R}^2$ défini par $f(t) = (\cos t, \sin t)$.

Interprétons autrement le théorème des accroissements finis. Si $f : [a, b] \to F$ est une application continue et dérivable sur $]a, b[$ dans un e.v. normé F, on peut interpréter $f(t)$ (resp. $f'(t)$) comme la position (resp. la vitesse) à l'instant t d'un point mobile dans F. La vitesse « au compteur » est $\| f'(t) \|$. Supposons qu'un second mobile parte de $f(a)$ en même temps que le premier, qu'il décrive une droite, et que sa vitesse « au compteur » $g'(t)$ soit, à chaque instant, au moins égale à celle du premier. Il est intuitif que ce second mobile s'éloignera du point de départ $f(a)$ plus vite que le premier. Rendons cela rigoureux.

1. Théorème des accroissements finis

Théorème. 1.1. — ***Soient a et b deux réels tels que $a < b$, F un e.v. normé, $f : [a, b] \to F$ et $g : [a, b] \to \mathbf{R}$ deux applications continues sur $[a, b]$ et différentiables sur $]a, b[$. Supposons que $\| f'(t) \| \leqslant g'(t)$ pour $a < t < b$. Alors $\| f(b) - f(a) \| \leqslant g(b) - g(a)$.***

PREUVE. — Nous allons montrer que si u et v sont deux réels tels que $a < u < v < b$, on a $\| f(v) - f(u) \| \leqslant g(v) - g(u)$. Puisque f et g sont continues en a et b, le théorème s'en déduira en faisant tendre u vers a et v vers b.

Supposons le résultat en défaut, alors $\| f(v) - f(u) \| - [g(v) - g(u)] = M > 0$. Partageons $[a_0 = u, b_0 = v]$ en son milieu $m = (u + v)/2$. L'inégalité triangulaire montre que sur l'un des segments $[u, m]$, $[m, v]$ on a

$$\| f(v) - f(m) \| - [g(v) - g(m)] \geqslant M/2$$

ou

$$\| f(m) - f(u) \| - [g(m) - g(u)] \geqslant M/2 .$$

Désignons par $[a_1, b_1]$ celui de ces deux segments où l'inégalité est réalisée. Réitérons le procédé en partageant $[a_1, b_1]$ en son milieu, etc. On obtient une suite de segments $[a_n, b_n]$ tels que $a_0 \leqslant \cdots \leqslant a_n < b_n \leqslant \cdots \leqslant b_0$, $| b_n - a_n | = (b - a)/2^n$ et

$$\| f(b_n) - f(a_n) \| - [g(b_n) - g(a_n)] \geqslant M/2^n .$$

Il en résulte que a_n et b_n convergent vers un même point $w \in]a, b[$ et que

$$\begin{aligned} M/2^n &\leqslant \| f(b_n) - f(w) \| + \| f(w) - f(a_n) \| - [g(b_n) - g(w)] - [g(w) - g(a_n)] \\ &\leqslant \| Df(w).(b_n - w) \| + o(b_n - w) + \| Df(w).(w - a_n) \| + o(w - a_n) \\ &\quad - [g'(w).(b_n - w) + o(b_n - w)] - [g'(w).(w - a_n) + o(w - a_n)] \\ &\leqslant \| Df(w) \| .(| b_n - w | + | w - a_n |) - g'(w).(| b_n - w | + | w - a_n |) + o(b_n - w) + o(w - a_n) \\ &= [\| Df(w) \| - g'(w)].(b_n - a_n) + o(b_n - a_n) . \end{aligned}$$

Divisons les deux membres par $b_n - a_n = (b - a)/2^n$ et faisons tendre n vers ∞. On obtient la contradiction $M/(b - a) \leqslant \| Df(w) \| - g'(w)$. □

L'intuition, qui conduit au théorème précédent, fait deviner que des pauses suivies d'un départ dans une nouvelle direction, sont permises dans le mouvement du mobile $f(t)$. Cela se traduit par un théorème plus fort, dont on trouvera la preuve dans (H. Cartan) ou (J. Dieudonné) :

Théorème. 1.2. — ***Soient $f : [a, b] \to F$, où F est un* e.v. *normé, et $g : [a, b] \to \mathbf{R}$ deux applications continues. Supposons que pour tout $t \in [a, b]$, à l'exclusion peut-être de ceux d'un ensemble dénombrable, f et g soient différentiables et que $\| f'(t) \| \leqslant g'(t)$. Alors $\| f(b) - f(a) \| \leqslant g(b) - g(a)$.***

Corollaire. 1.3. — ***Soit $f : [a, b] \to F$ une application continue dans un* e.v. *normé, possédant une dérivée $f'(t)$ pour tout $t \in]a, b[$. Supposons qu'il existe une constante k telle que $\| f'(t) \| \leqslant k$ pour tout $t \in]a, b[$. Alors $\| f(b) - f(a) \| \leqslant k.(b - a)$.***

PREUVE. — Prendre $g(t) = k.t$ dans 1.1.

Nous allons supposer maintenant que f est définie sur un ouvert U d'un e.v. normé E, qui n'est plus nécessairement $\mathbf{R}$.

Corollaire. 1.4. — ***Si $f : U \to F$ est différentiable dans U et si le segment $[a, b] = \{ (1 - t).a + t.b : 0 \leqslant t \leqslant 1 \}$, d'extrémités $a, b \in U$, est contenu dans U, alors***

$$\| f(b) - f(a) \| \leqslant \sup_{0 \leqslant t \leqslant 1} \| Df[(1 - t).a + t.b] \| . \| b - a \| .$$

PREUVE. — Le théorème de différentiation des applications composées, appliqué à

$$t \in [0, 1] \rightsquigarrow h(t) = f[(1 - t).a + t.b] ,$$

donne $h'(t) = Df[(1 - t).a + t.b].(b - a)$. On a donc $\| h'(t) \| \leqslant \sup_{0 \leqslant t \leqslant 1} \| Df[\ ] \| . \| b - a \|$ et il suffit d'appliquer 1.3., en remplaçant a par 0, b par 1, f par h et k par $\sup \| Df[\ ] \| . \| b - a \|$.

□

Convexité. 1.5. — On dit qu'un sous-ensemble U d'un e.v. est convexe si, quels que soient $a, b \in U$, le segment qui les joint est dans U.

Théorème des accroissements finis pour les convexes. 1.6. — ***Soient U un ouvert convexe d'un* e.v. *normé E, f une application différentiable de U dans un* e.v. *normé F. S'il existe $k \geqslant 0$ tel que $\| Df(u) \| \leqslant k$ pour tout $u \in U$, alors $\| f(b) - f(a) \| \leqslant k . \| b - a \|$, pour tous $a, b \in U$.***

PREUVE. — Conséquence immédiate de 1.4.

Corollaire. 1.7. — ***Soient U un ouvert convexe d'un* e.v. *normé E, f une application différentiable de U dans un* e.v. *normé F. Alors, pour tous $a, b, c \in U$, on a***

$$\| f(b) - f(a) - Df(c).(b - a) \| \leqslant \sup_{u \in U} \| Df(u) - Df(c) \| . \| b - a \| .$$

PREUVE. — Appliquer 1.4. à $u \rightsquigarrow f(u) - Df(c).u$, dont la différentielle est $Df(u) - Df(c)$. □

Passons à des applications des résultats précédents.

2. Réciproque du théorème 2.2. Chap. 1

Théorème. — ***Soient U un ouvert connexe d'un* e.v. *normé E, f une application différentiable de U dans un* e.v. *normé F. Si $Df(u) = 0$ pour tout u de U, alors f est constante.***

PREUVE. — Fixons a dans U et désignons par B l'ensemble des $b \in U$ tels que $f(b) = f(a)$. Puisque f est continue (1.3. chap. 1), B est fermé dans U. D'autre part, tout x de B est centre d'une boule ouverte contenue dans U et de rayon positif. Cette boule est convexe. D'après 1.6., $f(y) = f(x) = f(a)$ en tout point y de cette boule. Donc B est aussi ouvert dans U. Comme B n'est pas vide et que U est connexe, alors $B = U$. □

3. Condition nécessaire et suffisante pour qu'une application soit de classe C^1

Revenons au 2.10. du chapitre 1. Soient U un ouvert de la somme directe $E = E_1 \oplus \cdots \oplus E_n$, d'e.v. normés et f une application de U dans un e.v. normé F. On a vu que, si f est C^1 dans U, alors les différentielles partielles $D_k f : U \to \mathscr{L}(E_k ; F)$ sont C^1 dans U.

Nous allons prouver la réciproque.

Théorème. — ***Avec les notations ci-dessus, supposons que les différentielles partielles existent en tout point $x = (x_1, \ldots, x_n)$ de U et que les applications $D_k f : U \to \mathscr{L}(E_k ; F)$ soient continues en $a \in U$. Alors f est C^1 en a.***

PREUVE. — D'après la formule $Df(a).h = \sum_{k=1}^{n} D_k f(a).h_k$ de (2.10. chap. 1), il suffit de montrer que $Df(a)$ existe. Définissons

$$g(x) = f(x) - f(a) - \sum_{k=1}^{n} D_k f(a).(x_k - a_k).$$

Alors $D_k g(x) = D_k f(x) - D_k f(a)$.

Puisque les différentielles partielles sont continues en a, à tout $\varepsilon > 0$ correspond $r > 0$ tel que $\| D_k g(x) \| \leqslant \varepsilon$, $k = 1, \ldots, n$, si x reste dans la boule ouverte de centre a et de rayon r. Il résulte de 1.6 que

$$\begin{aligned} \| g(x) \| &\leqslant \sum_{k=1}^{n} \| g(x_1, \ldots, x_k, a_{k+1}, \ldots, a_n) - g(x_1, \ldots, x_{k-1}, a_k, \ldots, a_n) \| \\ &\leqslant \sum_{k=1}^{n} \varepsilon . \| x_k - a_k \| \leqslant n.\varepsilon . \| x - a \| . \end{aligned}$$

Comme ε est arbitraire $f(x) = f(a) + \sum_{k=1}^{n} D_k f(a).(x_k - a_k) + o(\| x - a \|)$. □

4. Un critère de convergence uniforme

Théorème. 4.1. — ***Soient U un ouvert connexe d'un* e.v. *normé E et $f_n : U \to F$ une suite d'applications différentiables dans un* espace de Banach *F. Supposons que :* 1) *il existe un point a de U tel que la suite $f_n(a)$ converge ;* 2) *la suite $Df_n : U \to \mathfrak{L}(E ; F)$ converge uniformément sur chaque borné de U vers une application $g : U \to \mathscr{L}(E ; F)$. Alors, pour chaque $x \in U$, la suite $f_n(x)$ converge vers une limite, que l'on notera $f(x)$. Cette convergence est uniforme sur chaque partie bornée convexe de U. Enfin $f = \lim_{n=\infty} f_n$ est différentiable et $Df = g$.***

PREUVE. — Soit B une boule ouverte de centre a et de rayon $r > 0$ contenue dans U. D'après 1.4., pour tout $x \in B$ on a :

$$\| f_p(x) - f_q(x) - [f_p(a) - f_q(a)] \| \leqslant \sup_{u \in B} \| Df_p(u) - Df_q(u) \| . \| x - a \| . \tag{4.2}$$

Puisque la suite $f_n(a)$ converge et que la suite $Df_n(u)$ converge uniformément sur B, la suite $f_n(x)$ est une suite de Cauchy de F. Comme F *est complet*, $\lim_{n=\infty} f_n(x) = f(x)$ existe. Ce raisonnement montre aussi que si f_n converge en un point d'une boule ouverte de U, elle converge uniformément sur cette boule. L'ensemble des points u de U où $f_n(u)$ converge est donc ouvert et fermé dans U. Comme il contient a et que U est connexe, il coïncide avec U. Autrement dit $\lim f_n(u) = f(u)$ existe pour tout $u \in U$.

Désignons encore par B une partie bornée *convexe* de U, de diamètre $d = \sup_{x,y \in B} \| x - y \|$. D'après 1.4. la relation (4.2) est encore valable si a, $x \in B$, et par suite

$$\| f_p(x) - f_q(x) \| \leqslant \| f_p(a) - f_q(a) \| + \sup_B \| \quad \| . d .$$

Cela montre que f_n converge uniformément dans B.

Reprenons l'inégalité (4.2) et faisons tendre p vers $+ \infty$. Puisque $\lim Df_n = g$, on a

$$\| f(x) - f(a) - [f_q(x) - f_q(a)] \| \leqslant \sup_{u \in B} \| g(u) - Df_q(u) \| . \| x - a \| .$$

Puisque la convergence de Df_q est uniforme sur B, à tout $\varepsilon > 0$ correspond un N tel que $q > N$ implique $\sup_B \| g(u) - Df_q(u) \| < \varepsilon$. D'autre part, pour q ainsi choisi, il existe $r' \leqslant r$ tel que $\| x - a \| \leqslant r'$ implique

$$\| f_q(x) - f_q(a) - Df_q(a).(x - a) \| \leqslant \varepsilon . \| x - a \| .$$

L'inégalité triangulaire et les inégalités précédentes montrent alors que

$$\| f(x) - f(a) - g(a).(x - a) \| \leqslant 3. \varepsilon . \| x - a \| .$$

Ainsi f est différentiable en a et $Df(a) = g(a)$. □

On en déduit le théorème suivant :

Théorème. 4.3. — ***Soient U un ouvert connexe d'un* e.v. *normé E, f_n une suite d'applications différentiables de U dans un* espace de Banach *F. Supposons que* 1) *il existe un point a de U tel que la série $\sum f_n(a)$ converge ;* 2) *la série $\sum Df_n$ converge uniformément sur chaque borné de U vers une application $S : U \to \mathscr{L}(E ; F)$. Alors la série $\sum f_n(x)$ converge pour chaque x de U vers une limite, que l'on notera h(x). Cette convergence est uniforme sur chaque partie bornée convexe de U. Enfin, h est différentiable et Dh = S.***

5. Théorème de Sard

Définition. 5.1. — Soit $f : U \to \mathbf{R}^n$ une application différentiable d'un ouvert U de $\mathbf{R}^p$ dans $\mathbf{R}^n$. On dit que $a \in U$ est un *point critique* de f si le rang de l'application linéaire $Df(a)$ est $< n$.

Définition. 5.2. — Munissons $\mathbf{R}^n$ de son produit scalaire usuel. Ce produit scalaire définit la norme et la distance usuelles. Un déplacement de $\mathbf{R}^n$ sera une application affine de $\mathbf{R}^n$ dans $\mathbf{R}^n$ conservant cette distance.

Donnons-nous $a_1, \dots, a_n \in \mathbf{R}$ et $h_1 > 0, \dots, h_n > 0$. Formons

$$P = [a_1, a_1 + h_1] \times \cdots \times [a_n, a_n + h_n] .$$

On appellera pavé de $\mathbf{R}^n$ tout transformé de P par un déplacement. Le volume du pavé est, par définition, $h_1, \dots, h_n$.

Définition. 5.3. — Un sous-ensemble E de $\mathbf{R}^n$ est de mesure nulle si, à tout $\varepsilon > 0$, correspond un recouvrement de E par des pavés dont la somme des volumes est inférieure à ε.

Le lecteur montrera, à titre d'exercice, que la réunion d'une famille dénombrable d'ensembles E_n, $n = 1, 2, \ldots$, de mesure nulle est un ensemble de mesure nulle [recouvrir E_n par des pavés dont la somme des volumes est moindre que $\varepsilon/2^n$].

Théorème de Sard. 5.4. — ***Soit $f : U \to \mathbf{R}^n$ une application de classe C^1 d'un ouvert U de $\mathbf{R}^p$ dans $\mathbf{R}^n$. Alors l'image $f(C)$ de l'ensemble C des points critiques de f est un ensemble de mesure nulle de $\mathbf{R}^n$.***

La preuve est difficile si $p > n$ [voir J. Milnor]. *Nous nous bornerons au cas $p = n$.*

PREUVE. — 1$^{\text{re}}$ *étape.* — Pavons $\mathbf{R}^n$ par les cubes de côté 1 et dont les sommets ont des coordonnées entières; $\mathbf{R}^n$ est la réunion dénombrable de ces cubes. D'après 5.3., il suffit donc de prouver que l'image par f de la partie des points critiques contenue dans l'un d'entre eux est de mesure nulle. Quitte à effectuer une translation, on peut encore supposer que ce cube est le cube unité $I = [0, 1]^n$. Il reste à prouver que la mesure de $f(C \cap I)$ est nulle.

2$^{\text{e}}$ *étape.* — Soit $x \in C \cap I$. Puisque rang $Df(x) < n$, $Df(x)\,\mathbf{R}^n$ est un sous-espace de dimension $\leqslant n - 1$. Il existe donc un hyperplan P, passant par $f(x)$ et dont la direction $\vec{P}$ contient $Df(x)\,\mathbf{R}^n$. Fixons $\varepsilon > 0$ et prenons y dans la boule ouverte $B(x, \varepsilon)$ de centre x et de rayon ε. D'après 1.7. on a

$$f(y) = f(x) + Df(x).(y - x) + b(\varepsilon).\| y - x \|\,,$$

où $b(\varepsilon) \leqslant \sup\limits_{u \in B} \| Df(x) - Df(u) \|$ tend vers zéro avec ε, puisque Df est continue, donc *uniformément continue sur le compact I.* Comme $f(x) + Df(x)(y - x) \in P$, on voit que la distance de $f(y)$ à l'hyperplan P est dominée par $\varepsilon.b(\varepsilon)$. Ainsi $f[B(x, \varepsilon)]$ est situé entre les deux hyperplans parallèles à P et dont la distance à P est $\varepsilon.b(\varepsilon)$.

D'autre part, 1.4. montre que $\| f(y) - f(x) \| \leqslant \sup\limits_{I} \| Df(u) \|.\| y - x \|$. En sorte que $f(y)$ est dans la boule de centre $f(x)$ et de rayon $a.\varepsilon$, où $a = \sup\limits_{I} \| Df(u) \|$.

En résumé $f[B(x, \varepsilon)]$ est situé dans un cylindre droit, dont la base est l'intersection de P avec la boule de centre $f(x)$ et de rayon $a.\varepsilon$, et dont la hauteur est $2\,\varepsilon.b(\varepsilon)$. Il est clair que $f[B(x, \varepsilon)]$ est encore dans un pavé de $\mathbf{R}^n$, dont les côtés parallèles à P sont de longueur $\sqrt{n}.a.b$ et dont le côté perpendiculaire à P est de longueur $2\,\varepsilon.b(\varepsilon)$. Le volume d'un tel pavé est $2.n^{(n-1)/2}.a^{n-1}.\varepsilon^n.b(\varepsilon)$.

3$^{\text{e}}$ *étape.* — Partageons chacun des n côtés du cube unité I en k parties égales. On obtient k^n cubes de côté $1/k$. Chacun d'eux est contenu dans une boule de rayon $\varepsilon = \sqrt{n}/k$ [théorème de Pythagore !]. Certains d'entre eux coupent l'ensemble C des points critiques. Leur image par l'application f est donc contenue, d'après la 2$^{\text{e}}$ étape, dans un pavé de volume

$$2.n^{(n-1)/2}.a^{n-1}.b(\sqrt{n}/k).k^{-n}\,.$$

Il en résulte que $f(C)$ est contenu dans la réunion d'au plus k^n pavés, dont la somme des volumes $2.n^{(n-1)/2}.a^{n-1}.b(\sqrt{n}/k)$ tend vers zéro si $k \to +\infty$. □

Voici une ultime application, de grande importance pratique, du théorème des accroissements finis :

6. Intégration des fonctions réglées et théorème fondamental du calcul intégral

Tous les espaces de ce paragraphe sont des *espaces de Banach.*

Fonctions en escalier. 6.1. — Soient F un espace de Banach et $[a, b]$, $a < b$, un intervalle fermé et borné de $\mathbf{R}$. Découpons $[a, b]$ en n intervalles par des points $a = a_0 < a_1 < \cdots < a_n = b$:

donnons-nous n vecteurs $C_1, \dots, C_n$ de F. La fonction $f: [a, b] \to F$, définie par $f(x) = C_1$ pour $a_0 \leqslant x < a_1, f(x) = C_2$ pour $a_1 \leqslant x < a_2, \dots, f(x) = C_n$ pour $a_{n-1} \leqslant x \leqslant a_n$, s'appelle une fonction en escalier. L'ensemble de ces fonctions sera noté $([a, b], F)$.

La somme de deux fonctions en escalier, le produit par un scalaire d'une fonction en escalier, sont des fonctions en escalier. En sorte que $([a, b], F)$ est une e.v. Prouvons la première assertion. Soient $f, g \in ([a, b], F)$; supposons que $f(x) = C_i$ pour $a_i \leqslant x < a_{i+1}, 0 \leqslant i \leqslant n$, et $g(x) = C'_j$ pour $b_j \leqslant x < b_{j+1}, 0 \leqslant j \leqslant m$. Formons une subdivision de $[a, b]$ en insérant entre a et b les points a_i, b_j. Dans un intervalle de cette subdivision f et g sont constantes, donc aussi $f + g$.

Si $f \in ([a, b], F)$ posons $\| f \|_{C^0} = \sup\limits_{[a,b]} \| f(x) \|$. Il est clair que $\| \ \|_{C^0}$ est une norme : la norme de la convergence uniforme sur $[a, b]$.

Fonctions réglées. 6.2. — L'espace $\mathscr{R} = \mathscr{R}([a, b], F)$ des fonctions réglées est le complété de l'espace normé $([a, b], F]$. Une fonction réglée est donc limite uniforme d'une suite de fonctions en escalier. *Exemple* : une fonction continue $h : [a, b] \to F$ est réglée. En effet, h étant continue sur le compact $[a, b]$, h est uniformément continue. A tout entier $n > 0$ correspond donc $p > 0$ tel que $\| h(x) - h(y) \| \leqslant 1/n$ si $x, y \in [a, b]$ et $| x - y | < p$. Choisissons donc une subdivision $a = a_0 < a_1 < \cdots < a_m = b$ telle que $| a_{i+1} - a_i | < p$ pour $i = 0, 1, \dots, m - 1$, et définissons f_n par $f_n(x) = h(a_1)$ pour $a_0 \leqslant x < a_1$, $f_n(x) = h(a_2)$ pour $a_1 \leqslant x < a_2, \dots, f_n(x) = h(a_m)$ pour $a_{m-1} \leqslant x \leqslant a_m$.
Evidemment $\| h - f_n \|_{C^0} \leqslant 1/n$, et la suite de fonctions en escalier f_n converge uniformément vers h sur $[a, b]$.

Intégrale d'une fonction en escalier. 6.3. — Si $f(x) = C_1$ pour $a_0 \leqslant x < a_1, \dots, f(x) = C_n$ pour $a_{n-1} \leqslant x \leqslant a_n$, l'intégrale de f entre a et b est par définition

$$I(f) = \int_a^b f(x).\mathrm{d}x = \sum_{i=0}^{n-1} (a_{i+1} - a_i).C_{i+1}.$$

On vérifie sans peine que $I : ([a, b], F) \to F$ est linéaire. D'autre part l'inégalité triangulaire dans F implique $\| I(f) \| \leqslant (b - a).\| f \|_{C^0}$. L'application linéaire I est donc continue.

Enfin, si l'on insère un nombre u entre a et b, on voit sans peine que

$$\int_a^b f(x).\mathrm{d}x = \int_a^u f(x).\mathrm{d}x + \int_u^b f(x).\mathrm{d}x.$$

Si l'on convient de poser $\displaystyle\int_u^v = - \int_v^u$ lorsque $v < u$, on en déduit la relation (dite de Chasles)

$$\int_u^v f(x).\mathrm{d}x = \int_u^w f(x).\mathrm{d}x + \int_w^v f(x).\mathrm{d}x \text{ pour tous } u, v, w \in [a, b].$$

Intégrale d'une fonction réglée. 6.4. — Si $f \in \mathscr{R}$ est limite uniforme d'une suite de fonctions en escalier f_n, la suite $I(f_n)$ converge. En effet, I étant linéaire continue, $I(f_n)$ est une suite de Cauchy de F. *Puisque F est complet*, $I(f_n)$ converge (c'est pourquoi les espaces de ce paragraphe sont des espaces de Banach). En fait $\lim\limits_{n=\infty} I(f_n)$ ne dépend pas de la suite utilisée pour approcher f uniformément : si g_p est une autre suite convergeant vers f, on a

$$\| f_n - g_p \|_{C^0} \leqslant \| f_n - f \|_{C^0} + \| f - g_p \|_{C^0},$$

qui tend vers zéro si $n, p \to +\infty$; donc $\| I(f_n) - I(g_p) \| \leqslant (b - a).\| f_n - g_p \|_{C^0}$ tend vers zéro si $n, p \to +\infty$. Il est par conséquent légitime de définir l'intégrale de $f \in \mathscr{R}$ entre a et b par

$$I(f) = \int_a^b f(x).\mathrm{d}x = \lim_{n=\infty} I(f_n).$$

Les propriétés suivantes résultent de 6.3. :

a) $I : \mathscr{R} \to F$ *est linéaire.*

En effet, si f_n (resp. g_n) est une suite de fonctions en escalier convergeant uniformément vers la fonction réglée f (resp. g), alors $f_n + g_n$ converge uniformément vers $f + g$ et

$$I(f+g) = \lim I(f_n + g_n) = \lim [I(f_n) + I(g_n)] = \lim I(f_n) + \lim I(g_n) = I(f) + I(g).$$

De même $I(k.f) = k.I(f)$ si k est un scalaire.

b) $\| I(f) \| \leqslant (b-a). \| f \|_{C^0}$ qui résulte sans peine de l'inégalité analogue pour les fonctions en escalier et de la continuité de $X \rightsquigarrow \| X \|_F$. En particulier I est une application linéaire continue.

c) La relation de Chasles subsiste pour les fonctions réglées (vérification aisée).

d) Enfin, si $L : F \to K$ (corps des scalaires) est une application linéaire continue et si $f \in \mathscr{R}$, alors $L \circ f$ est une fonction réglée et $L[I(f)] = I[L \circ f]$. En effet, si f_n est une suite de fonctions en escalier convergeant uniformément vers f, alors $L(f_n)$ est réglée et converge uniformément vers $L(f)$.

Primitive d'une fonction réglée. 6.5. — Soit g une fonction définie sur un intervalle (a, b) [qu'on ne suppose plus nécessairement borné], à valeurs dans F, et réglée sur tout intervalle fermé borné de (a, b).

Tout d'abord, convenons de poser, comme en 6.3., $\int_u^v g = -\int_v^u g$ si $u, v \in (a, b)$ et si $v < u$.

En particulier $\int_u^u g = 0$.

Soit x_0 un nombre fixe (arbitraire) de (a, b). On appelle primitive de g la fonction G définie en $x \in (a, b)$ par $G(x) = \int_{x_0}^x g(t).\mathrm{d}t$.

Remarquons d'abord que, bien qu'il existe a priori une infinité de primitives (G dépend de x_0), deux primitives G_1 et G_2 de g ne diffèrent que par une constante. En effet, la relation de Chasles entraîne

$$G_2(x) - G_1(x) = \int_{x_1}^x g(t).\mathrm{d}t - \int_{x_2}^x g(t).\mathrm{d}t = \int_{x_1}^{x_2} g(t).\mathrm{d}t.$$

D'autre part, si g est continue, on a :

Théorème. — ***Une primitive G d'une fonction continue g est de classe C^1 et $G' = g$.***

PREUVE. — D'après la relation de Chasles et 6.4. on a

$$G(x+h) - G(x) - g(x).h = \int_x^{x+h} [g(t) - g(x)].\mathrm{d}t;$$

et le théorème résulte alors de

$$\left\| \int_x^{x+h} [g(t) - g(x)].\mathrm{d}t \right\| / |h| \leqslant \sup_{[x, x+h]} \| g(t) - g(x) \|,$$

qui tend vers zéro si $h \to$ o, puisque g est continue. □

Théorème fondamental du calcul intégral. 6.6. — ***Soient E et F deux espaces de Banach, U un ouvert de E et $f : U \to F$ une application de classe C^1. Si $x + ty \in U$ pour tout $t \in [0, 1]$, alors***

$$f(x+y) = f(x) + \int_0^1 Df(x+ty).y.\mathrm{d}t.$$

PREUVE. — Posons $g(t) = f(x + ty)$ pour $0 \leqslant t \leqslant 1$. Si $0 < t < 1$, le théorème de différentiation des applications composées entraîne $g'(t) = Df(x + t.y).y$. Définissons $h(t) = f(x) + \int_0^t Df(x + sy).y.ds$ pour $0 \leqslant t \leqslant 1$. D'après le théorème précédent, on a $h'(t) = Df(x + ty).y$ pour $0 < t < 1$. Par suite $g'(t) - h'(t) = 0$ pour $0 < t < 1$ et le théorème du paragraphe 2 entraîne $g - h =$ constante pour $0 < t < 1$. Relation qui vaut encore pour $0 \leqslant t \leqslant 1$, puisque g et h sont continues. Mais $g(o) = h(o) = f(x)$, donc $g(1) = h(1)$; ce qui prouve le théorème. □

3 NOTION DE DIFFÉOMORPHISME RÉSOLUTION D'ÉQUATIONS

Ce chapitre est consacré à la résolution d'équation $f(x) = 0$. Puisque l'existence de solutions est obtenue grâce à la convergence d'approximations successives, le cadre est celui des espaces complets. C'est pourquoi les espaces de ce chapitre sont tous des espaces de Banach.

Préliminaires

Soit $f : \mathbf{R} \to \mathbf{R}$ une application continûment différentiable, telle que $f'(a) \neq 0$. Il existe donc un intervalle ouvert I, contenant a, où f' garde un signe constant (positif, pour fixer les idées). Ainsi f est croissante sur I et réalise une bijection de I sur l'intervalle ouvert $J = f(I)$.

Si y est « assez voisin » de $f(a)$, c'est-à-dire si $y \in J$, l'équation $f(x) = y$ a donc une solution $x = f^{-1}(y)$. D'autre part, on sait que $f^{-1} : J \to I$ est continûment différentiable et que $(f^{-1})'(y) = [f'(x)]^{-1}$. La solution $x = f^{-1}(y)$ est donc une fonction C^1 de y.

C'est ce résultat, classique, que nous allons généraliser.

1. Difféomorphismes

Définition. 1.1. — Soient U un ouvert d'un e.v. normé E, V un ouvert d'un e.v. normé F. On dit qu'une application $f : U \to V$ est un difféomorphisme de U sur V si f est une bijection et si f et l'application inverse f^{-1} sont de classe C^1.

EXEMPLES. — 1) $x \leadsto \operatorname{tg} x$ est un difféomorphisme de $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ sur $\mathbf{R}$. 2) $z \leadsto z^2$ est un difféomorphisme du demi-plan $\{ z \in \mathbf{C} : \text{partie imaginaire } z > 0 \}$ sur $\mathbf{C} - \mathbf{R}^+$.

Remarque. 1.2. — Un difféomorphisme $f : U \to V$ est un homéomorphisme, car f et f^{-1} sont différentiables, donc continues (1.3. chap. 1). Mais un homéomorphisme de classe C^1 n'est pas nécessairement un difféomorphisme, car f^{-1} peut ne pas être différentiable. Ainsi $x \leadsto x^3$ est un homéomorphisme de classe C^1 de $\mathbf{R}$ sur $\mathbf{R}$. Toutefois $f^{-1} : y \leadsto y^{1/3}$ n'est pas dérivable à l'origine.

Remarque. 1.3. — Si U n'est pas vide et si $f : U \to V$ est un difféomorphisme, alors E et F sont isomorphes. En particulier, leurs dimensions sont égales, si elles sont finies.

En effet, le théorème de différentiation des applications composées, appliqué à $f^{-1} \circ f = \mathrm{id}_E$, entraîne $Df^{-1}(f(u)) \circ Df(u) = \mathrm{id}_E$, si $u \in U$. Ainsi $Df(u) \in \mathscr{L}(E; F)$ possède un inverse $Df^{-1}(f(u)) \in \mathscr{L}(F; E)$.

Définition et remarque. 1.4. — On dit que $f : U \to V$ est étale et de classe C^1 si f est de classe C^1 et si $Df(u) \in \mathscr{L}(E; F)$ est un isomorphisme de E sur F pour tout $u \in U$.

Un difféomorphisme de classe C^1 est donc étale. La réciproque est inexacte : $f : (\rho, \theta) \in \mathbf{R}^+ \times \mathbf{R} \to (\rho.\cos\theta, \rho.\sin\theta) \in \mathbf{R}^2 - \{ o \}$ est étale, car

$$\text{dét}\, Df(\rho, \theta) = \text{dét}\begin{pmatrix} \cos\theta & -\rho.\sin\theta \\ \sin\theta & \rho.\cos\theta \end{pmatrix} = \rho \neq 0\,.$$

Toutefois f n'est pas un difféomorphisme, car elle n'est pas injective. La raison en est que l'ouvert $\mathbf{R}^+ \times \mathbf{R}$ est « trop grand »; en lui substituant l'ouvert $U = \mathbf{R}^+ \times \{ 0 < \theta < 2\pi \}$, l'application f devient un difféomorphisme de U sur $f(U)$.

Cela conduit au résultat fondamental suivant :

2. Enoncé du théorème d'inversion locale

Soient U un ouvert non vide d'un espace de Banach E, f une application de classe C^1 de U dans un espace de Banach F. Si, en un point $a \in U$, $Df(a)$ est un isomorphisme de E sur F, alors il existe un voisinage ouvert I de a et un voisinage ouvert J de $f(a)$ tels que $f|_I : I \to J$ soit un difféomorphisme de classe C^1. De plus, la différentielle au point $y = f^{-1}(x) \in J$ de l'application inverse f^{-1} est donnée par $Df^{-1}(y) = [Df(x)]^{-1}$.

Remarque. 2.1. — Posons $b = f(a)$. Le théorème précédent affirme que l'équation $f(x) = y$ admet une solution x unique, pourvu que y soit choisi assez voisin de b et que x soit cherché assez voisin de a.

3. Cas de la dimension finie

Nous allons démontrer le théorème d'inversion locale lorsque la dimension de E et de F est finie.

1re *étape.* — Avec les notations du théorème, observons qu'on peut supposer $a = 0, f(a) = 0$, $E = F$ et $Df(a) = \mathrm{id}_E$: il suffit de remplacer $f(x)$ par $h(x) = [Df(a)]^{-1}.[f(a + x) - f(a)]$.

2e *étape.* — Puisque Df est continue, à $\delta \in]0, 1[$ correspond $\varepsilon > 0$ tel que $\| u \| < \varepsilon$ implique $\| Df(u) - Df(\mathrm{o}) \| < \delta$. Prenons x et $x + y$ dans la boule ouverte $B(\mathrm{o}, \varepsilon)$ de centre o et de rayon ε, et supposons $f(x) = f(x + y)$. Puisque $Df(\mathrm{o}) = \mathrm{id}_E$, le théorème fondamental du calcul intégral (chap. 2) implique

$$(3.1) \quad 0 = \| f(x + y) - f(x) \| = \left\| \int_0^1 Df(x + t.y).y.\mathrm{d}t \right\|$$
$$= \left\| y + \int_0^1 [Df(x + t.y) - Df(\mathrm{o})].y.\mathrm{d}t \right\| \geqslant \| y \| - \delta.\| y \| = (1 - \delta).\| y \| .$$

Ainsi $y = 0$ et f est injective sur $B(\mathrm{o}, \varepsilon)$.

Montrons maintenant que l'image par f de la boule $B(\mathrm{o}, \varepsilon/2)$ contient une boule $B(\mathrm{o}, r)$, pourvu que $o < r < \varepsilon.(1 - \delta)/4$.

3e *étape.* — Donnons-nous $z \in B(\mathrm{o}, r)$. La boule fermée $\overline{B}(\mathrm{o}, \varepsilon/2)$ de l'espace E de *dimension finie* est compacte. La fonction continue $x \in \overline{B}(\mathrm{o}, \varepsilon/2) \rightsquigarrow \| f(x) - z \|$ atteint donc sa borne inférieure en un point x_0.

Je dis que x_0 est un point intérieur de la boule. Si $\| x \| = \varepsilon/2$, le théorème fondamental du calcul intégral et $Df(\mathrm{o}) = \mathrm{id}_E$ entraînent

$$\| f(x) \| = \left\| x + \int_0^1 [Df(t.x) - Df(\mathrm{o})].x.\mathrm{d}t \right\| \geqslant \| x \| - \delta.\| x \|$$
$$= (1 - \delta).\| x \| = \varepsilon.(1 - \delta)/2 > 2\,r .$$

Ainsi $\| f(x) - z \| \geqslant \| f(x) \| - \| z \| > r > \| z \| = \| f(\mathrm{o}) - z \| \geqslant \| f(x_0) - z \|$. La borne inférieure ne peut donc être atteinte sur la frontière de $\overline{B}(\mathrm{o}, \varepsilon/2)$.

4e *étape.* — Montrons que $f(x_0) = z$. Posons $y = k.[f(x_0) - z]$, où $k < \mathrm{o}$ et $| k |$ assez petit pour que $\| x_0 + y \| < \varepsilon/2$ [c'est possible, car $x_0 \in B(o, \varepsilon/2)$]. Le théorème fondamental du calcul intégral entraîne

$$\| f(x_0 + y) - f(x_0) - y \| = \left\| \int_0^1 [Df(x_0 + t.y) - Df(o)].y.dt \right\| \leqslant \delta . \| y \|$$

et par suite

$$\begin{aligned} \| f(x_0 + y) - z \| &\leqslant \| f(x_0 + y) - f(x_0) - y \| + \| f(x_0) - z + y \| \\ &\leqslant \delta . \| y \| + (1 + k) . \| f(x_0) - z \| = [1 + k - \delta . k] . \| f(x_0) - z \| . \end{aligned}$$

Puisque $1 + k - \delta k < 1$, la définition de x_0 montre que $f(x_0) = z$.

5e *étape.* — Posons $I = B(o, \varepsilon/2) \cap f^{-1}[B(o, r)]$. C'est un voisinage ouvert de o, car f est continue. D'après ce qui précède $f : I \to J = f(I)$ est une bijection. Montrons que l'application inverse est continue.

Si $y, y + k \in J$, il existe $x, x + h \in I$ tels que $y = f(x), y + k = f(x + h)$. D'après (3.1) on a $\| f(x + h) - f(x) \| \geqslant (1 - \delta) . \| h \|$, d'où

$$\| f^{-1}(y + k) - f^{-1}(y) \| \leqslant \| k \| / (1 - \delta) .$$

En résumé $f : I \to J$ est une homéomorphie de classe C^1. Il reste à montrer que f^{-1} est différentiable et que sa différentielle en y est $\{ Df[f^{-1}(y)] \}^{-1}$. Nous ne le ferons pas ici, car la preuve n'est pas plus difficile en dimension infinie, et elle sera donnée au paragraphe suivant.

Remarque. 3.2. — On peut simplifier la 4e étape. Démontrons d'abord un lemme, qui nous sera utile ultérieurement pour d'autres fins.

Lemme. — ***Soit A un sous-ensemble d'un* e.v. *normé E. Si la fonction f : A → R est différentiable (voir 1.4. chap. 1) et présente un minimum en un point a* intérieur *à A, alors Df(a) = 0.***

PREUVE. — Soit $h \in E$. Puisqu'il existe un ouvert contenant a et contenu dans A, alors $a + t.h \in A$ pour $t \in \mathbf{R}$ et $|t|$ assez petit. La fonction différentiable $t \leadsto f(a + t.h) \in \mathbf{R}$ est donc minimale pour $t = 0$. Donc $0 = \left.\frac{d}{dt} f(a + t.h)\right|_{t=0} = Df(a).h$, et $Df(a) = 0$, car h est arbitraire.

Ceci prouvé, et puisque E est de dimension finie, toutes les normes de E sont équivalentes (Appendice A.2.7.). On peut donc supposer qu'il existe sur E un produit intérieur tel que $\langle X, X \rangle$ définisse le carré de la norme. La fonction $x \leadsto \langle f(x) - z, f(x) - z \rangle$ est donc minimale pour $x = x_0$. Ecrivons que sa différentielle en x_0 est nulle. On obtient (règle de Leibniz, 2.9. chap. 1) : $Df(x_0).[f(x_0) - z] = 0$. Et, puisque $Df(x_0)$ est inversible, $f(x_0) = z$.

Remarque. 3.3. — La preuve précédente présente deux inconvénients. D'abord, elle utilise la compacité locale de E, qui est de dimension finie [voir J. Dieudonné, p. 106, pour la réciproque : un espace normé localement compact est de dimension finie]. Ensuite, elle ne fournit pas un algorithme permettant d'approcher effectivement la solution. La preuve qui suit va nous affranchir de ces faiblesses.

4. Preuve du théorème d'inversion locale

Cette fois E et F sont des espaces de Banach quelconques.

La preuve va nécessiter un certain nombre d'étapes.

Inversion d'un isomorphisme d'espaces de Banach. 4.1. — L'ensemble GL $(E; F)$ des isomorphismes de E sur F est un ouvert de $\mathscr{L}(E; F)$ et l'application $\mathfrak{I} : u \leadsto u^{-1}$ de GL $(E; F)$ dans GL $(F; E)$ est continue.

PREUVE. — On peut supposer $E = F$. En effet, si $v \in \mathrm{GL}(E; F)$ [supposé non vide, sinon tout ce qui suit est trivial], l'application $u \leadsto v^{-1} \circ u$ de $\mathscr{L}(E; F)$ dans $\mathscr{L}(E; E)$ est continue et $\mathrm{GL}(E; F)$ n'est autre que l'image inverse de $\mathrm{GL}(E; E)$ par cette application.

Soit $u \in \mathrm{GL}(E; E)$ et $h \in \mathscr{L}(E; E)$. Nous allons montrer que $u + h \in \mathrm{GL}(E; E)$ si $\| h \| < 1/\| u^{-1} \|$. Pour simplifier, désignons par 1 l'application identique id_E. Puisque u est inversible et que $u + h = u_0[1 + u^{-1} \circ h]$, il suffit de prouver (poser $v = - u^{-1} \circ h$) que $1 - v$ est inversible si $\| v \| = \| u^{-1} \circ h \| \leqslant \| u^{-1} \| . \| h \| < 1$. Pour cela inspirons-nous de la formule $(1 - x)^{-1} = 1 + x + x^2, ..$ où $x \in \mathbf{R}$, $| x | < 1$, et considérons la suite $X_0 = 1$, $X_1 = 1 + v, ..., X_n = 1 + v + \cdots + v^n$. C'est une suite de Cauchy de $\mathscr{L}(E; E)$ car [voir A.2.4.] on a $\| X_{p+q} - X_p \| = \| v^{p+1} + \cdots + v^{p+q} \| \leqslant \| v \|^{p+1} + \cdots + \| v \|^{p+q}$, qui tend vers zéro si $p \to + \infty$, car $\| v \| < 1$. Puisque $\mathscr{L}(E; E)$ est complet [voir A.2.3.], X_n converge vers une limite X. Mais $(1 - v) \circ X_n = 1 - v^{n+1}$, la composition est continue [voir A.3.4.] et $v^{n+1} \to 0$ si $n \to + \infty$; donc $(1 - v) \circ X = 1$ et X est l'inverse cherché de $1 - v$.

Gardons les mêmes notations et montrons que $\mathfrak{J}$ est continue.

$$\mathfrak{J}(u + h) - \mathfrak{J}(u) = [(1 - v)^{-1} - 1] \circ u^{-1} = [\lim_{n=\infty} \mathrm{X}_n - 1] \circ u^{-1}$$
$$= \lim_{n=\infty} (v + \cdots + v^n) \circ \mathrm{u}^{-1} .$$

Puisque

$$\| v + \cdots + v^n \| \leqslant \| v \| + \cdots + \| v \|^n \leqslant \| v \|/(1 - \| v \|) \leqslant \| h \| . \| u^{-1} \|/(1 - \| u^{-1} \circ h \|)$$

tend vers zéro si $n \to + \infty$. On a donc bien $\lim_{h \to 0} \| \mathfrak{J}(u + h) - \mathfrak{J}(u) \| = 0$.

PREUVE DU THÉORÈME. 4.2. — Reprenons les notations du paragraphe 2.

1re *étape.* — Comme au paragraphe 3, on peut supposer $a = 0, f(a) = 0, E = F$ et $Df(\mathrm{o}) = \mathrm{id}_E$.

2e *étape.* — Posons $g(x) = x - f(x)$. On a $g(\mathrm{o}) = 0$, $Dg(\mathrm{o}) = 0$. Puisque Dg est continue, il existe donc $r > \mathrm{o}$, tel que $x \in B(\mathrm{o}, 2\,r)$ implique $\| Dg(x) \| \leqslant 1/2$. Le théorème de la moyenne [1.6. chap. 2], applicable dans la boule $B(\mathrm{o}, 2\,r)$, qui est convexe, montre que

$$\| g(x) \| = \| g(x) - g(\mathrm{o}) \| \leqslant \| x \|/2 < r .$$

Ainsi $g[B(\mathrm{o}, 2\,r)] \subset B(\mathrm{o}, r)$.

Prenons $y \in B(\mathrm{o}, r)$; nous allons voir qu'il existe $x \in B(\mathrm{o}, 2\,r)$ unique tel que $f(x) = y$, c'est-à-dire $h(x) = x$, si l'on a posé $h(x) = y + g(x)$.

Etudions h. Si $x \in B(\mathrm{o}, 2\,r)$, on a $\| h(x) \| \leqslant \| y \| + \| g(x) \| \leqslant 2\,r$; h est donc une application de $B(\mathrm{o}, 2\,r)$ dans elle-même. D'autre part, (1.6. chap. 2), appliqué à g dans la boule $B(\mathrm{o}, 2\,r)$, donne $\| h(u) - h(v) \| = \| g(u) - g(v) \| \leqslant \| u - v \|/2$. Ainsi h est une contraction de $B(\mathrm{o}, 2\,r)$ dans elle-même. D'après le théorème du point fixe de Banach [Appendice B], il existe donc un $x \in B(\mathrm{o}, 2\,r)$ et un seul tel que $h(x) = x$, c'est-à-dire $f(x) = y$. Il en résulte que $f^{-1} : B(\mathrm{o}, r) \to B(\mathrm{o}, 2\,r)$ existe.

3e *étape.* Montrons que f^{-1} est 2-lipschitzienne, c'est-à-dire que, pour tous $x, y \in B(\mathrm{o}, r)$, $\| f^{-1}(x) - f^{-1}(y) \| \leqslant 2 . \| x - y \|$. D'après la 2e étape, on peut écrire $x = f(u)$, $y = f(v)$, où $u, v \in B(\mathrm{o}, 2\,r)$. Avec la définition $w = g(w) + f(w)$ de g, on en déduit

$$\| u - v \| \leqslant \| g(u) - g(v) \| + \| f(u) - f(v) \| ,$$

et. puisque $\| g(u) - g(v) \| \leqslant \| u - v \|/2$, $\| u - v \| \leqslant 2 . \| f(u) - f(v) \|$; c'est-à-dire

$$\| f^{-1}(x) - f^{-1}(y) \| \leqslant 2 \| x - y \| .$$

4e *étape.* — Montrons que f^{-1} est de classe C^1. Puisque $x \leadsto [Df(x)]^{-1}$ est la composée des applications continues Df et $\mathfrak{J}$ [voir 4.1.], elle est continue. On a donc pu choisir le nombre r initial assez petit pour que $[Df(x)]^{-1}$ existe sur $B(\mathrm{o}, 2\,r)$; de plus, sa continuité montre qu'il existe un $K > \mathrm{o}$ tel que $\| [Df(x)]^{-1} \| \leqslant K$ sur $B(\mathrm{o}, 2\,r)$. Ceci posé, soient $y, y + k \in B(\mathrm{o}, \mathrm{r})$, alors $x = f^{-1}(y)$ et $x + h = f^{-1}(y + h)$ sont dans $B(\mathrm{o}, 2\,r)$. On a

$$\| f^{-1}(y + k) - f^{-1}(y) - [Df(x)]^{-1}.k \| = \| (x + h) - x - [Df(x)]^{-1}.[f(x + h) - f(x)] \|$$
$$= \| [Df(x)]^{-1}.[f(x + h) - f(x) - Df(x).h] \| \leqslant K.\| f(x + h) - f(x) - Df(x).h \| .$$

Mais $\| f(x + h) - f(x) - Df(x).h \|/\| h \| \to$ o si $h \to$ o et, puisque f^{-1} est 2-lipschitzienne, $\| h \|/\| k \| = \| (x + h) - x \|/\| k \| = \| f^{-1}(y + k) - f^{-1}(y) \|/\| k \| \leqslant 2$. Il en résulte que f^{-1} est différentielle et que $Df^{-1}(y) = [Df(x)]^{-1}$. Cela s'écrit encore $Df^{-1} = \mathfrak{I} \circ Df \circ f^{-1}$, qui est continue, comme composée d'applications continues. □

Remarque. 4.3. — La continuité de Df est essentielle. Ainsi $f : \mathbf{R} \to \mathbf{R}$, définie par

$$f(x) = \begin{cases} (x/2) + x^2.\sin(1/x), & \text{si } x \neq \text{o}, \\ \text{o} & \text{si } x = \text{o}, \end{cases}$$

est différentiable et $f'(\text{o}) \neq 0$. Cependant il n'existe pas d'intervalle ouvert I contenant o sur lequel f soit inversible (le prouver).

Un corollaire : Théorème d'invariance du domaine. 4.4. — ***Soient U un ouvert d'un espace de Banach E, $f : U \to F$ une application étale de classe C^1 dans un espace de Banach F. Alors $f(U)$ est un ouvert de F.***

PREUVE. — Conséquence immédiate du théorème précédent : par hypothèse $Df(u)$ est un isomorphisme de E sur F, pour tout u de U.

5. Le théorème des fonctions implicites

Etudions la fonction $f : \mathbf{R}^2 \to \mathbf{R}$ définie par $f(x, y) = x^2 + y^2 - 1$.

Si $| a | \neq 1$, $b > 0$ et $f(a, b) = 0$, il existe un intervalle A contenant a et un intervalle B contenant b, tels qu'à chaque $x \in A$ corresponde un unique $y \in B$ satisfaisant $f(x, y) = 0$. Cela définit une fonction $x \in A \rightsquigarrow y = g(x) \in B$ vérifiant $f(x, g(x)) = 0$. Dans le cas présent $g(x) = (1 - x^2)^{1/2}$.

On aurait pu associer au nombre a un autre nombre c, égal ici à $- b$, tel que $f(a, c) = 0$. Nous aurions obtenu une autre fonction h, égale ici à $- g$, telle que $f(x, h(x)) = 0$. On dit que chacune des fonctions g et h est définie implicitement par l'équation $f(x, y) = 0$.

Si nous avions choisi $| a | = 1$, il eut été impossible de trouver une telle fonction définie sur un intervalle ouvert contenant a.

Plus généralement, soient m équations $f_i(x, y) = 0$, $i = 1, ..., m$, à m inconnues $y = (y_1, ..., y_m)$. dépendant de n paramètres $x = (x_1, ..., x_n)$. Supposons que $f_i(a, b) = 0$, $i = 1, ..., m$, pour $a = (a_1, ..., a_n)$ et $b = (b_1, ..., b_m)$. Sous quelles conditions peut-on attacher à chaque x, voisin de a, un y unique et voisin de b, vérifiant les m équations $f_i(x, y) = 0$? Le théorème qui va nous occuper donne un critère simple pour répondre à cette question.

Théorème des fonctions implicites. 5.1. — ***Soient E et F deux espaces de Banach, U un ouvert de E, V un ouvert de F et f une application de classe C^1 de $U \times V$ dans un espace de Banach G.***

Supposons qu'en $(a, b) \in U \times V$, la différentielle partielle $D_2 f(a, b) \in \mathcal{L}(F ; G)$ soit un isomorphisme de F sur G. Alors il existe un voisinage A de a, un voisinage W de $f(a, b)$ et une application unique $g_1 : A \times W \to V$, de classe C^1, tels que, pour tous $(x, w) \in A \times W$, on ait $f(x, g_1(x, w)) = w$.

PREUVE. — L'application $\varphi : U \times V \to E \oplus G$, définie par $\varphi(x, y) = (x, f(x, y))$, est de classe C^1, et sa différentielle en (a, b) est donnée par

$$D\varphi(a, b).(h, k) = (h, D_1 f(a, b).h + D_2 f(a, b).k),$$

pour $h \in E$, $k \in F$. Puisque $D_2 f(a, b)$ est un isomorphisme, $D\varphi(a, b)$ en est également un, d'inverse $(h', k') \rightsquigarrow (h', D_2 f(a, b)^{-1}.[k' - D_1 f(a, b).h'])$, pour $h' \in E$, $k' \in G$.

D'après le théorème d'inversion locale, φ est un difféomorphisme de classe C^1 d'un voisinage de (a, b) sur un voisinage de $\varphi(a, b)$. Quitte à réduire ce voisinage de $\varphi(a, b)$. on peut supposer que c'est le produit d'un voisinage ouvert $A \subset U$ de a par un voisinage ouvert W de $f(a, b)$. Le difféomorphisme inverse φ^{-1} est évidemment de la forme $(x, w) \rightsquigarrow (x, g_1(x, w))$. L'application $g_1 : A \times W \to V$ est l'application cherchée. □

Le corollaire qui suit répond à la question posée dans l'introduction.

Corollaire. 5.2. — ***Gardons les hypothèses de 5.1. et supposons, en outre, que $f(a, b) = 0$. Alors il existe un voisinage A de a, un voisinage B de b et une application unique $g : A \to B$ de classe C^1, tels que $f(x, g(x)) = 0$ pour tout $x \in A$. De plus***

$$Dg(x) = -[D_2 f(x, g(x))]^{-1} \circ D_1 f(x, g(x)) .$$

PREUVE. — Il suffit de prendre $w = 0$, $g(x) = g_1(x, o)$ et $B = g_1(A \times \{ o \})$. D'autre part, le théorème de différentiation des applications composées, appliqué à $f(x, g(x)) = 0$, donne $D_1 f[x, g(x)) + D_2 f(x, g(x)) \circ Dg(x) = 0$. Mais, puisque $D_2 f$ est continue, $D_2 f(a, b)$ inversible et (x, y) voisin de (a, b), alors (4.1. chap. 3) implique que $D_2 f(x, g(x))$ est inversible. La dernière partie du corollaire en résulte. □

Remarque. 5.3. — Si $K = \mathbf{R}$ ou $\mathbf{C}$, si $E = K^n$, $F = K^m$, et si $f = (f_1, ..., f_m)$, la condition « $D_2 f(a, b)$ inversible » exprime que le déterminant de la matrice $m \times m$, dont l'élément occupant la i-ième ligne et la j-ième colonne est $D_{n+j} f_i(a, b) = \dfrac{\partial f_i}{\partial y_j}(a, b)$, n'est pas nul.

Théorème. 5.4. — ***Soient E et F deux espaces de Banach, U un ouvert de E et $f : U \to F$ une application de classe C^1.***

Faisons les hypothèses suivantes :

a) ***$Df(a)$ est surjective en $a \in U$;***

b) ***il existe un sous-espace de Banach E_2 de E tel que E soit la somme directe du noyau $E_1 = \mathrm{Ker}\, Df(a)$ et de E_2.***

Alors $f(U)$ contient un voisinage ouvert de $f(a)$.

PREUVE. — Observons d'abord que, puisque $Df(a) \in \mathscr{L}(E; F)$ est continue, son noyau E_1 est un sous-espace fermé de E. C'est donc un espace de Banach.

Remarquons ensuite que l'hypothèse b est superflue si E est de dimension finie ou si c'est un espace de Hilbert [d'après le théorème de la projection, on peut prendre pour E_2 l'orthocomplément de E_1].

Ceci posé, d'après la définition de E_2, l'application $f : U \subset E_1 \oplus E_2 \to F$ possède une différentielle partielle continue $D_2 f(a) : E_2 \to F$, qui est un isomorphisme. Les hypothèses des théorèmes 5.1. sont donc vérifiées et $f(U)$ contient l'ouvert W exhibé en 5.1. □

Une application. 5.5. — Nous allons montrer, sur un exemple, comment le théorème 5.4. peut servir à prouver l'existence de solutions d'équations différentielles.

E est l'espace de Banach $C^1([o, 1])$ des fonctions $u : [o, 1] \to \mathbf{R}$ de classe C^1, normé par $\| u \|_{C^1} = \sup_{[0,1]} | u(t) | + \sup_{[0,1]} | u'(t) |$, étudié en (1.5. Appendice A) ; F est l'espace de Banach $C^0([o, 1])$ des fonctions continues $g : [o, 1] \to \mathbf{R}$, muni de la norme de la convergence uniforme.

L'application $f : E \to F$ est définie par $f(u) : t \rightsquigarrow u'(t) + t.u^2(t)$.

Nous avons vu (1.5. chap. 1) qu'elle est de classe C^1 et que $Df(o) = \dfrac{d}{dt}$. Par conséquent $Df(o)$ est surjective : d'après (6.5. chap. 2), $g \in C^0([o, 1])$ est l'image par $Df(o)$ d'une primitive de g, dont on sait qu'elle est dans $C^1([o, 1])$. D'autre part $E_1 = \mathrm{Ker}\, Df(o)$ est l'ensemble des fonctions constantes. Si E_2 est le sous-espace de Banach de E dont les éléments sont les fonctions u d'intégrale nulle sur $[o, 1]$, on a évidemment $E = E_1 \oplus E_2$.

D'après le théorème 5.4., il existe donc $\varepsilon > 0$ tel que si $g \in C^0([0, 1])$ vérifie $|g(x)| < \varepsilon$ pour tout $x \in [0, 1]$, alors il existe $u \in C^1([0, 1])$ telle que $u'(t) + t.u^2(t) = g(t)$ pour $t \in [0, 1]$.

Note. 5.6. — Nous reviendrons sur le théorème des fonctions implicites et sur ses corollaires au chapitre « Conjugaison et coordonnées locales ».

Le lecteur intéressé trouvera à l'appendice C un algorithme de construction de la racine d'une équation $f(x) = 0$, qui améliore considérablement l'algorithme du théorème de Banach utilisé dans la preuve du théorème d'inversion locale.

Enfin, l'appendice D donne deux théorèmes d'*inversion globale*.

4 DIFFÉRENTIELLES D'ORDRE SUPÉRIEUR

Ce chapitre est consacré aux différentielles d'ordre supérieur d'une application différentiable, aux règles de calcul les concernant et à la formule de Taylor.

1. Différentielles successives. Théorème de Schwarz

Soient E et F des e.v. normés, U un ouvert de E et $f: U \to F$ une application différentiable. La différentielle de f est une application de l'ouvert U dans l'e.v. normé $\mathscr{L}(E; F)$. On peut donc se demander si Df est différentiable en $a \in U$, ou même en tout point de U.

Définitions. 1.1. — On dit que *f est deux fois différentiable en $a \in U$* si Df est différentiable en a. La différentielle de Df en a est notée $D^2 f(a)$. C'est un élément de $\mathscr{L}(E; \mathscr{L}(E; F))$, qu'on appelle la *différentielle seconde de f en a.*

Bien entendu la définition n'exige pas que Df existe en tout point de U. Il suffit que Df soit définie sur un voisinage ouvert $U' \subset U$ de a, et soit différentiable en a.

Si $Df: U \to \mathscr{L}(E; F)$ est différentiable en chaque point de U, on dit que f est deux fois différentiable dans U. Si c'est le cas, $u \in U \rightsquigarrow D^2 f(u)$ est une application $D^2 f: U \to \mathscr{L}(E; \mathscr{L}(E; F))$, qu'on appelle la *différentielle seconde de f.* Il se peut alors que $D^2 f$ soit continue dans U; s'il en est ainsi, on dit que *f est de classe C^2 dans U.*

Evidemment, les deux propriétés suivantes sont équivalentes :

a) f est la classe C^2 dans U ;

b) f est différentiable dans U et Df est de classe C^1 dans U.

Interprétation de $D^2 f(a)$. 1.2. — Rappelons que $\mathscr{L}(E; \mathscr{L}(E; F))$ est canoniquement isomorphe à l'e.v. normé $\mathscr{L}^2(E; F)$ [voir 4.2. Appendice A]. L'image de $D^2 f(a) \in \mathscr{L}(E; \mathscr{L}(E; F))$ dans cet isomorphisme est donc une application bilinéaire continue de $E \times E$ dans F, qui est définie par $(h, k) \rightsquigarrow (D^2 f(a).h).k$: $D^2 f(a)$ fait correspondre à $h \in E$ l'élément $D^2 f(a).h$ de $\mathscr{L}(E; F)$, et l'image de $k \in E$ par cette dernière application est $(D^2 f(a).h).k$, que l'on notera $D^2 f(a).(h, k)$.

Calcul de $D^2 f(a)$. 1.3. — Supposons f deux fois différentiable en a. Appliquons la formule (1.2) du chapitre 1 à Df :

$$D^2 f(a).h = D(Df)(a).h = \frac{\mathrm{d}}{\mathrm{d}t} Df(a + t.h) \Big|_{t=0} .$$

On obtient ainsi un élément de $\mathscr{L}(E; F)$, dont la valeur en $k \in E$ est

$$D^2 f(a).(h, k) = \frac{\mathrm{d}}{\mathrm{d}t} Df(a + t.h).k \Big|_{t=0} .$$

Mais, toujours d'après la formule ((1.2) chap. 1), $Df(a + t.h).k = \dfrac{\mathrm{d}}{\mathrm{d}s} f(a + t.h + s.k) \Big|_{s=0}$.

Par conséquent, $D^2 f(a).(h, k) = \dfrac{\mathrm{d}}{\mathrm{d}t} \cdot \dfrac{\mathrm{d}}{\mathrm{d}s} f(a + t.h + s.k) \Big|_{t=s=0}$.

EXEMPLE : APPLICATIONS BILINÉAIRES CONTINUES. 1.4. — Reprenons l'application bilinéaire continue $b : E_1 \oplus E_2 \to F$ de (2.3. chap. 1). Si $a = (a_1, a_2)$ et $h = (h_1, h_2)$ sont deux éléments de $E_1 \oplus E_2$, on a vu que $Db(a).h = b(h_1, a_2) + b(a_1, h_2)$. L'application Db :

$$E_1 \oplus E_2 \to \mathscr{L}(E_1, E_2 ; F)$$

est donc une application linéaire continue. D'après (2.2. chap. 1) sa différentielle existe donc et elle est constante. Cette constante est l'élément $D^2 f(a) \in \mathscr{L}^2(E_1 \oplus E_2 ; F)$ dont la valeur sur les éléments $h = (h_1, h_2)$ et $k = (k_1, k_2)$ de $E_1 \oplus E_2$ est $D^2 f(a).(h, k) = b(h_1, k_2) + b(k_1, h_2)$, comme on le voit aussi en utilisant 1.3.

Théorème de Schwarz. 1.5. — ***Si $f : U \in E \to F$ est deux fois différentiable en $a \in U$, alors $\| f(a + u + v) - f(a + u) - f(a + v) + f(a) - D^2 f(a).(u, v) \|/(\| u \| + \| v \|)^2$ tend vers zéro si u et v tendent vers zéro.***

En particulier $D^2 f(a)$ est une forme bilinéaire symétrique : $D^2 f(a).(u, v) = D^2 f(a).(v, u)$ pour tous $u, v \in E$.

PREUVE. — Puisque Df est différentiable en a, à tout $\varepsilon > o$ correspond un $\delta > o$ tel que $\| x \| < 2\delta$ entraîne

$$\| Df(a + x) - Df(a) - D^2 f(a).x \| \leqslant \varepsilon . \| x \| . \tag{1.6}$$

Choisissons u et v de norme inférieure à δ et posons

$$g_v(u) = f(a + u + v) - f(a + u) - f(a + v) + f(a) - D^2 f(a).(u, v) .$$

Tenons compte de la règle de différentiation d'une application linéaire continue $u \rightsquigarrow D^2 f(a).(u, v)$, on a :

$$\begin{aligned} Dg_v(u) &= Df(a + u + v) - Df(a + u) - D^2 f(a).v = \\ &= [Df(a + u + v) - Df(a) - D^2 f(a).(u + v)] - [Df(a + u) - Df(a) - D^2 f(a).u] , \end{aligned}$$

et (1.6) entraîne $\| Dg_v(u) \| \leqslant 2\varepsilon.(\| u \| + \| v \|)$.

Appliquons le théorème (1.4. chap. 2) à g_v dans la boule $B(o, \| u \|)$. Puisque $g_v(o) = o$, on obtient $\| g_v(u) \| = \| g_v(u) - g_v(o) \| \leqslant 2\varepsilon.(\| u \| + \| v \|).\| u \| \leqslant 2\varepsilon.(\| u \| + \| v \|)^2$, et la première partie du théorème en résulte.

Puisque $f(a + u + v) - f(a + u) - f(a + v) + f(a)$ est symétrique en u et v, il en est donc de même de $D^2 f(a).(u, v)$. □

Définissons maintenant les différentielles d'ordre quelconque, en reprenant les notations du début : f est une application d'un ouvert U d'un e.v. normé E dans un e.v. normé F.

Définitions. 1.7. — ***Convenons que l'expression « f est une fois différentiable » signifie « f est différentiable ». Nous allons définir, par récurrence sur l'entier $n > o$, l'expression « f est n fois différentiable dans U » et la différentielle $D^n f(a)$ d'ordre n de f en $a \in U$.***

Rappelons, à cet effet, que l'e.v. normé des applications linéaires continues de E dans l'e.v. normé $\mathscr{L}^{n-1}(E; F)$ des applications $(n - 1)$-linéaires continues de $E \times \cdots \times E (n - 1$ fois$)$ dans F est canoniquement isomorphe à $\mathscr{L}^n(E ; F)$ [voir 4.2. Appendice A].

Ceci rappelé, on dit que f **est n fois différentiable en** ***$a \in U$ s'il existe un voisinage ouvert $U' \subset U$ de a où f est $n - 1$ fois différentiable en chaque point, et si l'application $u \in U' \rightsquigarrow D^{n-1} f(u)$ de U' dans $\mathscr{L}^{n-1}(E; F)$ est différentiable en a. La différentielle de $D^{n-1} f$ au point a se note $D^n f(a)$ et*** **s'appelle la différentielle d'ordre n de f en a.** ***C'est un élément de $\mathscr{L}(E ; \mathscr{L}^{n-1}(E; F)) = \mathscr{L}^n(E ; F)$, et sa valeur en $(h_1, ..., h_n) \in E^n$ est notée $D^n f(a).(h_1, ..., h_n)$. Expression égale, d'après la définition même, à $(D(D^{n-1} f)(a).h_1).(h_2, ..., h_n)$.***

Le lecteur vérifiera l'équivalence de la définition précédente avec la suivante : f est n fois différentiable en a si elle est différentiable dans un voisinage ouvert $U' \subset U$ de a, et si $Df : U' \to \mathscr{L}(E; F)$ est $(n - 1)$ fois différentiable en a.

Si f est n fois différentiable en chaque point de U, on dit que f **est n fois différentiable dans U.** ***S'il en est ainsi, $u \rightsquigarrow D^n f(u)$ définit une application $D^n f : U \to \mathscr{L}^n(E; F)$, qu'on appelle***

la* différentielle d'ordre *n* de *f*. *Si l'application $D^n f$ est continue dans U, on dit que f* est de classe C^n dans *U. Le lecteur vérifiera l'équivalence de cette définition avec la suivante : f est de classe C^n dans U si f est de classe C^1 dans U et si Df est de classe C^{n-1} dans U.

Observons que, puisqu'une application différentiable en un point est continue en ce point, une application n fois différentiable est de classe C^{n-1}.

Les espaces $C^n(U\,;F)$. 1.8. — La linéarité de la différentiation montre que les applications $f : U \to F$ de classe C^n forment un e.v. ; on le notera $C^n(U\,; F)$. En particulier $C^0(U\,; F)$ est l'espace des applications continues de U dans F.

L'intersection $\bigcap\limits_{n>0} C^n(U\,; F)$ se note $C^\infty(U\,; F)$; ses éléments sont les applications $f : U \to F$ qui sont de classe C^n pour tout entier $n > 0$.

On dit qu'un tel élément est de *classe C^∞*, ou, ce qui revient au même, est *indéfiniment différentiable*, c'est-à-dire n fois différentiable pour tout n.

Généralisation du théorème de Schwarz. 1.9. — ***Si $f : U \subset E \to F$ est n fois différentiable en $a \in U$, alors $D^n f(a) \in \mathscr{L}^n(E\,; F)$ est une application n-linéaire symétrique. Autrement dit, pour tous $h_1, \ldots, h_n \in E$ et pour toute permutation s des entiers $\{1, 2, \ldots, n\}$, on a***

$$D^n f(a).(h_1, \ldots, h_n) = D^n f(a).(h_{s(1)}, \ldots, h_{s(n)}) .$$

PREUVE. — La question ne se pose que pour $n \geqslant 2$, et le théorème 1.5. y répond si $n = 2$.

Supposons donc $n \geqslant 3$ et prouvons le théorème par récurrence, en le supposant vrai pour $n - 1$. Par hypothèse $D^n f(a) = D(D^{n-1} f)(a)$. D'autre part $D^{n-1} f$ est à valeurs dans le sous-espace $\mathscr{L}_s^{n-1}(E\,; F)$ de $\mathscr{L}^{n-1}(E\,; F)$ formé des applications $(n-1)$-linéaires symétriques. Par suite $D^n f(a).h_1 = D(D^{n-1} f)(a).h_1 \in \mathscr{L}_s^{n-1}(E\,; F)$ pour $h_1 \in E$, et

$$D^n f(a).(h_1, \ldots, h_n) = (D^n f(a).h_1).(h_2, \ldots, h_n)$$

est une fonction symétrique de $h_2, \ldots, h_n$.

Si la permutation s laisse le nombre 1 fixe, le théorème est donc démontré.

Si la permutation s ne laisse pas le nombre 1 fixe, elle est le produit de permutations laissant 1 fixe par une permutation échangeant les entiers 1 et 2, en laissant $3, \ldots, n$ fixes. Le théorème sera donc démontré si l'on prouve que $D^n f(a).(h_1, h_2, \ldots, h_n)$ ne change pas lorsqu'on échange h_1 et h_2. Mais $D_n f(a) = D^2(D^{n-2} f(a))$; donc, en appliquant le théorème 1.5. à $D^{n-2} f$,

$$(D^n f(a).h_1).h_2 = D^2(D^{n-2} f)(a).(h_1, h_2) = D^2(D^{n-2} f)(a).(h_2, h_1) = (D^n f(a).h_2).h_1 . \quad \square$$

Calcul de $D^n f(a)$. 1.10. — Supposons que $f : U \subset E \to F$ soit n fois différentiable en $a \in U$. En appliquant de façon répétée la formule (1.2. chap. 1) on obtient la généralisation suivante de (1.3) :

$$D^n f(a).(h_1, \ldots, h_n) = \frac{\mathrm{d}}{\mathrm{d}t_n} \cdots \frac{\mathrm{d}}{\mathrm{d}t_1} f\left(a + \sum_{i=1}^{n} t_i.h_i\right)\Bigg|_{t_1 = \ldots = t_n = 0}, \tag{1.11}$$

où $h_1, \ldots, h_n \in E$.

Cette formule ramène le calcul de $D^n f(a)$ à celui des dérivées successives d'une fonction de n variables réelles $t_1, \ldots, t_n$ et à valeurs dans F.

Le nombre de droite de (1.11) généralise la notion de dérivée de f dans la direction d'un vecteur h de E. Notons $\mathrm{d}^n f(a)$ l'opération qui associe à $(h_1, \ldots, h_n) \in E^n$ ce nombre de droite [on l'appelle quelquefois la dérivée de Gâteaux d'ordre n de f au point a]. Nous allons démontrer une généralisation du théorème du paragraphe 3 du chapitre 2.

Théorème. 1.12. — Supposons que

a) ***$\mathrm{d}^n f(x)$ existe dans un voisinage ouvert U de a ;***

b) ***$\mathrm{d}^n f(x) \in \mathscr{L}^n(E\,; F)$ pour tout x de U ;***

c) ***$x \in U \mapsto \mathrm{d}^n f(x)$ soit continue. Alors f est de classe C^n en a et $\mathrm{d}^n f(a) = D^n f(a)$.***

PREUVE. — Montrons-le par récurrence sur n.

Si $n = 1$, $t \rightsquigarrow f(x + t.h)$ est dérivable en $t = 0$ pour tout $x \in U$ et pour tout $h \in E$, et sa dérivée $df(x).h$ dépend continûment de x. Posons $x = a + s.h$, où $0 \leqslant s \leqslant 1$ et $\| h \|$ assez petit pour que $x \in U$. Alors $t \rightsquigarrow f(a + s.h + t.h)$ est dérivable en $t = 0$ et sa dérivée $df(a + s.h).h$, qui n'est autre que la dérivée en s de $s \rightsquigarrow f(a + sh)$, est continue. Le théorème fondamental du calcul intégral entraîne donc $f(a + h) - f(a) = \int_0^1 df(a + s.h).h.ds$.

Par suite

$$\| f(a + h) - f(a) - df(a).h \| = \left\| \int_0^1 [df(a + s.h) - df(a)].h.ds \right\|$$
$$\leqslant \| h \| \cdot \int_0^1 \| df(a + sh) - df(a) \|.ds,$$

qui est un $o(\| h \|)$ puisque $x \rightsquigarrow df(x)$ est continue en a. Ainsi $Df(a)$ existe, elle est égale à $df(a)$ et Df est donc continue en a.

Supposons le théorème établi pour $n - 1$. Par hypothèse, et d'après le théorème fondamental du calcul intégral,

$$[D^{n-1} f(a + h_n) - D^{n-1} f(a)].(h_1, ..., h_{n-1}) = [d^{n-1}f(a + h_n) - d^{n-1}f(a)].(h_1, ..., h_{n-1}) =$$
$$\int_0^1 (d^n f(a + s.h_n).h_n).(h_1, ..., h_{n-1}).ds.$$

On en déduit

$$\| [D^{n-1} f(a + h_n) - D^{n-1} f(a) - d^n f(a).h_n].(h_1, ..., h_{n-1}) \| \leqslant$$
$$\left\| \int_0^1 [d^n f(a + s.h_n) - d^n f(a)].(h_1, ..., h_m).ds \right\| \leqslant$$
$$\| h_1 \| ... \| h_n \| \cdot \int_0^1 \| d^n f(a + s \cdot h_n) - d^n f(a) \| ds.$$

Puisque $d^n f$ est continue en a, l'intégrale tend vers zéro avec $\| h \|$. Il en résulte que $D^{n-1} f(a + h_n) - D^{n-1} f(a) - d^n f(a).h_n = o(| h_n |)$; $D^n f(a)$ existe donc, elle est égale à $d^n f(a)$ et elle est donc continue en a.

2. Règles de calcul

Applications linéaires continues. 2.1. — Si f est une application linéaire continue de l'e.v. normé E dans l'e.v. normé F, on sait (2.1. chap. 1) que sa différentielle est constante. Donc $D^n f = 0$ pour $n \geqslant 2$ et f est de classe C^∞.

Applications bilinéaires continues. 2.2. — Soient E_1, E_2 et F trois e.v. normés et $b : E_1 \times E_2 \to F$ une application bilinéaire continue. On sait (1.4.) que Db est différentiable et que $D^2 b$ est une constante. Donc $D^n b = 0$ pour $n \geqslant 3$ et b est de classe C^∞.

Règle de Leibniz. 2.3. — Soient E, F_1, F_2 et G des e.v. normés, U un ouvert de E, $f : U \to F_1$ et $g : U \to F_2$ des applications de classe C^k et $b : F_1 \times F_2 \to G$ une application bilinéaire continue. Définissons, comme en (2.9. chap. 1), l'application $p : U \to G$ par $x \rightsquigarrow b(f(x), g(x))$. Alors p est de classe C^k.

PREUVE. — Supposons f et g de classe C^1. Selon (2.9. chap. 1) p est différentiable en $a \in U$ et si $h \in E$ on a

$$(2.4) \qquad Dp(a).h = b(Df(a).h, g(a)) + b(f(a), Dg(a).h).$$

Par conséquent Dp est continue et p est de classe C^1.

Nous allons démontrer le théorème par récurrence en le supposant vrai pour $k - 1$. Supposons f et g de classe C^k. L'application $v : \mathscr{L}(E ; F_1) \times F_2 \to \mathscr{L}(E ; G)$, qui fait correspondre à $A \in \mathscr{L}(E ; F_1)$ et $l \in F_2$ l'application linéaire continue $h \rightsquigarrow b(A(h), l)$ de E dans G, est continue et bilinéaire. D'autre part Df et g sont de classe C^{k-1}. Puisque le théorème est admis pour $k - 1$, il en résulte que $x \rightsquigarrow v(Df(x), g(x)) = b(Df(x).., g(x))$ est de classe C^{k-1}.

On montre de même que $x \rightsquigarrow b(f(x), Dg(x).)$ est de classe C^{k-1}. Donc, d'après la formule (2.4), Dp est de classe C^{k-1} et p est de classe C^k.

La preuve montre que l'on peut remplacer « de classe C^k » par « k fois différentiable » dans l'énoncé du théorème.

Différentielles d'ordres supérieures d'un produit. 2.5.

Si f et g sont deux fonctions à valeurs numériques définies sur un ouvert où elles sont de classe C^n, la règle précédente montre que leur produit $f.g$ est de classe C^n. En appliquant n fois la règle de dérivation d'un produit, on voit que la dérivée n-ième est de la forme

$$(f.g)^{(n)} = \sum_{q=0}^{n} A_{qn}.f^{(n-q)}.g^{(q)},$$

où les coefficients A_{qn} sont des constantes indépendantes de f et g. Spécialisons ces fonctions en choisissant $f(x) = e^{ax}$, $g(x) = e^{bx}$, où a et b sont des scalaires. Après simplification la formule précédente devient $(a + b)^n = \sum A_{qn}.a^{n-q}.b^q$. Puisque a et b sont arbitraires les coefficients A_{qn} sont ceux de la formule du binôme : $A_{qn} = C_n^q = n\,!/(n - q)\,!\,q\,!$. Ainsi

$$(2.6) \qquad (f.g)^{(n)} = \sum_{q=0}^{n} C_n^q.f^{(n-q)}.g^{(q)}.$$

La preuve de la généralisation suivante est laissée à la discrétion du lecteur : sous les hypothèses de (2.3) la différentielle d'ordre n de p est donnée par

$$D^n p(x).(h_1, ..., h_n) = \sum_{q=0}^{n} \sum b[D^q f(x).(h_{i_1}, ..., h_{i_q}), D^{n-q} g(x).(h_{j_1}, ..., h_{j_{n-q}})]$$

où $\sum$ est étendue aux C_n^q partitions de $(h_1, ..., h_n) \in E^n$ en deux sous-ensembles tels que $i_1 < \cdots < i_q$ et $j_1 < \cdots < j_{n-q}$.

On peut donner à cette expression une forme semblable à celle de (2.6). Soit sym l'opération de symétrisation, qui fait correspondre à $A \in \mathscr{L}^n(E; F)$ l'application n-linéaire symétrique sym (A) définie par

$$\operatorname{sym}(A).(h_1, ..., h_n) = \frac{1}{n\,!} \cdot \sum_s A[h_{s(1)}, ..., h_{s(n)}],$$

où la sommation est étendue à toutes les permutations s de $\{ 1, 2, ..., n \}$. En tenant compte du théorème de Schwarz (1.9) l'expression précédente se condense en

$$D^n p = \operatorname{sym} \sum_{q=0}^{n} C_n^q.b(D^{n-q} f, D^q g).$$

Différentielles d'applications composées. 2.7. — ***Soient E, F et G trois* e.v. *normés, U un ouvert de E et V un ouvert de F. Supposons que f : U → V et g : V → G soient deux applications de classe C^n, alors g ∘ f est aussi de classe C^n.***

PREUVE. — On sait déjà (2.6. chap. 1) que si f et g sont différentiables, alors $g \circ f$ l'est aussi et que $D(g \circ f)(x) = Dg[f(x)] \circ Df(x)$.

Si f et g sont de classe C^1, Dg et f sont continues; l'application composée $x \rightsquigarrow Dg[f(x)]$ est donc continue. Comme $Df(x)$ l'est aussi, la formule précédente montre que $D(g \circ f)$ est continue ; $g \circ f$ est donc de classe C^1.

Démontrons le théorème par récurrence en le supposant vrai pour $n - 1$. Supposons f et g de classe C^n, Dg et f sont donc de classe C^{n-1} et, d'après l'hypothèse de récurrence, l'application composée $x \rightsquigarrow Dg[f(x)]$ est de classe C^{n-1}. Df l'est aussi. D'autre part l'application $\mathscr{L}(F ; G) \times \mathscr{L}(E ; F) \to \mathscr{L}(E ; G)$ qui fait correspondre à $A \in \mathscr{L}(F ; G)$ et $B \in \mathscr{L}(E ; F)$ leur composé $A \circ B$, est continue et bilinéaire. Donc, d'après la règle de Leibniz,

$$x \rightsquigarrow Dg[f(x)] \circ Df(x)$$

est de classe C^{n-1} et $g \circ f$ est de classe C^n.

Remarque. — Sous les hypothèses du théorème précédent, et avec la notion de foncteur tangent T introduite en (2.6. chap. 1), on a $T^n(g \circ f) = T^n g \circ T^n f$ par application répétée de la formule $T(g \circ f) = Tg \circ Tf$.

Un résultat analogue en termes de D est passablement plus compliqué :

$$D^n(g \circ f)(x).(h_1, \ldots, h_n) = \sum_{q=1}^{n} \sum D^q g[f(x)].(D^{r_1} f(x).(h_1^{(1)}, \ldots, h_{r_1}^{(1)}), \ldots, D^{r_q} f(x).(h_1^{(q)}, \ldots, h_{r_q}^{(q)})) ,$$

où la somme $\sum$ est étendue aux $n!/r_1! \ldots r_q!$ partitions de $h_1, \ldots, h_n$ en q sous-ensembles. comptant respectivement $r_1, \ldots, r_q$ éléments, et dont les h sont rangés dans l'ordre croissant de leurs indices. Si l'on pose, pour abréger (!),

$$(D^{r_1} f, \ldots, D^{r_q} f)(x).(h_1, \ldots, h_n) =$$
$$(D^{r_1} f(x).(h_1, \ldots, h_{r_1}), D^{r_2} f(x).(h_{r_1+1}, \ldots, h_{r_1+r_2}), \ldots, D^{r_q} f(x)(h_{r_1+\cdots+r_{q-1}+1}, \ldots, h_n)) ,$$

en utilisant le théorème de Schwarz et la symétrisation sym, la formule ci-dessus prend la forme condensée

$$D^n(g \circ f) = \operatorname{sym} \sum_{q=1}^{n} \sum_{r_1+\cdots+r_q=n} \frac{n!}{r_1! \ldots r_q!} (D^q g \circ f)(D^{r_1} f, \ldots, D^{r_q} f)$$

[voir : L. E. Fraenkel et T. Ratiu, Math. Proc. Camb. Phil. Soc. 83 (1978), page 159. Voir aussi : H. Federer. Geometric measure theory. Springer-Verlag (1969), page 222].

Inversion d'un isomorphisme d'espaces de Banach. 2.8.

Reprenons les notations de (4.1. chap. 3) : GL $(E ; F)$ désigne l'ouvert formé par les isomorphismes de l'espace de Banach E sur l'espace de Banach F ; $\mathfrak{J}$ désigne l'inversion $u \rightsquigarrow u^{-1}$. qui applique GL $(E ; F)$ sur GL $(F ; E)$.

Théorème. — ***L'application $\mathfrak{J}$ est de classe C^∞ et $D\mathfrak{J}(u).h = -u^{-1} \circ h \circ u^{-1}$ pour $h \in \mathscr{L}(E ; F)$.***

Preuve. — Posons $L.h = -u^{-1} \circ h \circ u^{-1}$. Evidemment L est linéaire et elle est continue car. d'après (2.4, Appendice A), $\| L.h \| \leqslant \| u^{-1} \| . \| u \| . \| h \|$. Etudions $A = \mathfrak{J}(u + h) - \mathfrak{J}(u) - L.h$. On a $A = (u+h)^{-1} \circ [1 - (u+h) \circ u^{-1} + (u+h) \circ u^{-1} \circ h \circ u^{-1}] = (u+h)^{-1} \circ h \circ u^{-1} \circ h \circ u^{-1}$. d'où $\| A \| \leqslant \| (u + h)^{-1} \| . \| u^{-1} \|^2 . \| h \|^2$. Puisque $\mathfrak{J}$ est continue (4.1. chap. 3), $(u+h)^{-1} \to u^{-1}$ si $h \to 0$ et par conséquent $\| A \| / \| h \| \to 0$ si $h \to 0$. Cela montre que $\mathfrak{J}$ est différentiable et que $D\mathfrak{J}(u).h = -u^{-1} \circ h \circ u^{-1}$.

Montrons que $\mathfrak{J}$ est de classe C^1. A cet effet introduisons une notation : si $a, b \in \mathscr{L}(F ; E)$ notons $f(a, b)$ l'application linéaire $h \rightsquigarrow -a \circ h \circ b$ de $\mathscr{L}(E ; F)$ dans $\mathscr{L}(F ; E)$. En sorte que $D\mathfrak{J}(u) = f(u^{-1}, u^{-1})$. On voit que

$$(a, b) \in \mathscr{L}(F ; E) \times \mathscr{L}(F ; E) \rightsquigarrow f(a, b) \in \mathscr{L}[\mathscr{L}(E ; F) ; \mathscr{L}(F ; E)]$$

est bilinéaire et qu'elle est continue, car $\| f(a, b).h \| = \| a \circ h \circ b \| \leqslant \| a \| . \| b \| . \| h \|$. Par conséquent $D\mathfrak{J}$ est continue, car c'est la composée de l'application continue $u \rightsquigarrow (u^{-1}, u^{-1})$

de $\mathrm{GL}(E;F)$ dans $\mathscr{L}(F;E) \times \mathscr{L}(F;E)$ et de l'application continue $(a,b) \rightsquigarrow f(a,b)$. On a donc prouvé que $\mathfrak{J}$ est de classe C^1.

Supposons le théorème établi jusqu'à l'ordre n et supposons $\mathfrak{J}$ de classe C^n, $n \geqslant 1$. L'application $u \rightsquigarrow u^{-1}$ est donc de classe C^n. D'autre part l'application bilinéaire continue $(a,b) \rightsquigarrow f(a,b)$ est de classe C^∞ d'après 2.2. Il résulte de la règle de Leibniz que $D\mathfrak{J} : u \rightsquigarrow f(u^{-1}, u^{-1})$ est de classe C^n ; l'application $\mathfrak{J}$ est donc de classe C^{n+1}. Cela démontre que $\mathfrak{J}$ est de classe C^∞. □

Remarque. — Soient $a \in \mathrm{GL}(E;F)$ et $h \in \mathscr{L}(E;F)$ tels que $\| a^{-1} \circ h \| < 1$. La formule

$$(a+h)^{-1} = \sum_{k=0}^{\infty} (-a^{-1} \circ h)^k \circ a^{-1}$$

démontrée au (4.1. chap. 3), permet d'établir que

$$D^k \mathfrak{J}(a).(h_1, \dots, h_k) = \sum_s (-1)^k (a^{-1} \circ h_{s(1)} \circ a^{-1}) \circ \cdots \circ (a^{-1} \circ h_{s(k)} \circ a^{-1}),$$

où la sommation est étendue à toutes les permutations s de $\{1, 2, \dots, k\}$ [le prouver].

Inverse d'un difféomorphisme de classe C^n. 2.9. — ***Soit $f : U \to V$ un difféomorphisme de classe C^n, $n \geqslant 1$, d'un ouvert U d'un* e.v. *normé E sur un ouvert V d'un* e.v. *normé F. Alors le difféomorphisme inverse f^{-1} est de classe C^n.***

Preuve. — Si $n = 1$ le résultat est réduit à une définition ; il n'y a rien à démontrer et l'on sait (1.2. chap. 3) que $Df^{-1} = [Df \circ f^{-1}]^{-1}$. Cela s'écrit encore $Df^{-1} = \mathfrak{J} \circ Df \circ f^{-1}$.

Supposons le théorème établi pour $n - 1 \geqslant 1$. Afin de le prouver pour n, nous allons montrer que si f est de classe C^n alors Df^{-1} est de classe C^{n-1}. Puisque f est, en particulier, de classe C^{n-1}, l'hypothèse de récurrence dit que f^{-1} est de classe C^{n-1}. Puisque f est de classe C^n, Df est de classe C^{n-1}. Quant à $\mathfrak{J}$, nous venons de voir qu'elle est de classe C^∞. D'après le théorème 2.7. et la formule $Df^{-1} = \mathfrak{J} \circ Df \circ f^{-1}$ on conclut que Df^{-1} est bien de classe C^{n-1}. □

Corollaire. 2.10. — ***Si dans le théorème d'inversion locale* (2. chap. 3) *on suppose en outre f de classe C^n, $n \geqslant 1$, alors f est un difféomorphisme local de classe C^n [son inverse local f^{-1} est de classe C^n].***

Application définie sur une somme directe. 2.11. — Supposons que E soit la somme directe $E_1 \oplus \cdots \oplus E_n$ d'e.v. normés et soit f une application d'un ouvert U de E dans un e.v. normé F.

Supposons f de classe C^k, $k \geqslant 1$. En utilisant (2.10. chap. 1) et le théorème 2.7. concernant les applications composées, on voit que chaque différentielle partielle $D_r f : U \in \mathscr{L}(E_r; F)$ existe et qu'elle est de classe C^{k-1}. De plus si $h = (h_1, \dots, h_n)$, $h_r \in E_r$, on sait que

$$Df(a).h = \sum_{r=1}^{n} D_r f(a).h_r \quad \text{pour} \quad a \in U. \tag{2.12}$$

On se propose de trouver une formule semblable pour $D^k f(a)$, et tout d'abord pour $k = 2$. A cette fin utilisons la formule (1.11) :

$$D^2 f(a).(h^{(1)}, h^{(2)}) = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}s} f(a + t.h^{(1)} + s.h^{(2)}) \Big|_{t=s=0},$$

pour $h^{(1)}, h^{(2)} \in E$. Pour t assez petit $a + t.h^{(1)} \in U$ et l'on a

$$\frac{\mathrm{d}}{\mathrm{d}s} f(a + t.h^{(1)} + s.h^{(2)}) \Big|_{s=0} = Df(a + t.h^{(1)}).h^{(2)}.$$

D'après la formule (2.12), appliquée au point $a + t.h^{(1)}$, cela s'écrit encore

$$\sum_{r=1}^{n} D_r f(a + t.h^{(1)}).h_r^{(2)}.$$

Finalement, dérivons par rapport à t et faisons $t = 0$, on obtient

$$(2.13) \qquad D^2 f(a).(h^{(1)}, h^{(2)}) = \sum_{r=1}^{n} \frac{d}{dt} D_r f(a + t.h^{(1)})\Big|_{t=0} .h_2^{(2)} .$$

Explicitons le membre de droite. L'application $x \in E \rightsquigarrow D_r f(x) \in \mathscr{L}(E_r; F)$ est différentiable et sa différentielle $D(D_r f)(x) \in \mathscr{L}(E; \mathscr{L}(E_r; F))$ s'exprime à son tour par la formule (2.12) :

$$D(D_r f)(x).h^{(1)} = \sum_{s=1}^{n} D_s(D_r f)(x).h_s^{(1)}, \quad \text{où} \quad D_s(D_r f)(x) \in \mathscr{L}(E_s; \mathscr{L}(E_r; F)).$$

En combinant ce résultat avec $\frac{d}{dt} D_r f(a + t.h^{(1)})\Big|_{t=0} = D(D_r f)(a).h^{(1)}$ et la formule (2.13) on obtient

$$(2.14) \qquad D^2 f(a).(h^{(1)}, h^{(2)}) = \sum_{r,s=1}^{n} D_s(D_r f)(a).(h_s^{(2)}, h_r^{(1)}).$$

On pose $D_s(D_r f)(a) = D_{sr}^2 f(a)$; c'est un élément de $\mathscr{L}(E_s; \mathscr{L}(E_r; F))$, qu'on sait être canoniquement isomorphe à $\mathscr{L}(E_s, E_r; F)$. [4.2. Appendice A]. On l'appelle une différentielle partielle d'ordre 2 de f au point a.

D'après le théorème de Schwarz, $D^2 f(a)$ est une forme bilinéaire symétrique. La formule (2.14) entraîne donc

$$\sum_{r,s} D_{sr}^2 f(a).(h_s^{(1)}, h_r^{(2)}) = \sum_{r,s} D_{sr}^2 f(a).(h_s^{(2)}, h_r^{(1)}).$$

En échangeant les indices de sommation r et s dans le membre de droite on obtient

$$\sum_{r,s} D_{sr}^2 f(a).(h_s^{(1)}, h_r^{(2)}) = \sum_{r,s} D_{rs}^2 f(a).(h_r^{(2)}, h_s^{(1)}).$$

Comme les $h_s^{(1)}$ et les $h_r^{(2)}$ sont arbitraires, on en déduit

$$D_{sr}^2 f(a).(h_s^{(1)}, h_r^{(2)}) = D_{rs}^2 f(a).(h_r^{(2)}, h_s^{(1)}).$$

L'application $D_{rs}^2 f(a) \in \mathscr{L}(E_r, E_s; F)$ est donc la composée de l'application

$$(h_r^{(2)}, h_s^{(1)}) \in E_r \times E_s \rightsquigarrow (h_s^{(1)}, h_r^{(2)}) \in E_s \times E_r$$

et de l'application $D_{sr}^2 f(a) \in \mathscr{L}(E_s, E_r; F)$. En particulier $D_{rr}^2 f(a)$ est une application bilinéaire symétrique. Par contre il faut se garder de croire que $D_{rs}^2 f(a) = D_{sr}^2 f(a)$: ce ne sont même pas des éléments du même espace si $s \neq r$.

Si maintenant f est de classe C^k, $k \geqslant 2$, la formule (2.14) se généralise sans peine :

$$D^k f(a).(h^{(1)}, \ldots, h^{(k)}) = \sum_{r_1,\ldots,r_k} D_{r_1,\ldots,r_k}^k f(a).(h_{r_1}^{(1)}, \ldots, h_{r_k}^{(k)}),$$

avec des notations évidentes.

Cas où $E = \mathbf{R}^n$ ou $\mathbf{C}^n$. 2.16. — Dans le paragraphe précédent prenons $E_1 = \cdots = E_n = \mathbf{R}$ (*ou* $\mathbf{C}$). Alors $\mathscr{L}(E_r; F)$ s'identifie canoniquement à F d'après (4.1. Appendice A); donc $\mathscr{L}(E_s; \mathscr{L}(E_r; F))$ s'identifie à $\mathscr{L}(E_s; F)$, c'est-à-dire encore à F. Il résulte du paragraphe précédent que $D_{sr} f(a)$ et $D_{rs} f(a)$ s'identifient au même élément de F. *Dans ce cas particulier* on pourra donc écrire $D_{sr} f(a) = D_{rs} f(a)$, souvent noté $\frac{\partial^2 f}{\partial x^r\, \partial x^s}(a)$.

Si, en outre, $F = \mathbf{R}^m$ (ou $\mathbf{C}^m$) et si f est donnée par les composantes dans la base canonique $[f(x) = f_1(x), \ldots, f_m(x)]$, les composantes de la forme bilinéaire $D^2 f(a)$, définie sur $\mathbf{R}^n \times \mathbf{R}^n$ et à valeurs dans $\mathbf{R}^m$ sont $\frac{\partial^2 f_i}{\partial x^r\, \partial x^s}(a)$, $i = 1, \ldots, m$, $1 \leqslant r, s \leqslant n$, dans les bases canoniques. Plus généralement, les composantes dans les bases canoniques de $D^k f(a)$ sont $D_{r_1,\ldots,r_k}^k f_i(a)$, également notées $\frac{\partial^k f_i}{\partial x^{r_1} \ldots \partial x^{r_k}}(a)$, $i = 1, \ldots, m$, $1 \leqslant r_1, \ldots, r_k \leqslant n$.

3. Formule de Taylor

On se propose d'étendre à « l'ordre n » le théorème fondamental du calcul intégral :

$$f(a+h)-f(a)=\int_0^1 Df(a+t.h).h.\mathrm{d}t\,,$$

qui met en cause la différentielle d'ordre 1 de f.

Lemme. 3.1. — ***Si u est une fonction $n+1$ fois différentiable d'une variable réelle t et à valeurs dans un espace de Banach F, on a***

$$D\left[u(t)+(1-t).Du(t)+\cdots+\frac{1}{n\,!}(1-t)^n.D^n u(t)\right]=\frac{1}{n\,!}(1-t)^n.D^{n+1}u(t)\,.$$

PREUVE. — On applique la règle de Leibniz (2.9. chap. 1) en prenant $E=\mathbf{R}$, $F_1=\mathbf{R}$, $F_2=F$, $f(t)=(1-t)^k$, $g(t)=D^k u(t)$, $0\leqslant k\leqslant n$, et pour application bilinéaire continue $b:F_1\times F_2=\mathbf{R}\times F\to F$ le produit de $r\in\mathbf{R}$ par $y\in F$. La preuve en résulte par « télescopage ». □

Lemme. 3.2. — ***Si u est une fonction de classe C^{n+1} d'une variable réelle t, définie sur un ouvert contenant*** $[0,1]$ ***et à valeurs dans un espace de Banach F, on a***

$$u(1)-u(0)-u'(0)-\frac{1}{2}u''(0)-\cdots-\frac{1}{n\,!}u^{(n)}(0)=\int_0^1\frac{(1-t)^n}{n\,!}\cdot u^{(n+1)}(t)\cdot\mathrm{d}t\,,$$

où $u^{(k)}(t)=D^k u(t).1\in F$ est le vecteur dérivé d'ordre k.

PREUVE. — On applique le théorème fondamental du calcul intégral à la fonction

$$t\rightsquigarrow\left[u(t)+(1-t).Du(t)+\cdots+\frac{1}{n\,!}(1-t)^n.D^n u(t)\right].1\,,$$

qui est de classe C^1, et on utilise le lemme précédent. □

Affaiblissons les hypothèses de ce lemme :

Lemme. 3.3. — ***Si u est une fonction $n+1$ fois différentiable d'une variable réelle t, définie sur un ouvert contenant*** $[0,1]$, ***à valeurs dans un espace de Banach F, et telle qu'il existe une constante C majorant $\|D^{n+1}u(t)\|$ pour $0\leqslant t\leqslant 1$, on a***

$$\left\|u(1)-u(0)-u'(0)-\cdots-\frac{1}{n\,!}u^{(n)}(0)\right\|\leqslant\frac{C}{(n+1)\,!}\,.$$

PREUVE. — On applique le théorème 1.1. du chapitre 2 en prenant $[a,b]=[0,1]$,

$$f(t)=u(t)+(1-t)\cdot u'(t)+\cdots+\frac{1}{n\,!}(1-t)^n\cdot u^{(n)}(t)\quad\text{et}\quad g(t)=-C\cdot\frac{(1-t)^{n+1}}{(n+1)\,!}\,.$$

Le lemme 3.1. nous place en effet dans les hypothèses de ce théorème :

$$\|f'(t)\|=\left\|\frac{1}{n\,!}\cdot(1-t)^n\cdot u^{(n+1)}(t)\right\|\leqslant\frac{C}{n\,!}(1-t)^n=g'(t)\,.$$

Il en résulte $\|f(1)-f(0)\|\leqslant g(1)-g(0)$; c'est l'inégalité annoncée. □

Formule de Taylor avec reste intégral. 3. 4. — Soit f une application de classe C^{n+1} d'un ouvert U d'un e.v. normé E dans un espace de Banach F. Si le segment $[a, a + h]$ est contenu dans U on a

$$f(a + h) - f(a) - Df(a)\cdot h - \cdots - \frac{1}{n\,!} D^n f(a)\cdot(h)^n =$$

$$= \int_0^1 \frac{1}{n\,!} \cdot (1 - t)^n \cdot D^{n+1} f(a + t\cdot h)\cdot(h)^{n+1} \cdot \mathrm{d}t\,.$$

où $D^k f(a)\cdot(h)^k = D^k f(a)\cdot\underbrace{(h, \ldots, h)}_{k \text{ fois}}$.

PREUVE. — Supposons que $[a, a + h] = \{ a + th : 0 \leqslant t \leqslant 1 \}$ soit dans U. D'après le théorème 2.7. des applications composées, la fonction $u(t) = f(a + t.h)$ est de classe C^{n+1} pour $0 \leqslant t \leqslant 1$ et sa dérivée d'ordre k est, comme on le voit par récurrence sur k,

$$u^{(k)}(t) = D^k u(t).1 = D^k f(a + th).(h)^k\,.$$

Il suffit maintenant d'utiliser le lemme 3.2. □

Formule de Taylor avec reste de Lagrange. 3.5. — Dans le théorème précédent remplaçons l'hypothèse « f est de classe C^{n+1} » par les hypothèses suivantes : f est $n + 1$ fois différentiable et il existe une constante C qui majore $\| D^{n+1} f(x) \|$ pour tout $x \in U$. Alors

$$\left\| f(a + h) - f(a) - Df(a)\cdot h - \cdots - \frac{1}{n\,!} D^n f(a)\cdot(h)^n \right\| \leqslant \frac{C}{(n + 1)\,!} \cdot \| h \|^{n+1}\,.$$

PREUVE. — Reprenons la fonction $u(t) = f(a + th)$. Elle est $n + 1$ fois différentiable et

$$\| u^{(n+1)}(t) \| = \| D^{n+1} f(a + t.h).(h)^{n+1} \| \leqslant \| D^{n+1} f(a + t.h) \|.\| h \|^{n+1} \leqslant C.\| h \|^{n+1}\,.$$

Il suffit alors d'appliquer le lemme 3.3. □

Théorème. 3.6. — ***Soit $f : U \subset E \to F$ une application $n - 1$ fois différentiable dans l'ouvert U de l'e.v. normé E et à valeurs dans un espace de Banach F. Si f est n fois différentiable en $a \in U$, alors***

$$\left\| f(a + h) - f(a) - Df(a).h - \cdots - \frac{1}{n\,!} D^n f(a).(h)^n \right\| = o(\| h \|^n)\,.$$

PREUVE. (H. Cartan). — Pour $n = 1$, ce n'est rien d'autre que la définition de la différentielle en a. On va prouver le théorème par récurrence sur n en le supposant vrai pour $n - 1$. Considérons l'application

$$g(h) = f(a + h) - f(a) - \cdots - \frac{1}{n\,!} D^n f(a).(h)^n$$

et calculons sa différentielle. A cet effet cherchons la différentielle de $\varphi : h \rightsquigarrow D^k f(a).(h)^k$, où $k = 0, 1, \ldots, n$. Puisque $D^k f(a) \in \mathscr{L}^k(E; F)$, si $l \in E$ on a

$$D\varphi(h).l = \frac{\mathrm{d}}{\mathrm{d}t} D^k f(a).(h + tl)^k \Big|_{t=0} = D^k f(a).(l, k, \ldots, k) + D^k f(a).(k, l, k, \ldots, k) + \cdots +$$

$$+ D^k f(a).(k, \ldots, k, l)\,.$$

Comme, selon le théorème de Schwarz, $D^k f(a)$ est symétrique, tous les termes de la dernière somme sont égaux. Ainsi $D\varphi(h).l = k.D^k f(a).(l, h, \ldots, h)$. Si l'on note $D^k f(a).(h)^{k-1}$ l'application linéaire continue $l \rightsquigarrow D^k f(a).(l, h, \ldots, h)$, on en déduit

$$Dg(h) = Df(a + h) - Df(a) - \cdots - \frac{1}{(n - 1)\,!} D^n f(a).(h)^{n-1}\,.$$

Appliquons l'hypothèse de récurrence à Df, on en déduit $\| Dg(h) \| = o(\| h \|^{n-1})$; à tout $\varepsilon > 0$ correspond donc $\delta > 0$ tel que $\| h \| \leqslant \delta$ entraîne $\| Dg(h) \| \leqslant \varepsilon . \| h \|^{n-1}$. En utilisant le théorème des accroissements finis (1.4. chap. 2) on obtient $\| g(h) \| = \| g(h) - g(0) \| \leqslant \varepsilon . \| h \|^{n}$ et l'on a bien $\| g(h) \| = o(\| h \|^{n})$. □

Le même résultat aurait pu être déduit du théorème précédent, mais au prix de l'hypothèse trop forte : « f possède une différentielle d'ordre $n + 1$ bornée sur un voisinage de a ».

L'estimée du théorème 3.6. constitue la généralisation à l'ordre n de celle qui définit la différentielle d'ordre 1 :

$$\| f(a + h) - f(a) - Df(a).h \| = o(\| h \|).$$

C'est-à-dire que

$$I = f(a) + Df(a).h + \cdots + \frac{1}{n\,!} D^{n} f(a).(h)^{n}$$

approche $f(a + h)$ au voisinage de a à l'ordre n en $\| h \|$. On peut se demander si cette propriété caractérise I. Voici une première réponse.

Unicité de la formule de Taylor. 3.7. — Soient E et F deux espaces de Banach, f une application de classe C^{n} d'un ouvert U de E dans F. Supposons qu'il existe des applications $A_k \in \mathscr{L}^{k}(E ; F)$, $k = 0, .,., n$, symétriques telles qu'en $a \in U$ on ait

$$f(a + h) - \sum_{k=0}^{n} \frac{1}{k\,!} A_k.(h)^{k} = o(\| h \|^{n}).$$

Alors $D^{k} f(a) = A_k$ pour $k = 0, \ldots, n$.

Preuve. — Posons $B_k = \frac{1}{k\,!}(A_k - D^{k} f(a))$. La formule de Taylor 3.6. et la relation que nous supposons donnent, par soustraction membre à membre,

$$\sum_{k=0}^{n} B_k.(h)^{k} = o(\| h \|^{n}). \tag{3.8}$$

Faisons $h \to 0$ dans (3.8), on obtient $B_0 = 0$. Supposons que $B_0 = \cdots = B_k = 0$ pour $k < n$. Alors (3.8) entraîne

$$- B_{k+1}.(h)^{k+1} = \sum_{i=k+2}^{n} B_i.(h)^{i} + o(\| h \|^{n}).$$

$$\| B_{k+1} \| . \| h \|^{k+1} \leqslant \sum_{i=k+2}^{n} \| B_i \| . \| h \|^{i} + o(\| h \|^{n}).$$

Divisons les deux membres par $\| h \|^{k+1}$ et faisons $h \to 0$, on obtient $B_{k+1} = 0$. Les B sont donc tous nuls : c'est le résultat annoncé. □

Ce résultat facile permet souvent d'identifier les différentielles successives de f *lorsqu'on sait que f est de classe C^{n}*.

Non seulement le développement de Taylor à l'ordre n est unique, mais sa propriété d'approcher $f(a + h)$ à $o(\| h \|^{n})$ près le caractérise. C'est ce que montre le théorème suivant, dû à R. Abraham et J. Robbin (Transversal mappings and flows. Benjamin. 1967).

Réciproque de la formule de Taylor. 3.9. — Soit f une application d'un ouvert U d'un e.v. normé E dans un e.v. normé F. Faisons les hypothèses suivantes :

a) il existe des applications continues $a_j : U \to \mathscr{L}^{j}(E ; F)$, $j = 0, 1, \ldots, n$, telles que $a_j(x)$ soit symétrique pour tout x de U ;

b) soit $x \in U$ et $h \in E$ tels que $x + h$ soit contenu dans une boule ouverte de centre x appartenant à U. Posons

$$R_n(x, h) = f(x + h) - \sum_{k=0}^{n} \frac{1}{k\,!} a_k(x).(h)^k , \tag{3.10}$$

où $(h)^k = \underbrace{(h, \dots, h)}_{k \text{ fois}}$, et supposons que pour tout $x_0 \in U$, $\| R_n(x, h) \| / \| h \|^n \to \mathrm{o}$ si $(x, h) \to (x_0, 0)$.

Alors f est de classe C^n et $D^k f = a_k$ pour $k = 0, 1, \dots, n$.

PREUVE. (E. Nelson). — *Cas* $n = 1$.

Par hypothèse $f(x + h) = a_0(x) + a_1(x).h + R_1(x + h)$. Puisque $\| R_1(x, h) \| = \mathrm{o}(\| h \|)$. on a $a_0(x) = f(x)$, puis $f(x + h) = f(x) + a_1(x).h + \mathrm{o}(\| h \|)$. Donc $a_1 = Df$ et, puisque a_1 est continue, f est de classe C^1.

Nous allons démontrer le théorème pour n quelconque en le supposant vrai pour $n - 1$.

Puisque $\frac{1}{n\,!} a_n(x).(h)^n + R_n(x, h) = R_{n-1}(x, h)$ vérifie

$$\| R_{n-1}(x, h) \| / \| h \|^{n-1} \leqslant \left[\frac{1}{n\,!} \| a_n(x) \| + \| R_n(x, h) \| / \| h \|^n \right] \cdot \| h \| \to \mathrm{o}$$

si $h \to \mathrm{o}$, l'hypothèse de récurrence entraîne $a_0 = f, \dots, a_{n-1} = D^{n-1} f$. Ceci posé fixons $x \in U$ et choisissons $\varepsilon > \mathrm{o}$ assez petit pour que la boule de centre x et de rayon $2\,\varepsilon$ soit dans U. Prenons $y, z \in E$ tels que $\| y \|, \| z \| < \varepsilon$ et écrivons $f(x + y + z)$ de deux façons différentes :

$$f(x + y + z) = f(x + y) + Df(x + y).z + \cdots + \frac{1}{(n-1)\,!} D^{n-1} f(x + y).(z)^{n-1} + \\ + \frac{1}{n\,!} a_n(x + y).(z)^n + R_n(x + y, z) ,$$

puis

$$f(x + y + z) = f(x) + Df(x).(y + z) + \cdots + \frac{1}{(n-1)\,!} D^{n-1} f(x).(y + z)^{n-1} + \\ + \frac{1}{n\,!} a_n(x).(y + z)^n + R_n(x, y + z) .$$

Par soustraction membre à membre on en déduit

$$g_0(y) + g_1(y).z + \cdots + g_{n-1}(y).(z)^{n-1} + \\ + \left[\frac{1}{n\,!} a_n(x + y).(z)^n - \frac{1}{n\,!} a_n(x).(z)^n + R_n(x + y, z) - R_n(x, y + z) \right] = 0 \tag{3.11}$$

où, en tenant compte de la symétrie de la forme n-linéaire $a_n(x)$,

$$g_{n-1}(y).(z)^{n-1} = \frac{1}{(n-1)\,!} [D^{n-1} f(x + y) - D^{n-1} f(x) - a_n(x).y].(z)^{n-1} . \tag{3.12}$$

Imposons maintenant $\| z \| < \| y \|$. Le crochet de (3.11) vérifie $\| [\ ] \| / \| y \|^n \leqslant \frac{1}{n\,!} \| a_n(x+y) - a_n(x) \| + \| R_n(x + y, z) \| / \| z \|^n + 2^n . \| R_n(x, y + z) \| / \| y + z \|^n$.

Puisque a_n est continue, le premier terme du membre de droite tend vers zéro si $y \to 0$. Puisque $(x + y, z) \to (x, 0)$ si $y \to 0$, l'hypothèse *b* montre que le second terme tend vers zéro si $y \to 0$. Il en est de même pour le troisième. Il résulte alors de (3.11) que, si $y \to 0$,

$$\| g_0(y) + g_1(y).z + \cdots + g_{n-1}(y).(z)^{n-1} \| / \| y \|^n \to 0 .$$

En faisant $z = 0$ on obtient $g_0(y) = \mathrm{o}(\| y \|^n)$, donc

$$g_1(y).z + \cdots + g_{n-1}(y).(z)^{n-1} = \mathrm{o}(\| y \|^n) .$$

Changeons z en $z/2$ dans cette relation, puis multiplions les deux membres par 2 et retranchons membre à membre la relation obtenue de la relation initiale ; multiplions enfin par 2 les deux membres. On obtient

$$g_2(y).(z)^2 + \cdots + \left(2 - \frac{1}{2^{n-3}}\right) \cdot g_{n-1}(y) \cdot (z)^{n-1} = o(\| y \|^n).$$

Réitérons le procédé en chassant $g_2(y).(z)^2$... etc. On obtient finalement

$$g_{n-1}(y).(z)^{n-1} = o(\| y \|^n).$$

D'après (3.12) cela entraîne $\| D^{n-1} f(x+y) - D^{n-1} f(x) - a_n(x).y \| / \| y \| \to 0$ si $y \to 0$. Par définition de la différentielle, cela signifie $D^n f(x) = a_n(x)$ et, puisque a_n est continue, f est de classe C^n. □

Application. 3.13. — Ce résultat est un critère commode pour décider si une application est de classe C^n.

Considérons, par exemple, l'inversion $\mathfrak{I} : u \rightsquigarrow u^{-1}$ vue en 2.8. Si $x \in \mathrm{GL}(E; E)$, $h \in \mathscr{L}(E; E)$ et $\| x^{-1} \circ h \| < 1$, on peut écrire $\mathfrak{I}(x+h) = \sum_{k=0}^{n} (-x^{-1} \circ h)^k \circ x^{-1} + R_n(x, h)$, où

$$R_n(x, h) = \sum_{n+1}^{\infty} (-x^{-1} \circ h)^k \circ x^{-1}.$$

Posons

$$A_k(x).(h_1, \ldots, h_k) = \sum_s (-1)^k.(x^{-1} \circ h_{s(1)} \circ x^{-1}) \circ \cdots \circ (x^{-1} \circ h_{s(k)} \circ x^{-1}),$$

$h_1, \ldots, h_k \in \mathscr{L}(E; F)$ et où la sommation s'étend à toutes les permutations s de $\{1, 2, \ldots, k\}$.

Evidemment $A_k(x) \in \mathscr{L}^k(E; F)$ est symétrique et, puisque $x \rightsquigarrow x^{-1}$ est continue, $x \rightsquigarrow A_k(x)$ est continue. Enfin $R_n(x, h)$ est de la forme $\rho_n(x, h).(h)^n$ où $\rho_n(x, h) \in \mathscr{L}^n(E; F)$ dépend continûment de (x, h) et vérifie $\rho_n(x, o) = o$. Puisque $\mathfrak{I}(x+h) = \sum_{k=0}^{n} A_k(x).(h)^k + R_n(x, h)$, le théorème 3.9 montre que $\mathfrak{I}$ est de classe C^n et que $D^k \mathfrak{I} = A_k$, $k = o, \ldots, n$, pour tout n.

4. Série de Taylor. Analyticité

Définition. 4.1. — Supposons f de classe C^∞. Il est alors loisible de former la série $\sum_0^\infty \frac{1}{k!} D^k f(a).(h)^k$. On l'appelle la série de Taylor de f en a [si $a = o$ on l'appelle quelquefois la série de Mac-Laurin de f].

Si, étant donné un nombre $r > 0$, la série de Taylor en a converge pour tout $h \in E$ de norme $\| h \| < r$, et si sa somme est égale à $f(a+h)$, on dit que f est analytique en a.

Puisque, pour tout entier $n > o$, on a $f(a+h) = \sum_0^n \frac{1}{k!} D^k f(a).(h)^n + R_n(a, h)$, f est analytique en a si, et seulement si, le reste $R_n(a, h)$ tend vers zéro avec $1/n$.

Sans s'attarder sur l'étude des applications analytiques, qui constituerait l'objet d'un cours tout entier, notons que la série de Taylor peut converger sans que sa somme soit égale à $f(a+h)$. Voici des exemples intéressants pour eux-mêmes.

Fonctions de classe C^∞ et non analytiques. 4.2. — Considérons la fonction $f : \mathbf{R} \to \mathbf{R}$ définie par

$$f(x) = \begin{cases} e.e^{-1/x^2} & \text{si } x > 0, \\ 0 & \text{ailleurs}. \end{cases}$$

Nous allons montrer qu'elle est C^∞. Evidemment la question ne se pose pas pour $x < 0$, car alors $D^n f(x) = 0$ pour tout n. Si $x > 0$, on voit par récurrence sur n que $D^n f(x) = x^{-3n}.Q_n(x).f(x)$, où $Q_0(x) = 1$, $Q_1(x) = 2$, $Q_2(x) = 4 - 6x^2$ et $Q_n(x)$, pour $n \geqslant 1$, est un polynôme de degré $2n - 2$ satisfaisant la relation de récurrence

$$Q_{n+1}(x) = (2 - 3nx^2).Q_n(x) + x^3.Q_n'(x).$$

La fonction f est donc C^∞ pour $x > 0$.

Montrons, par récurrence sur n, que $D^n f(0) = 0$. C'est évident si $n = 0$. Supposons-le établi pour n. Alors

$$D^{n+1} f(0) = \lim_{\substack{x > 0 \\ \rightarrow}} D^n f(x)/x = \lim_{\substack{x > 0 \\ \rightarrow}} x^{-3n-1}.Q_n(x).e^{-1/x^2} = 0.$$

En particulier la dérivée n-ième de f est continue partout, y compris pour $x = 0$. La fonction f est donc C^∞.

La série de Mac-Laurin de f converge évidemment vers zéro puisque $D^n f(0) = 0$. Sa somme n'est donc pas $f(h)$: f n'est pas analytique en zéro.

A partir de f on obtient de nombreuses fonctions analogues.

Puisque f est C^∞ et croissante pour $x > 0$, la fonction $g(x) = f[1 - f(x)]$ est de classe C^∞, décroissante sur $\mathbf{R}$ et telle que $g(x) = 1$ pour $x \leqslant 0$, $g(x) = 0$ pour $x \geqslant 1$.

Prenons $a < b$ et posons $h(x) = g(a - x).g(x - b)$. Cette fonction est C^∞, nulle hors de $]a - 1, b + 1[$, positive ailleurs et égale à 1 sur $[a, b]$. On l'appelle une fonction « plateau » [dessiner son graphe].

5 FONCTION EXPONENTIELLE. ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES A COEFFICIENTS CONSTANTS

Ce chapitre est consacré à l'exponentielle $\exp(A)$ d'un endomorphisme A d'un espace de Banach.

On sait que la fonction exponentielle usuelle $t \in \mathbf{R} \rightsquigarrow e^{at}$, où $a \in \mathbf{R}$, vérifie l'équation différentielle $y' = a.y$. Cette propriété s'étend à l'application $t \in \mathbf{R} \rightsquigarrow \exp(tA)$. C'est la clef de la résolution des équations différentielles linéaires à coefficients constants.

Dans tout ce chapitre E est un espace de Banach réel ou complexe. Pour simplifier on posera $\mathcal{L}(E; E) = \mathrm{End}\,(E)$, $\mathrm{GL}\,(E; E) = \mathrm{GL}\,(E)$ et, si $A, B \in \mathrm{End}\,(E)$, $A \circ B = A.B$. Enfin, si aucune ambiguïté n'en résulte, $\mathrm{id}_E = 1$.

1. Définitions de l'exponentielle

Première définition. 1.1. — Inspirons-nous de la propriété bien connue de l'exponentielle usuelle : $e^x = \sum_0^\infty \frac{x^n}{n!}$ pour $x \in \mathbf{R}$. Si $A \in \mathrm{End}\,(E)$ formons la suite $S_n = 1 + A + \cdots + \frac{A^n}{n!}$.

Théorème. — ***La suite S_n converge uniformément sur chaque borné de* End (E) *et sa limite* $\exp(A)$ *vérifie* $\| \exp(A) \| \leqslant \exp(\| A \|)$.**

PREUVE. — Soient p et q deux entiers positifs. D'après (2.2. et 2.4. Appendice A) on a

$$\| S_{p+q} - S_p \| \leqslant \frac{\| A \|^{p+1}}{(p+1)!} + \cdots + \frac{\| A \|^{p+q}}{(p+q)!}, \tag{1.2}$$

qui tend vers zéro si $p, q \to \infty$, car c'est le reste tronqué de la série convergente $\sum_0^\infty \| A \|^n/n!$. La suite S_n est donc une suite de Cauchy et elle converge dans l'espace complet End (E) vers une limite que l'on note $\exp(A)$.

Faisons tendre q vers l'infini dans (1.2). Si $\| A \| \leqslant R =$ constante, on en déduit $\| \exp(A) - S_p \| \leqslant \sum_{p+1}^\infty R^k/k!$ et la convergence est uniforme sur la boule de centre o et de rayon R de End (E).

Enfin $\| S_n \| \leqslant 1 + \cdots + \| A \|^n/n!$ entraîne $\| \exp(A) \| \leqslant \exp \| A \|$. □

Seconde définition. 1.3. — Pour tout $A \in \mathrm{End}\,(E)$ on a $\exp(A) = \lim_{n=\infty} \left(1 + \frac{A}{n}\right)^n$.

PREUVE. — D'après la formule du binôme

$$\left(1 + \frac{A}{n}\right)^n = \sum_{k=0}^{k} C_n^k \cdot \frac{1}{n^k} \cdot A^k,$$

où $C_n^k = \frac{n!}{k!\,(n-k)!}$. Remarquons que $\frac{1}{k!} - \frac{C_n^k}{n^k} \geqslant 0$. On en déduit

$$\left\| \exp(A) - \left(1 + \frac{A}{n}\right)^n \right\| = \left\| \sum_0^n \left(\frac{1}{k\,!} - \frac{C_n^k}{n^k}\right) \cdot A^k + \sum_{n+1}^{\infty} \frac{A^k}{k\,!} \right\|$$

$$\leqslant \sum_0^n \left(\frac{1}{k\,!} - \frac{C_n^k}{n^k}\right) \cdot \| A \|^k + \sum_{n+1}^{\infty} \frac{\| A \|^k}{k\,!} = \exp \| A \| - \left(1 + \frac{\| A \|}{n}\right)^n \to 0$$

si $n \to \infty$, d'après la formule classique $e^x = \lim\limits_{n=\infty} \left(1 + \frac{x}{n}\right)^n$ pour $x \in \mathbf{R}$. □

2. Propriétés de l'exponentielle

Continuité. 2.1. — L'application exp est continue; elle est même localement lipschitzienne : $\| \exp(A) - \exp(B) \| \leqslant e^M . \| A - B \|$, si M majore $\| A \|$ et $\| B \|$.

PREUVE. — De (2.2. et 2.4. Appendice A) et de l'identité

$$A^n - B^n = A^{n-1}.(A - B) + A^{n-2}.(A - B).B + \cdots + A.(A - B).B^{n-2} + (A - B).B^{n-1}$$

on déduit $\| A^n - B^n \| \leqslant n.[\max(\| A \|, \| B \|)]^{n-1}.\| A - B \|$. Le théorème en résulte. □

On prouvera ultérieurement que exp est de classe C^∞.

Commutation avec la conjugaison. 2.2. — Si B est un isomorphisme de E et $A \in \mathrm{End}(E)$, on a $\exp(B^{-1}.A.B) = B^{-1}.\exp(A).B$.

PREUVE. — La conjugaison $X \rightsquigarrow B^{-1}.X.B$ est continue et commute avec $X \rightsquigarrow X^n$ pour tout entier $n \geqslant 0$.

Exponentielle d'une somme de deux opérateurs qui commutent. 2.3.

Soient $A, B \in \mathrm{End}(E)$ tels que $A.B = B.A$. Alors $\exp(A + B) = \exp(A).\exp(B)$. En particulier $\exp(A) \in \mathrm{GL}(E)$ et son inverse est $\exp(-A)$.

PREUVE. — Selon (1.3.), et puisque $A.B = B.A$, on a

$$\exp(A) \cdot \exp(B) = \lim_{n=\infty} \left(1 + \frac{A}{n}\right)^n \cdot \lim_{n=\infty} \left(1 + \frac{B}{n}\right)^n = \lim_{n=\infty} \left[\left(1 + \frac{A}{n}\right) \cdot \left(1 + \frac{B}{n}\right)\right]^n.$$

Posons $v = \left(1 + \frac{A}{n}\right) \cdot \left(1 + \frac{B}{n}\right)$, $u = 1 + \frac{A + B}{n}$ et utilisons l'estimée de 2.1. :

$$\| v^n - u^n \| \leqslant n.[\max(\| u \|, \| v \|)]^{n-1}.\| v - u \| \leqslant$$

$$\leqslant \left[1 + \frac{\| A \| + \| B \|}{n} + \frac{\| AB \|}{n^2}\right]^{n-1} \cdot \frac{\| AB \|}{n}.$$

Si n est assez grand $\| AB \| \leqslant n(\| A \| + \| B \|)$, donc $[\ ]^{n-1} \leqslant \left(1 + \frac{2 \| A \| + 2.\| B \|}{n}\right)^n$ qui tend vers $\exp 2(\| A \| + \| B \|)$ si $n \to \infty$. On en déduit $\lim\limits_{n=\infty} \| v^n - u^n \| = 0$, donc

$$\exp(A).\exp(B) = \exp(A + B).$$

Si l'on remarque que $\exp(0) = 1$, on voit que $\exp(-A)$ est l'inverse de $\exp(A)$. □

MISE EN GARDE. — Le résultat est en défaut si $A.B \neq B.A$ [prendre $E = \mathbf{R}^2$, A de matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et B de matrice $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$].

Remarque. — Toute exponentielle est un carré. En effet $\exp(A) = [\exp(A/2)]^2$.

Théorème. 2.4. — ***Soit E un espace de Banach de dimension finie d. Désignons par* dét (A) *et* tr (A) *le déterminant et la trace de* $A \in$ End (E). *Alors* dét $[\exp(A)] = \exp[\text{tr}\,(A)]$.**

PREUVE. — Choisissons une base de E. Le déterminant de A est un polynôme en les coefficients de sa matrice, dét est donc une application continue. La définition 1.3. entraîne

$$\text{dét}\,[\exp(A)] = \text{dét} \lim_{n=\infty} \left(1 + \frac{A}{n}\right)^n = \lim_{n=\infty} \left[\text{dét}\left(1 + \frac{A}{n}\right)\right]^n.$$

Etudions dét $\left(1 + \frac{A}{n}\right)$ lorsque n est grand. Pour cela, regardons la matrice $d \times d$ de $B \in$ End (E) comme un point de E^d, en considérant chaque ligne $B_1, \ldots, B_d$ comme un élément de E. Alors dét : $E^d \to \mathbf{R}$ (ou $\mathbf{C}$) est une application d-linéaire continue. D'après (2.3. chap. 1) elle est différentiable et si $h = (h_1, \ldots, h_d) \in E^d$ on a

$$D(\text{dét})\,(B).h = \sum_{k=1}^{d} \text{dét}\,(B_1, \ldots, B_{k-1}, h_k, B_{k+1}, \ldots, B_d).$$

Dans le cas qui nous occupe $B = 1$, $h = A/n$, donc

$$\text{dét}\left(1 + \frac{A}{n}\right) = \text{dét}\,1 + D(\text{dét})\,(1) \cdot \frac{A}{n} + o\left(\frac{1}{n}\right) = 1 + \frac{1}{n}\text{tr}\,A + o\left(\frac{1}{n}\right).$$

Par conséquent

$$\text{dét}\,[\exp(A)] = \lim_{n=\infty} \left[1 + \frac{1}{n}\text{tr}\,(A) + o\left(\frac{1}{n}\right)\right]^n = \lim_{n=\infty} \left(1 + \frac{1}{n} \cdot \text{tr}\,A\right)^n = \exp[\text{tr}\,(A)]. \quad \square$$

Corollaire. 2.5. — ***Si E est un espace de Banach réel de dimension finie, alors* dét $[\exp(A)] > 0$. *En particulier* dét : Exp$(E) \to$ GL (E) *n'est pas surjective.***

Théorème. 2.6. — ***Soit A un endomorphisme d'un espace de Banach E. L'application* $f : \mathbf{R} \to$ GL (E) *définie par* $f(t) = \exp(t.A)$ *est de classe* C^∞ *et* $f'(t) = A.f(t)$ *pour tout t.***

PREUVE. — Montrons d'abord que f est différentiable en zéro. Puisque $f(0) = \text{id}_E$, on a

$$\| f(s) - f(0) - s.A \| \leqslant \sum_{n=2}^{\infty} |s|^n.\|A\|^n/n\,! = \exp(|s|.\|A\|) - 1 - |s|.\|A\| = o(|s|).$$

Donc $Df(0)$ existe et $f'(o) = Df(0).1 = A$.

Evidemment $t.A$ et $s.A$ commutent quels que soient les réels t et s. D'après (2.3.) on a donc $f(t + s) = f(t).f(s)$. Le membre de droite est dérivable en s pour $s = 0$, le membre de gauche l'est donc aussi et $f'(t) = f(t).f'(0) = f(t).A$. Puisque A et $\exp(t.A)$ commutent cela s'écrit encore $f'(t) = A.f(t)$. Ainsi f est de classe C^1.

La relation $f' = A.f$ montre que si f est de classe C^n il en est de même de f' ; f est donc de classe C^∞.

3. Groupe à un paramètre d'automorphismes linéaires

Définition. 3.1. — On appelle groupe à un paramètre d'automorphismes linéaires d'un espace de Banach E, un homomorphisme h du groupe additif des réels dans le groupe GL (E). Si l'application h est continue, on dit que le groupe est continu.

EXEMPLE. — $f : t \rightsquigarrow \exp(t.A)$, où $A \in$ End (E). En effet $f(0) = \text{id}_E$ et, d'après (2.6.), $f(s + t) = f(s).f(t)$ pour tous $t, s \in \mathbf{R}$.

Montrons que cet exemple est le cas général.

Théorème. 3.2. — ***Tout groupe continu h à un paramètre d'automorphismes linéaires d'un espace de Banach E est de la forme*** $t \rightsquigarrow \exp(t.A)$, ***où*** $A \in \operatorname{End}(E)$.

Il y a donc correspondance biunivoque entre ces groupes et les endomorphismes A. On dit que A est le générateur du groupe $\exp(t.A)$.

PREUVE. — Puisque h est continu on peut intégrer chacun des deux membres de $h(s+t) = h(s).h(t)$ entre 0 et $a > 0$. Effectuons le changement de variable $s = u - t$ dans le membre de gauche et utilisons la propriété d de (6.4. chap. 2); on obtient

$$\int_t^{t+a} h(u).\mathrm{d}u = h(t) \circ \int_0^a h(s).\mathrm{d}s\,.$$

Puisque h est continu on peut choisir $a > 0$ assez proche de 0 pour que $L = a^{-1} . \int_0^a h(s) . \mathrm{d}s$ soit proche de $h(0) = 1$. Il résulte alors de (4. 1. chap. 3) que L est inversible ; comme L commute manifestement avec $h(t)$, l'égalité ci-dessus s'écrit $h(t) = L^{-1} \circ \int_t^{t+a} h(u).\mathrm{d}u$. D'après (6.5. chap. 2) le membre de droite est dérivable par rapport à t, il en est donc de même de h et, si l'on pose $C = L^{-1} \circ [h(a) - 1]$, on obtient $h'(t) = C.h(t)$, où C commute avec $h(t)$. Cela permet de calculer la dérivée de $g(t) = h(t) \circ \exp(-C.t)$. En utilisant 2.6. on obtient

$$g'(t) = h'(t).\exp(-C.t) + h(t).[-C.\exp(-C.t)] = 0\,.$$

D'après (2. chap. 2) g est donc constante et égale à $g(0) = h(0).\exp(0) = \mathrm{id}_E$. Par conséquent $h(t) = \exp(C.t)$.

EXEMPLES. — *a*) Pour $A = \mathrm{id}_E$ on trouve le groupe $t \rightsquigarrow e^t.\mathrm{id}_E$ des homothéties de centre O et de rapport positif.

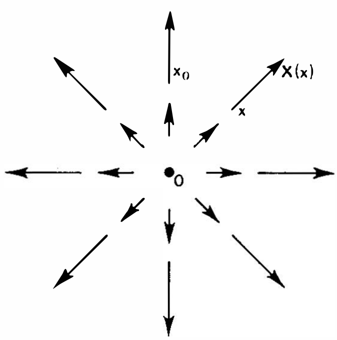


Fig. 1.

b) Pour $A \in \operatorname{End}(\mathbf{R}^2)$ dont la matrice est $\left(\begin{array}{c|c} 0 & -1 \\ \hline 1 & 0 \end{array}\right)$, on a $A^2 = -\mathrm{id}_E$. On en déduit que $\exp(tA)$ a pour matrice $\left(\begin{array}{c|c} \cos t & -\sin t \\ \hline \sin t & \cos t \end{array}\right)$. On obtient le groupe des rotations de centre O.

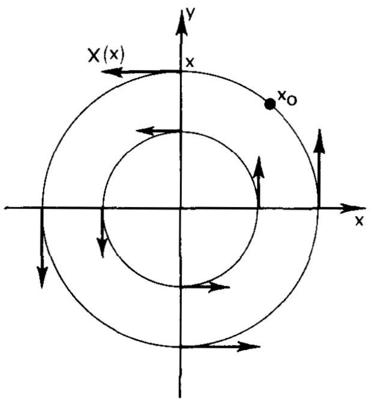


Fig. 2.

c) Pour $A \in \mathrm{End}(\mathbf{R}^2)$ dont la matrice est $\left(\begin{array}{c|c} 0 & 1 \\ \hline 1 & 0 \end{array}\right)$, on a $A^2 = \mathrm{id}_E$. On en déduit que $\exp(t.A)$ a pour matrice $\left(\begin{array}{c|c} \mathrm{ch}\, t & \mathrm{sh}\, t \\ \hline \mathrm{sh}\, t & \mathrm{ch}\, t \end{array}\right)$. On obtient le groupe des rotations hyperboliques.

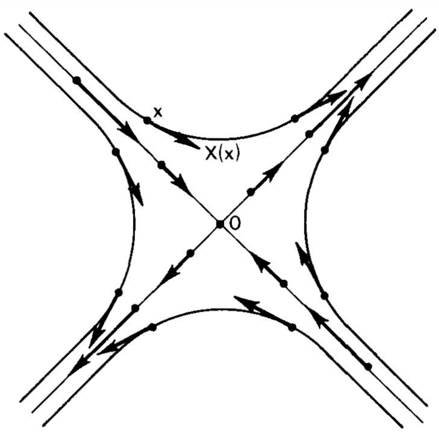


Fig. 3.

Les figures montrent, dans chacun de ces cas, l'orbite $\{ \exp(t.A).x_0 : t \in \mathbf{R} \}$ d'un point x_0 de $E = \mathbf{R}^2$. Si l'on interprète t comme le temps, la vitesse X à l'instant t ne dépend que de la position $x = \exp(t.A).x_0$ à cet instant :

$$X(x) = \frac{\mathrm{d}}{\mathrm{d}t} \exp(t.A).x_0 = A.\exp(t.A).x_0 = A.x\,.$$

4. Equations différentielles linéaires homogènes à coefficients constants

Définition. 4.1. — On appelle équation différentielle homogène du premier ordre à coefficients constants une expression de la forme

$$\frac{dx}{dt} = A.x\,, \tag{4.2}$$

où A est un endomorphisme d'un espace de Banach E.

Une solution de cette équation est une application différentiable f d'un intervalle I de $\mathbf{R}$ dans E, vérifiant $f'(t) = A.f(t)$ pour tout $t \in I$.

Il y a donc correspondance biunivoque entre l'ensemble des équations que l'on vient de définir et l'ensemble des endomorphismes A de E.

Exemple. — Si $E = K^n$ ($K = \mathbf{R}$ ou $\mathbf{C}$), l'équation $x' = A.x$ est équivalente à un système de n équations scalaires à n inconnues :

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_i^j.x_j\,, \qquad i = 1, 2, ..., n\,.$$

Remarques. 4.3. — *a*) Désignons par S_I l'ensemble des solutions $f : I \to E$ de (4.2). Il est clair que si $f, g \in S_I$ et que si k est un scalaire, $f + g \in S_I$ et $k.f \in S_I$.

L'ensemble des solutions *définies sur un même intervalle* I est donc un espace vectoriel. Si la dimension de S_I est finie, on aura résolu l'équation (4.2), c'est-à-dire trouvé toutes ses solutions, si l'on exhibe une base de S_I. Une telle base s'appelle un *système fondamental de solutions.*

b) Une solution f est continue, car on l'a supposée différentiable. Par conséquent $f' = A.f$ est continue et f est donc de classe C^1. On voit par récurrence que f est de classe C^∞.

Théorème fondamental. 4.4. — ***Soient E un espace de Banach, $A \in$ End (E), $x_0 \in E$, I un intervalle (ouvert ou fermé, fini ou infini) de R, $t_0 \in I$. Alors il existe une application différentiable unique $f : I \to E$ vérifiant $\frac{df}{dt} = A.f(t)$ pour $t \in I$ et $f(t_0) = x_0$. A savoir***

$$f(t) = \exp[(t - t_0)A].x_0\,.$$

On dit que f est la solution définie sur I de $x' = A.x$, qui vérifie la condition initiale $f(t_0) = x_0$.

Preuve. — Définissons l'application $f : I \to E$ par $f(t) = \exp[(t - t_0).A].x_0$. Evidemment $f(t_0) = x_0$, et le théorème 2.6. montre que f vérifie $x' = A.x$. L'existence est démontrée.

Démontrons l'unicité. Si g est un second candidat, $y = f - g$ vérifie $y'(t) = A.y(t)$ pour $t \in I$, et $y(t_0) = 0$. Intégrons de t_0 à $s \in I$, en utilisant la propriété d de (6.4. chap. 1) :

$$y(s) = y(s) - y(t_0) = \int_{t_0}^{s} A.y(t).dt = A\left(\int_{t_0}^{s} y(t).dt\right).$$

On en déduit $\| y(s) \| \leqslant \| A \| . \int_{t_0}^{s} \| y(t) \|.dt$, qui permet d'estimer la dérivée de

$$F(s) = \exp(- \| A \|.s) . \int_{t_0}^{s} \| y \|.dt\,.$$

On trouve $F'(s) \leqslant 0$ pour tout $s \in I$. Puisque $F(s) \geqslant 0$ et $F(t_0) = 0$ il en résulte $F(s) = 0$, donc $y(s) = 0$, pour $s \in I$. On a bien $f = g$.

Conséquences. 4.5. — *a*) Prenons $I = \mathbf{R}$. Il existe donc une solution f de $x' = A.x$ définie pour tout $t \in \mathbf{R}$ et vérifiant $f(t_0) = x_0$ pour t_0 et x_0 arbitrairement fixés.

Si maintenant I est un intervalle quelconque de $\mathbf{R}$ contenant t_0, il est évident que la restriction $f|_I$ de f à I vérifie l'équation différentielle $x' = A.x$ et la condition initiale $f|_I(t_0) = x_0$. D'après l'unicité elle coïncide donc avec la solution de (4.4). On exprime cela en disant que les solutions $t \rightsquigarrow f(t)$ d'une équation $x' = A.x$ peuvent être prolongées à $\mathbf{R}$ tout entier. On dit encore que les solutions maximales (non prolongeables) de $x' = A.x$ sont définies sur $\mathbf{R}$ tout entier.

b) Pour chaque $t_0 \in I$ l'application $f \rightsquigarrow f(t_0) = x_0$ de l'espace vectoriel S_I des solutions de $x' = A.x$ dans l'espace E est un isomorphisme d'espace vectoriel.

En effet, l'application qui à $x_0 \in E$ associe $\exp[(t - t_0).A].x_0$ est injective [unicité de la solution], surjective [existence de la solution] et évidemment linéaire.

Il en résulte que la dimension de S_I est égale à celle de E.

Remarque. 4.6. — Une fois obtenue ce qu'on appelle la solution générale $\exp(t.A).x$ de l'équation $x' = A.x$ on pourrait croire la résolution achevée. Il n'en est rien, car la série $\exp(t.A) = \sum (t.A)^n/n\,!$ peut être malaisée à calculer. Si l'on garde la solution sous cette forme, il y a bien des questions auxquelles il est difficile de répondre : $\exp(t.A).x$ est-elle périodique ? Reste-t-elle bornée si $t \to +\infty$?

Nous allons montrer qu'il est possible de remplacer $\exp(t.A)$ par une expression plus simple, du moins si E est de dimension finie.

5. Calcul explicite des solutions

Dans tout ce paragraphe E est un espace de dimension finie n.

L'endomorphisme A est diagonalisable. 5.1. — E possède donc une base $\{e_k\}$ de vecteurs propres de A, de valeurs propres correspondantes a_k : $A(e_k) = a_k.e_k$. Il en résulte $A^n(e_k) = (a_k)^n.e_k$ pour tout n, donc $\exp(t.A).e_k = \exp(t.a_k).e_k$.

Les $\exp(t.a_k).e_k$ sont donc solutions de $x' = A.x$. Comme ils sont au nombre de $n = \dim E$ et qu'ils sont linéairement indépendants, d'après (*b*.4.5.) ils forment une base de l'espace des solutions.

Dans la pratique, les valeurs propres a_k étant déterminées, on cherchera les e_k par la méthode des coefficients indéterminés en écrivant que $\exp(t.a_k).e_k$ vérifie $x' = A.x$.

L'endomorphisme A est nilpotent. 5.2. — Rappelons que $A \in \operatorname{End}(E)$ est nilpotent d'indice N si $A^N = 0$, tandis que $A^p \neq 0$ pour $p = 1, 2, \ldots, N - 1$. S'il en est ainsi $\exp(t.A)$ se réduit au polynôme $1 + tA + \cdots + (tA)^{N-1}/(N-1)\,!$. La solution générale $\exp(t.A).x$ de $x' = A.x$ sera donc un vecteur dont les composantes sont des polynômes en t de degré inférieur à N. On cherchera les composantes de ce vecteur par la méthode des coefficients indéterminés.

Exemple d'opérateur nilpotent. — L'ensemble E des polynômes en u à coefficients dans $\mathbf{R}$ de degré $< n$ est un espace vectoriel de dimension n sur $\mathbf{R}$. Comme la dérivation par rapport à u diminue le degré d'un polynôme, $p' \in E$ si $p \in E$. L'application D, qui associe à chaque polynôme p de E son polynôme dérivé p', est donc un endomorphisme de E. Comme la dérivée n-ième d'un polynôme de degré $< n$ est nulle, D est nilpotent d'indice n. Par conséquent

$$\exp(t.D) = 1 + t.D + \cdots + (t.D)^{n-1}/(n-1)\,!$$

Identifions $\exp(tD)$ en faisant opérer le membre de droite sur $p \in E$; on obtient

$$p(u) + t \cdot p'(u) + \cdots + \frac{t^{n-1}}{(n-1)\,!} \cdot p^{(n-1)}(u)\,.$$

D'après la formule de Taylor (3. chap. 4), ce n'est autre que $p(u + t)$. Ainsi $\exp(t.D)$ est l'opérateur de translation $T_t : p \rightsquigarrow p(t + .)$.

Le cas général sur le corps $\mathbf{C}$. 5.3. — E est dorénavant un espace vectoriel complexe de dimension n. Par conséquent, d'après le théorème de d'Alembert-Gauss, le polynôme caractéristique $p(\lambda) = \text{dét}(A - \lambda.1)$ de $A \in \text{End}(E)$ se factorise : $p(\lambda) = (\lambda - \lambda_1)^{r_1} \ldots (\lambda - \lambda_m)^{r_m}$. L'endomorphisme A possède n valeurs propres $\lambda_1, \ldots, \lambda_m$ à condition de les compter avec leur ordre de multiplicité $r_1, \ldots, r_m$.

On trouvera à l'appendice E la preuve du résultat suivant : soit E_i le noyau de $(A - \lambda_i.1)^{r_i}$. Alors E_i est de dimension r_i et E est la somme directe des E_i.

Tirons quelques conséquences de ce résultat :

a) $x \in E_i$ entraîne $A(x) \in E_i$. En effet $(A - \lambda_i.1)^{r_i}.(A.x) = A.[(A - \lambda_i.1)^{r_i}.x] = 0$. L'opérateur A laisse donc E_i invariant et il induit un opérateur $A_i = A\,|_{E_i}$ sur E_i ;

b) $A = \sum A_i$, les A_i commutent, $A_i(E_k) = 0$ si $i \neq k$, et $(A_i - \lambda_i.1)^{r_i} = 0$.

La preuve est évidente.

Ceci posé, nous allons expliciter la solution générale $\exp(tA).x$ de l'équation $x' = A.x$. Puisque $A = \sum A_i$ et que les A_i commutent, d'après 2.3. on a

$$\exp(tA) = \exp(t.A_1) \ldots \exp(t.A_m)\,.$$

Décomposons $x \in E$ selon les E_i : $x = \sum x_i$. La factorisation de $\exp(t.A)$ que nous venons de mettre en évidence et le fait que, si $k \neq i$, $A_k(x_i) = 0$, donc $\exp(tA_k).x_i = x_i$, entraînent $\exp(tA).x = \sum_i \exp(t.A_i).x_i$.

Écrivons $\exp(t.A_i)$ sous la forme $\exp[t(A_i - \lambda_i.1) + t.\lambda_i.1]$. Puisque 1 et $A - \lambda_i.1$ commutent cela s'écrit encore $\exp(t.A_i) = \exp(t.\lambda_i).\exp[t(A_i - \lambda_i.1)]$. Mais $A_i - \lambda_i.1$ est nilpotent d'indice r_i ; d'après 5.2., $\exp[t.(A_i - \lambda_i.1)]$ est donc un polynôme en t à coefficients dans $\text{End}(E)$ et de degré $< r_i$. Notons-le $P_i(t)$.

En fin de compte $\exp(tA).x = \sum_i e^{t.\lambda_i}.P_i(t).x_i$. On a donc démontré que la solution générale de $x' = A.x$ est la somme de m vecteurs de la forme $e^{t.\lambda_i}.X_i(t)$, où $X_i(t)$ est un vecteur dont les composantes sont des polynômes en t de degré inférieur à l'ordre de multiplicité r_i de la valeur propre λ_i.

Corollaire. — ***Si $a_1, \ldots, a_m$ sont des nombres complexes distincts et $p_1(t), \ldots, p_m(t)$ des polynômes en t, l'égalité $e^{t.a_1} p_1(t) + \cdots + e^{t.a_m}.p_m(t) = 0$ pour tout $t \in \mathbf{R}$ entraîne $p_1 = \cdots = p_m = 0$.***

Méthode pratique de résolution. — Elle consiste à déterminer les valeurs propres λ_i avec leur ordre de multiplicité r_i, puis à chercher les polynômes $X_i(t)$ par la méthode des coefficients indéterminés. A cette fin on prend le polynôme $X_i(t)$, de degré $r_i - 1$ et à valeurs dans E, le plus général et on exprime que $e^{t.\lambda_i}\, X_i(t)$ vérifie $x' = A.x$. Après simplification par $e^{t.\lambda_i}$ des équations ainsi obtenues, il reste des équations polynomiales en t dont tous les coefficients peuvent être séparément égalés à zéro, car ces équations sont des identités. On obtient ainsi un système d'équations linéaires qui permet de déterminer ceux des coefficients des X_i qui ne peuvent demeurer arbitraires [a priori r_i d'entre eux sont arbitraires puisque $\dim E_i = r_i$].

Le cas général sur le corps des réels. 5.4. — Ici E est un espace vectoriel réel de dimension n et $A \in \text{End}(E)$.

Comme un polynôme de degré n ne possède pas nécessairement n racines réelles, le raisonnement du 5.3. ne s'applique plus. Nous allons nous ramener au cas complexe.

Commençons par construire un espace vectoriel complexe à partir de E. Définissons sur le produit $E \times E$ une addition et un produit par le nombre complexe $a + i.b$ par

$$(x, y) + (x', y') = (x + x', y + y'),\quad (a + i.b).(x, y) = (ax - by, bx + ay)\,.$$

On obtient un espace vectoriel complexe E^c de dimension n qu'on appelle le complexifié de E. Le sous-espace $\{ (x, 0) : x \in E \}$ est canoniquement isomorphe à E. On l'identifie à E et on l'appelle l'espace des vecteurs réels. Cela permet d'écrire x au lieu de $(x, 0)$ et comme $(x, y) = (x, 0) + i.(y, 0)$, on écrira $x + iy$ au lieu de (x, y).

Définissons maintenant le complexifié de $A \in \mathrm{End}\,(E)$. C'est l'endomorphisme A^c de E^c défini par $A^c(x + iy) = A(x) + i.A(y)$. Observons qu'il laisse invariant le sous-espace réel E.

Enfin, nous appellerons complexifiée de l'équation différentielle $x' = A.x$ l'équation $z' = A^c.z$. On vérifie sans peine les propriétés suivantes de ses solutions :

a) Une solution $t \rightsquigarrow z(t) \in E^c$ de l'équation complexifiée telle que $z(0) \in E$ est réelle ; c'est-à-dire $z(t) \in E$ pour tout t. De plus cette solution est également solution de l'équation réelle.

b) La fonction $t \rightsquigarrow z(t) = x(t) + i.y(t)$ est solution de l'équation complexifiée si, et seulement si, sa partie réelle $x(t)$ et sa partie imaginaire pure $y(t)$ vérifient l'équation réelle.

Venons-en à la résolution de l'équation $x' = A.x$ dans E avec la condition initiale $x(0) = x_0 \in E$. Résolvons l'équation complexifiée $z' = A^c.z$, avec la même condition initiale $z(0) = x_0 \in E$, selon la méthode du 5.3. D'après la propriété *a* ci-dessus la solution sera donc réelle et vérifiera $x' = A.x$. C'est donc la solution cherchée. Il nous reste à l'exprimer sous forme réelle, car elle est la somme de vecteurs de la forme $e^{\lambda.t}.X(t)$, où $X(t)$ est un polynôme en t, et où la valeur propre λ de A n'est pas forcément réelle. D'après la propriété *b* ci-dessus les parties réelle et imaginaire de $e^{\lambda.t}.X(t)$ sont solutions de $x' = A.x$. Si $\lambda = a+ib$, $a, b \in \mathbf{R}$, on obtiendra donc la solution de $x' = A.x$ sous forme réelle en écrivant qu'elle est la somme de termes de la forme $e^{at}.\cos(bt).X(t)$ et $e^{at}.\sin(bt).X(t)$, où $X(t)$ est un polynôme en t à coefficients réels, de degré inférieur à l'ordre de multiplicité de λ.

Donnons quelques applications des résultats précédents.

6. Equations différentielles linéaires homogènes d'ordre n à coefficients constants

Commençons par un exemple emprunté à la mécanique.

Oscillateur harmonique. 6.1. — Un point de masse $m = 1$ se meut sur une droite $\mathbf{R}$, rappelé par l'origine O avec une force proportionnelle à sa distance à O [à l'aide d'un ressort, par exemple]. Si t désigne le temps, l'abscisse du point est une fonction $t \rightsquigarrow q(t)$ qui obéit à la loi de Newton : $q'' = \omega^2.q$, où ω est une constante positive [qui dépend du ressort].

L'équation $q'' + \omega^2.q = 0$ contient la dérivée seconde de q. Nous allons la ramener à une équation du premier ordre par un artifice : *la réduction au premier ordre*. Posons $q' = mq' = p$ (c'est l'impulsion du point). L'équation précédente est équivalente au système différentiel linéaire du premier ordre :

$$q' = p\,, \qquad p' = -\,\omega^2.q\,, \tag{6.2}$$

c'est-à-dire qu'à toute solution q de l'équation correspond une solution $(q, p = q')$ de (6.2) et, qu'inversement, étant donnée une solution (q, p) de (6.2), q est une solution de l'équation $q'' + \omega^2 q = 0$.

La solution générale de (6.2) s'obtient aisément :

$$\begin{cases} q(t) = a.\cos\omega t + b.\sin\omega t\,, \\ p(t) = \omega.[-\,a.\sin\omega t + b.\cos\omega t]\,, \end{cases}$$

où la position initiale $q(0) = a$ et l'impulsion initiale $p(0) = \omega.b$ sont arbitraires.

L'espace $\mathbf{R}^2$ des (q, p) s'appelle l'espace des phases. Son intérêt réside dans le fait que par un point quelconque $(q(0), p(0))$ passe une courbe solution (on dit aussi une courbe intégrale) et une seule. Ces courbes intégrales sont de deux types : l'origine $(0, 0)$, et des ellipses de centre O et d'équation $q^2 + \omega^{-2}.p^2 = a^2 + b^2$.

Notons enfin que toutes les solutions sont de période $2\,\pi/\omega$.

Remarque. — Donnons-nous n oscillateurs harmoniques. L'espace des phases est

$$\mathbf{R}^{2n} = \{ q = (q_1, \ldots, q_n), p = (p_1, \ldots, p_n) \} ,$$

où q_k (resp. p_k) est la position (resp. l'impulsion) du k-ième point matériel. L'évolution du système mécanique est régie par le système différentiel linéaire homogène à coefficients constants $q'_k = p_k$, $p'_k = -(\omega_k)^2 . q_k$, $k = 1, \ldots, n$.

Les courbes intégrales $t \rightsquigarrow (q(t), p(t))$ sont encore bornées, mais elles ne sont périodiques que si, et seulement si, les périodes partielles $2\pi/\omega_k$ sont commensurables, i.e., s'il existe des entiers non tous nuls $c_1, \ldots, c_n$ tels que $c_1.\omega_1 + \cdots + c_n.\omega_n = 0$.

Cas général. 6.3. — On dit qu'une fonction n fois dérivable $y : \mathbf{R} \to \mathbf{C}$ est solution d'une équation différentielle linéaire homogène d'ordre n à coefficients $a_1, \ldots, a_n$ constants si elle vérifie

$$y^n + a_1 . y^{(n-1)} + \cdots + a_1 . y = 0 , \tag{6.4}$$

où $y^{(k)}$ est la dérivée k-ième de y.

Résoudre cette équation c'est trouver toutes les fonctions y qui la vérifient. Pour cela, utilisons la méthode de réduction au premier ordre. Posons $y = x_1$, $y' = x_2, \ldots, y^{(n-1)} = x_n$; évidemment $x = (x_1, \ldots, x_m)$ vérifie le système différentiel linéaire homogène du premier ordre à coefficients constants

$$\begin{cases} x'_1 = x_2 \\ \vdots \\ x'_{n-1} = x_n \\ x'_n = -a_n . x_1 - a_{n-1} . x_2 - \cdots - a_1 . x_n . \end{cases} \tag{6.5}$$

Réciproquement, il est clair que la première composante x_1 d'une solution de (6.5) est solution de (6.4). En posant

$$A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & & & & 1 \\ -a_n & -a_{n-1} & & \ldots & -a_1 \end{bmatrix}$$

(6.5) s'écrit $x' = AX$. L'équation caractéristique $\det(A - k.1) = 0$ n'est autre que $k^n + a_1 k^{n-1} + \cdots + a_n = 0$, comme on le voit aisément par récurrence sur n. D'après les résultats de 5.3., si k_i est une racine d'ordre r_i de ce polynôme, la solution générale de $x' = Ax$ est $\sum_1^m \exp(k_i . t) . P_i(t)$, où $P_i(t)$ est un polynôme de degré $r_i - 1$. En n'en gardant que la première composante, on voit que la solution générale de (6.4) est de la même forme, $P_i(t)$ étant un polynôme à valeurs scalaires.

On peut trouver choquant de ramener l'intégration de (6.4) à celle d'un système du premier ordre. Voici une autre méthode, dont le domaine d'application déborde largement le cadre présent.

Polynômes différentiels à coefficients constants. 6.6.

Revenons à l'équation (6.4). Une solution y est a priori n fois différentiable. En fait, puisque $y^{(n)} = -a_1 . y^{(n-1)} - \cdots - a_n . y$, $y^{(n)}$ est différentiable, donc y est $n + 1$ fois différentiable. Finalement, de proche en proche, on voit que y est C^∞ : toute solution de (6.4) appartient à $C^\infty = C^\infty(\mathbf{R} ; \mathbf{C})$.

Désignons par D l'application linéaire (non continue !) qui fait correspondre à $f \in C^\infty$ sa dérivée f et, plus généralement par $D^k f$ la dérivée k-ième de f : on peut considérer que D^k

est la puissance k-ième de l'opérateur D. L'équation (6.4) s'écrit, avec ces notations,

$$D^n f + a_1.D^{n-1} f + \cdots + a_n y = 0\,.$$

Sa résolution est ramenée au problème suivant : trouver le noyau de l'opérateur $p(D) = D^n + a_1\, D^{n+1} + \cdots + a_n$ de l'espace C^∞.

On dit que $p(D)$ est un polynôme différentiel à coefficients constants. Il est clair que l'ensemble de ces polynômes forme un anneau isomorphe à celui des polynômes en $x \in \mathbf{C}$. En particulier, si $p(x)$ se factorise en $p(x) = q(x).d(x)$, on a aussi $p(D) = q(D).d(D)$. En sorte que $y \in \operatorname{Ker} p(D)$ équivaut à $d(D)\, y \in \operatorname{Ker} q(D)$.

Plaçons-nous sur le corps des complexes et revenons à l'équation (6.4), écrite sous la forme $p(D)\, y = 0$. Le polynôme p se factorise complètement et, si k est l'une de ses racines, $p(D) = (D - k).q(D)$. On a donc $q(D)\, y \in \operatorname{Ker}(D - k)$. Or un élément u de $\operatorname{Ker}(D - k)$ vérifiant $u' = k.u$, on a $u(t) = A.\exp(k.t)$, où A est une constante arbitraire. Ainsi (6.4) s'écrit encore $q(D).y = A.\exp(kt)$. On peut continuer en factorisant q. Chaque étape introduit une constante arbitraire. On a terminé au terme de degré (p) étapes. On voit donc, a priori, que l'espace vectoriel des solutions de $p(D)\, f = 0$ est de dimension $n =$ degré (p).

Cherchons une base de cet espace. Nous allons montrer que si k est une racine multiple d'ordre r de p, alors $t \rightsquigarrow \exp(kt).t^s$ est dans $\operatorname{Ker}(p)$ pour $0 \leqslant s < r$. Puisque $(D - k)^r$ divise $p(D)$, il suffit de prouver que $(D - k)^r.(\exp(kt).t^s) = 0$. Or $(D - k)\,(\exp(kt).t^s) = s.\exp(kt).t^{s-1}$, d'où le résultat par itération. Il en résulte que si $k_1, \ldots, k_m$ sont les racines distinctes de p, d'ordres de multiplicité respectifs $r_1, \ldots, r_m$, alors

$$\exp(k_1.t), \ldots, \exp(k_1.t).t^{r_1-1}, \ldots, \exp(k_m.t).t^{r_m-1}$$

sont $r_1 + \cdots + r_m = n$ solutions.

Il reste à montrer qu'elles sont linéairement indépendantes, c'est-à-dire que si $p_1, \ldots, p_m$ sont des polynômes de degrés respectifs $r_1 - 1, \ldots, r_m - 1$, alors

$$\exp(k_1.t).p_1(t) + \cdots + \exp(k_m.t).p_m(t) = 0$$

entraîne $p_1 = \cdots = p_m = 0$. On le prouve par récurrence sur m, en le supposant vrai pour $m - 1$: on divise les deux membres par $\exp(k_1.t)$, puis on dérive (degré p_1) + 1 fois. Il vient $\exp(k_2 - k_1)\, t.P_2(t) + \cdots = 0$, où P_i, s'il n'est pas nul, est un polynôme de même degré que p_i. Puisque $k_2 - k_1 \neq 0$, etc., l'hypothèse de récurrence entraîne $p_2 = \cdots = p_m = 0$; donc $p_2 = \cdots = p_m = 0$ et par suite $p_1 = 0$.

7. Solutions bornées ou périodiques de $x' = A.x$

E est un espace vectoriel complexe de dimension n et $A \in \operatorname{End}(E)$.

Cherchons sous quelles conditions les solutions de $x' = Ax$ sont toutes bornées.

D'après (5.3) la solution générale est une combinaison linéaire des vecteurs de la forme $\exp(kt).P(t)$, où P est un polynôme de degré inférieur à l'ordre de multiplicité de la valeur propre k du polynôme caractéristique de A. Si toutes les solutions sont bornées, du fait que $\lim e^{at}.t^k = +\infty$ lorsque $t \to +\infty$ (resp. $-\infty$) si $a > 0$ (resp. $a < 0$), on voit que les valeurs propres sont toutes imaginaires pures : $k = i.\omega$, où $\omega \in \mathbf{R}$. Revenons au vecteur solution $\exp(kt).P(t)$; dans ces conditions $\|\exp(kt).P(t)\| = \|P(t)\|$, qui ne peut rester borné que si P est de degré nul. L'opérateur A est donc diagonalisable et sa matrice dans une base convenable est $\operatorname{diag}(i\omega_1, \ldots, i\omega_n)$ où $\omega_k \in \mathbf{R}$.

Réciproquement, s'il en est ainsi, la solution générale de $x' = Ax$, écrite dans la base ci-dessus, est $(\exp(i\omega_1\, t)\, x_1, \ldots)$. Elle demeure donc bornée. On a donc prouvé le résultat suivant :

Théorème. — ***Pour que toutes les solutions de $x' = Ax$ soient bornées, il faut, et il suffit que A soit diagonalisable et que ses valeurs propres soient imaginaires pures.***

Remarque. — Via la base propre ci-dessus, E s'identifie à $\mathbf{C}^n$. Ecrivons chacune des n copies de $\mathbf{C}$ sous la forme $\{ q_r + ip_r : q_r, p_r \in \mathbf{R} \}$. L'espace E s'identifie encore a $\mathbf{R}^{2n} = \{(q_1, p_1, \dots, q_n, p_n)\}$. et la solution générale de $x' = A.x$ s'écrit

$$q_r = a_r.\cos(\omega_r.t) - b_r.\sin(\omega_r.t)$$
$$p_r = a_r.\sin(\omega_r.t) + b_r.\cos(\omega_r.t)$$

Aux notations près, on reconnaît (voir 6.) la solution générale d'un système de n oscillateurs harmoniques de fréquences ω_r. On en déduit :

Théorème. — ***Pour que toutes les solutions de $x' = A.x$ soient périodiques (donc bornées), il faut et il suffit que A soit diagonalisable et que ses valeurs propres soient imaginaires pures et commensurables.***

Bibliographie. — Pour des exposés plus complets sur ces questions on renvoie le lecteur au traité de E. A. Coddington et N. Levinson, et à l'ouvrage de V. Arnold, nourri d'exemples et écrit dans un style vivifiant.

6 PRODUIT INTÉGRAL ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES

Ce chapitre est consacré aux équations différentielles linéaires à coefficients variables.

Préliminaires

Soit A un endomorphisme continu d'un espace de Banach E. Nous avons vu au chapitre 5 que la solution générale de l'équation différentielle linéaire $x' = A.x$ est $\exp(t.A).x_0$, $x_0 \in E$. D'autre part, on sait (1.3. chap. 5) que $\exp(t.A) = \lim\limits_{n=\infty} \left(1 + \dfrac{tA}{n}\right)^n$; $\left(1 + \dfrac{tA}{n}\right)^n.x_0$ est donc une solution approchée de $x' = A.x$; c'est la méthode d'Euler. Nous allons la généraliser suivant un procédé dû à V. Volterra (1887).

Cessons de supposer que A est une constante. Si A est une application continue d'un intervalle I dans End (E), proposons-nous de trouver une application dérivable $x : I \to E$ vérifiant $x'(t) = A(t).x(t)$ pour $t \in I$, et $x(t_0) = x_0$, où $t_0 \in I$, $x_0 \in E$. Cette équation peut être approchée par un système d'équations aux différences finies :

$$\frac{x(t_{i+1}) - x(t_i)}{t_{i+1} - t_i} = A(t_i).x(t_i)\,, \qquad i = 0, \dots, n-1\,,$$

où $t_0 < t_1 < \cdots < t_n = t$. A partir de ces équations on obtient

$$x(t) = \prod_{i=0}^{n} (1 + A(t_i)).(t_{i+1} - t_i).x_0\,.$$

Nous allons montrer que cette expression converge vers la solution cherchée de $x'(t) = A(t).x(t)$ lorsque le pas $t_{i+1} - t_i$ de la partition tend vers zéro. Si, en particulier, $t_0 = 0$, $t_{i+1} - t_i = (t-a)/n$ et A = constante, la limite n'est autre que $\exp(tA).x_0$; c'est la solution de $x' = A.x$ vérifiant $x(0) = x_0$.

1. Produit intégral

Produit intégral de fonctions en escalier. 1.1. — Donnons-nous un intervalle fermé $[a, b]$, $a < b$. Découpons-le en n intervalles par des points $a = t_0 < t_1 < \cdots < t_n = b$; donnons-nous n endomorphismes $A_1, \dots, A_n \in$ End (E). L'application $A : [a, b] \to$ End (E), définie par $A(t) = A_1$ pour $t_0 \leqslant t < t_1, \dots, A(t) = A_n$ pour $t_{n-1} \leqslant t \leqslant t_n$ est une fonction en escalier (voir 6. chap. 2). Son produit intégral entre a et b est, par définition,

$$P_a^b(A) = \prod_a^b (1 + A(t).\mathrm{d}t) = \exp(\Delta t_n.A_n) \dots \exp(\Delta t_1.A_1)$$

où $\Delta t_j = t_j - t_{j-1}$ pour $1 \leqslant j \leqslant n$.

Le lecteur prendra garde à l'ordre des facteurs, puisqu'a priori, $A_1, \dots, A_n$ ne commutent pas. Par contre, s'ils commutent, et en particulier si A est constante, on a

$$P_a^b(A) = \exp(\Delta t_n.A_n + \cdots + \Delta t_1.A_1) = \exp\left(\int_a^b A(t).\mathrm{d}t\right),$$

d'aprés (2.3. chap. 5).

Remarque. — On ne change pas le produit intégral en ajoutant des points de subdivision à ceux de la subdivision t_i. En effet $A(t)$ demeure constante entre t_i et t_{i+1} ; si $t_i < t' < t_{i+1}$, on a donc, d'après (2.3. chap. 5),

$$\exp(t_{i+1} - t_i)\,A = \exp[(t_{i+1} - t').A + (t' - t_i).A] = \exp(t_{i+1} - t')\,A.\exp(t' - t_i)\,A\,.$$

Lemme. 1.2. — ***Soient A et B deux fonctions en escalier définies sur [a, b] et à valeurs dans E. Alors***

$$\| P_a^b(A) - P_a^b(B) \| \leqslant (b - a).e^{(b-a).M}.\| A - B \|\,,$$

où $\| A \| = \sup \| A(t) \|$ ***pour*** $a \leqslant t \leqslant b$ ***et*** $M = \max(\| A \|, \| B \|)$.

Preuve. — D'après la remarque précédente on peut supposer que A et B sont constantes sur chacun des intervalles d'une même subdivision $a = t_0 < \cdots < t_n = b$.

Posons $a_i = \exp(\Delta t_i.A_i)$, $b_i = \exp(t.B_i)$. Alors

$$\| P(A) - P(B) \| = \| a_n \ldots a_1 - b_n \ldots b_1 \| \leqslant \| a_n \ldots a_1 - a_n \ldots a_2.b_1 \| +$$
$$+ \| a_n \ldots a_2.b_1 - a_n \ldots a_3\, b_2\, b_1 \| + \cdots + \| a_n.b_{n-1} \ldots b_1 - b_n\, b_{n-1} \ldots b_1 \| \leqslant$$
$$\| a_1 - b_1 \|.\| a_n \| \ldots \| a_2 \| + \| a_2 - b_2 \|.\| a_n \| \ldots \| a_3 \|.\| b_1 \| + \cdots +$$
$$+ \| a_n - b_n \|.\| b_{n-1} \| \ldots \| b_1 \|\,.$$

Utilisons l'estimée 2.1. du chapitre 5 et $\| \exp L \| \leqslant \exp \| L \|$ pour majorer cette dernière expression. On trouve

$$\| P(A)-P(B) \| \leqslant \Delta t_1.\| A_1 - B_1 \| \exp(\Delta t_1.M).\exp(\Delta t_n.M) \ldots \exp(\Delta t_2.M) + \cdots +$$
$$+ \Delta t_n.\| A_n - B_n \|.\exp(\Delta t_n.M).\exp(\Delta t_{n-1}.M) \ldots \exp(\Delta t_1.M) \leqslant (b-a).\| A - B \|.\exp(b-a)M.$$

□

Produit intégral de fonctions réglées. 1.3. — Supposons que $A : [a, b] \to \mathrm{End}(E)$ soit une fonction réglée, c'est-à-dire (6.2. chap. 2) qu'elle soit limite uniforme sur $[a, b]$ d'une suite A_k de fonctions en escalier. D'après le lemme précédent

$$\| P_a^b(A_p) - P_a^b(A_q) \| \leqslant (b - a).e^{(b-a).M} \| A_p - A_q \|\,,$$

où M est un majorant des $\| A_k \|$.

Il en résulte que $P(A_k)$ est une suite de Cauchy de End (E). Puisque E est un espace de Banach, End (E) est complet et $P(A_k)$ converge. On voit immédiatement que sa limite ne dépend pas de la suite A_k utilisée pour approcher A. Il est donc légitime de définir ce qu'on appelle le produit intégral de A entre a et b par

$$P_a^b(A) = \prod_a^b (1 + A(t).\mathrm{d}t) = \lim_{k=\infty} P_a^b(A_k)\,.$$

Propriétés du produit intégral. 1.4.

a) Si $\| A \| = \sup \| A(t) \|$ pour $a \leqslant t \leqslant b$, la relation $\| \exp(L) \| \leqslant \exp \| L \|$ entraîne immédiatement $\| P_a^b(A) \| \leqslant \exp(b - a) \| A \|$.

b) Le lemme 1.2. s'étend par continuité à toutes les fonctions réglées A et B.

c) $P_a^b(A)$ est inversible.

La relation $[\exp(L)]^{-1} = \exp(-L)$ le montre immédiatement pour une fonction en escalier :

$$[P_a^b(A)]^{-1} = [\exp(\Delta t_n.A_n) \ldots \exp(\Delta t_1.A_1)]^{-1} = \exp(-\Delta t_1.A_1) \ldots \exp(-\Delta t_n.A_n)\,.$$

[attention à l'ordre des facteurs].

La propriété s'étend à une fonction réglée quelconque par continuité.

d) Relation de Chasles.

Evidemment $P_a^a(A) = 1 (= \mathrm{id}_E)$.

Si $a > b$, la propriété précédente permet de définir $P_a^b(A)$ par $[P_b^a(A)]^{-1}$.

Avec cette convention, si $A : I \to \mathrm{End}\,(E)$ est une fonction réglée définie sur un intervalle I, on a l'analogue de la relation de Chasles :

$$P_a^c(A) = P_b^c(A).P_a^b(A) \quad \text{pour} \quad a, b, c \in I$$

[attention à l'ordre des facteurs]. On le vérifie sans peine pour une fonction en escalier, et le cas général s'obtient par continuité.

e) Le lecteur, curieux de s'expliquer la notation $\prod_a^b (1 + \mathrm{A}(t).\mathrm{d}t)$, pourra démontrer en s'inspirant de (1.3. chap. 5) que

$$P_a^b(A) = \lim \prod_{i=0}^{n} (1 + \Delta t_i . A(t_i))$$

lorsque le plus grand des pas $\Delta t_i = t_{i+1} - t_i$ de la subdivision $a = t_0 < \cdots < t_n = b$ tend vers zéro.

Enfin, voici l'analogue du théorème (6.5. chap. 2) qui dit qu'une fonction continue est la dérivée de l'une quelconque de ses primitives :

f) Si $a < t < b$ et si $A : [a, b] \to \mathrm{End}\,(E)$ est continue, alors $f(t) = P_a^t(A)$ est dérivable et $f'(t) = A(t).f(t)$.

Preuve. — Prenons $|h|$ assez petit pour que $a < t + h < b$. Le produit intégral entre t et $t + h$ de l'endomorphisme constant $A(t)$ est évidemment

$$P_t^{t+h}(A(t)) = \exp(h.A(t)) = 1 + h.A(t) + \mathrm{o}(h).$$

D'autre part le lemme 1.2. entraîne

$$\| P_t^{t+h}(A) - P_t^{t+h}(A(t)) \| \leqslant |h|.\exp(|h|.\|A\|) \sup_{t \leqslant s \leqslant t+h} \| A(s) - A(t) \|,$$

qui est un $\mathrm{o}(h)$ puisque A est continue. Il en résulte $P_t^{t+h}(A) = 1 + h.A(t) + \mathrm{o}(h)$. Et le théorème résulte de la relation de Chasles

$$P_a^{t+h}(A) - P_a^t(A) = (P_t^{t+h}(A) - 1).P_a^t(A). \qquad \square$$

2. Equations différentielles linéaires homogènes

Définition. 2.1. — Soient I un intervalle de $\mathbf{R}$ (éventuellement $\mathbf{R}$ tout entier) et $A : I \to \mathrm{End}\,(E)$. On dit qu'une application différentiable $f : I \to E$ est solution de l'équation différentielle linéaire homogène du premier ordre $x'(t) = A(t).x(t)$ si $f'(t) = A(t).f(t)$ pour $t \in I$.

Remarques. — *a*) Si A est constante, nous retrouvons les équations différentielles à coefficients constants du chapitre précédent.

b) Exactement comme au chapitre 5, on voit que l'ensemble S_I des solutions $f : I \to E$ est un espace vectoriel. On aura donc résolu l'équation, c'est-à-dire trouvé toutes ses solutions, si l'on exhibe une base de S_I.

c) Si A est de classe C^k, $k \geqslant 0$, alors toute solution est de classe C^{k+1}, comme on le voit immédiatement par récurrence sur k.

Théorème fondamental. 2.2. — ***Soient E un espace de Banach, $x_0 \in E$, I un intervalle de $\mathbf{R}$, $t_0 \in I$, et $A : I \to$ End (E) une fonction continue. Alors il existe une application différentiable $f : I \to E$, et une seule, telle que***

$$\begin{cases} f'(t) = A(t).f(t) & \textit{pour} \quad t \in I\,; \\ f(t_0) = x_0\,. \end{cases}$$

Cette application est de classe C^1 et elle est donnée par $f(t) = P_{t_0}^t(A).x_0$.

L'existence résulte immédiatement de la propriété f de 1.4.

Montrons que la solution donnée par le théorème est unique. Si g est un second candidat, formons $h(t) = \varphi(t)^{-1}.g(t)$, où $\varphi(t) = P_{t_0}^t(A)$. C'est une fonction dérivable et, d'après (2.8. chap. 4),

$$h'(t) = -\varphi(t)^{-1}.\varphi'(t).\varphi(t)^{-1}.g(t) + \varphi(t)^{-1}.g'(t).$$

Mais $\varphi'(t) = A(t).\varphi(t)$ et $g'(t) = A(t).g(t)$, donc $h'(t) = 0$. D'après (2. chap. 2) h est une constante et comme $\varphi(t_0) = 1$ on a $h(t_0) = x_0$. Ainsi $h(t) = x_0$ et l'on a bien

$$g(t) = \varphi(t).x_0 = f(t).$$

Corollaire. 2.3. — ***Exactement comme en (4.5. chap. 5), on voit, en prenant $I = \mathbf{R}$, que $x'(t) = A(t).x(t)$ possède une solution maximale $f : \mathbf{R} \to E$ vérifiant $f(t_0) = x_0$, et que toute autre solution $g : I \to E$ vérifiant $g(t_0) = x_0$ est la restriction de f à I.***

Corollaire. 2.4. — ***Exactement comme en (4.5. chap. 5), on voit que l'espace vectoriel S des solutions maximales $f : \mathbf{R} \to E$ de $x'(t) = A(t).x(t)$ est isomorphe à E.***

Si la dimension de E est finie, une base de S (on dit souvent : **un système fondamental de solutions*)*** ***comporte*** **dim** ***E vecteurs.***

La solution générale $P_{t_0}^t(A)$ de $x'(t) = A(t).x(t)$ dépend linéairement de la condition initiale x_0. Voyons comment elle dépend de A.

Théorème de comparaison. 2.5. — ***Soient E un espace de Banach, $x_0 \in E$, I un intervalle, $t_0 \in I$, $A : I \to \mathrm{End}\,(E)$ et $B : I \to \mathrm{End}\,(E)$ deux fonctions continues. Désignons par $f : I \to E$ la solution unique de $x'(t) = A(t).x(t)$ vérifiant $f(t_0) = x_0$, par $g : I \to E$ la solution unique de $x'(t) = B(t).x(t)$ vérifiant $g(t_0) = x_0$. Alors***

$$\| f(t) - g(t) \| \leqslant e^{|t-t_0|.M}.| t - t_0 |.\| A - B \|.\| x_0 \|,$$

où $\| A - B \|$ est la borne supérieure sur (t_0, t) de $\| A(s) - B(s) \|$ et M un majorant de $\| A \|$ et $\| B \|$.

PREUVE. — Le théorème fondamental donne les expressions de f et g. Il suffit ensuite d'appliquer le lemme 1.2. □

Remarque. — Proposons-nous d'approcher la solution f de $x'(t) = A(t).x(t)$ vérifiant $f(t_0) = x_0$. Approchons pour cela la fonction continue A par une fonction en escalier : étant donné $\varepsilon > 0$, l'uniforme continuité assure l'existence d'une subdivision $t_0 < t_1 < \cdots < t_n = t$ telle que $\| A - B \| < \varepsilon$ si $B : [t_0, t] \to \mathrm{End}\,(E)$ est la fonction en escalier définie par $B(t) = A(t_0)$ pour $t_0 \leqslant t \leqslant t_1, \ldots, B(t) = A(t_{n-1})$ pour $t_{n-1} \leqslant t \leqslant t_n$. En reprenant la preuve de 2.5., on voit que $\exp[\Delta t_n.A(t_{n-1})] \ldots \exp[\Delta t_1.A(t_0)].x_0$ approche f à $\varepsilon.e^{|t-t_0|.\|A\|}.| t-t_0 |.\| x_0 \|$ près.

Terminologie. 2.6. — On appelle souvent $P_a^t(A)$ la résolvante, ou le noyau résolvant de l'équation $x'(t) = A(t).x(t)$, et on la note $R(a, t)$.

Ce qui précède ne donne aucun moyen pratique pour la déterminer explicitement. Voici néanmoins, à titre indicatif, une forme théoriquement explicite.

Exponentielle de Dyson. 2.7. — Soit $A : I \to \mathrm{End}\,(E)$ une fonction continue. Donnons-nous un segment $[a, b]$, $a < b$, contenu dans l'intervalle I, et soit M un majorant de $\| A(t) \|$ pour $a \leqslant t \leqslant b$.

Si $c, t \in [a, b]$, définissons par récurrence des applications $R_n(c, t) : [a, b] \to \mathrm{End}\,(E)$:

$$(2.8) \qquad \begin{cases} R_0(c, t) = 1 (= \mathrm{id}_E) \\ \mathrm{R}_{n+1}(\mathrm{c}, \mathrm{t}) = \displaystyle\int_c^t A(s).R_n(c, s).\mathrm{d}s \quad \text{pour} \quad n \geqslant 0, \end{cases}$$

Montrons que $\| R_n(c, t) \| \leqslant \frac{|t - c|^n}{n\,!} \cdot M^n$. C'est évident pour $n = 0$. Supposons-le prouvé pour n ; (2.8) implique alors

$$\| R_{n+1}(c, t) \| \leqslant \left| \int_c^t M \cdot \frac{|s - c|^n}{n\,!} \cdot M^n \cdot \mathrm{d}s \right| \leqslant \frac{|t - c|^{n+1}}{(n + 1)\,!} \cdot M^{n+1}$$

et la propriété est établie.

La suite $R_0(c, t) + \cdots + R_n(c, t)$ converge donc uniformément sur $[a, b]$ vers une limite $S(c, t) = \sum_0^{\infty} R_n(c, t)$ et (2.8) entraîne

$$S(c, t) = 1 + \int_c^t A(s).S(c, s).\mathrm{d}s$$

Il en résulte que $f(t) = S(c, t)$ est dérivable et qu'elle vérifie l'équation différentielle $f'(t) = A(t).f(t)$, avec la condition initiale $f(c) = 1$. D'après (*f.* 1.4.) et l'unicité, on a donc $f(t) = P_c^t(A)$. En fin de compte $P_c^t(A) = \sum_0^{\infty} R_n(c, t)$, où les R_n sont déterminés de proche en proche grâce à (2.8.) :

$$R_n(c, t) = \int_{c \leqslant s_1 \leqslant \ldots \leqslant s_n \leqslant t} \cdots \int A(s_n) \cdots A(s_1).\mathrm{d}s_1 \cdots \mathrm{d}s_n\,, \quad \text{si} \quad c < t$$

(attention à l'ordre des facteurs).

Expression de la résolvante à l'aide d'un système fondamental de solutions. 2.9. — Choisissons une base dans l'espace E, supposé de dimension finie n. Soit $x_1(t), \ldots, x_n(t)$ un système fondamental de solutions de l'équation différentielle $x'(t) = A(t).x(t)$. Désignons par $X(t)$ la matrice $n \times n$ dont la k-ième colonne, lue de haut en bas, est formée par les composantes de $x_k(t)$. Puisque $x_k(t) = P_a^t(A).x_k(a)$, on en déduit $X(t) = R(a, t).X(a)$.

Puisque les vecteurs $x_1(a), \ldots, x_n(a)$ sont linéairement indépendants, la matrice $X(a)$ est inversible, donc $R(a, t) = X(t).X(a)^{-1}$.

Bien que ce soit la résolvante qui ait servi à démontrer l'existence des solutions de $x'(t) = A(t).x(t)$, c'est cette formule qui la déterminera explicitement si on a le bonheur d'exhiber n solutions linéairement indépendantes.

A défaut d'une expression explicite de la résolvante $R(a, t) = P_a^t(A)$, il se peut que des propriétés de $A(t)$ transparaissent dans $R(a, t)$. Voici quelques exemples importants :

Théorème. 2.10. — ***Si A est une constante, $R(a, t) = \exp(t - a).A$.***

Preuve. — Immédiate, d'après 1.1., et sans surprise. En effet, $\exp(t - a)\,A.x_0$ est la solution de l'équation différentielle à coefficients constants $x'(t) = A.x(t)$ vérifiant la condition initiale $x(a) = x_0$ (voir 4.4. chap. 5).

Théorème. 2.11. — ***Pour que $C \in \mathrm{End}\,(E)$ commute avec $R(a, t)$ pour tous a, t, il faut et il suffit que C commute avec $A(t)$ pour tout t.***

Preuve. — Supposons que $C.R(a, t) = R(a, t).C$ pour tous a, t. Puisque

$$\frac{\mathrm{d}}{\mathrm{d}t} R(a, t) = A(t).R(a, t)\,,$$

on en déduit $C.A(t).R(a, t) = A(t).R(a, t).C$; et comme $R(a, a) = 1$, on a bien $C.R(a) = R(a).C$ pour tout a.

Réciproquement, supposons que $C.A(t) = A(t).C$ pour tout t. Posons $f(t) = C.R(a, t) - R(a, t).C$. Puisque $\frac{\mathrm{d}}{\mathrm{d}t} R(a, t) = A(t).R(a, t)$, on a

$$f'(t) = C.A(t).R(a, t) - A(t).R(a, t).C = A(t).[C.R(a, t) - R(a, t).C] = A(t).f(t).$$

Donc f est la solution, à valeurs dans End (E), de $x'(t) = A(t).x(t)$ qui vérifie $f(a) = C.R(a, a) - C.R(a, a) = 0$. D'après l'unicité, $f(t) = 0$ pour tout t et l'on a bien $C.R(a, t) = R(a, t).C$.

Exemple. — Supposons que l'espace de Banach soit le complexifié $E^c = E + i.E$ de l'espace réel E et que, pour chaque t, $A(t)^c$ soit le complexifié de $A(t) \in \mathrm{End}(E)$ [voir 5.4. chap. 5]. Alors $A(t)^c$ commute pour tout t avec la projection canonique de E^c sur l'espace réel E. Il résulte de 2.11. que la solution de $z'(t) = A(t)^c.z(t)$ vérifiant $z(t_0) = z_0 \in E$ reste réelle pour tout t.

Théorème. 2.12. — ***Soit b une forme bilinéaire continue sur E. Pour que***

$$\textbf{(2.13)} \qquad b(R(a, t).x, R(a, t).y) = b(x, y) \quad \textit{pour } x, y \in E \textit{ et pour tous } a, t,$$

il faut et il suffit que

$$\textbf{(2.14)} \qquad b(A(t).x, y) + b(x, A(t).y) = 0 \quad \textit{pour } x, y \in E \textit{ et pour tout } t.$$

Preuve. — Appelons $f(t)$ le membre de gauche de (2.13). D'après la règle de Leibniz (2.9. chap. 1), $f'(t) = b(A(t).R(a, t).x, R(a, t).y) + b(R(a, t).x, A(t).R(a, t).y)$.

Si $f(t) = b(x, y)$, on a $f'(a) = 0$, donc (2.14). Réciproquement, si (2.14) est vérifié, $f'(t) = 0$ et $f(t)$ se réduit à la constante $f(a) = b(x, y)$. □

Exemples. — *a)* Si E est un espace de Hilbert réel (resp. complexe), $R(a, t)$ est orthogonal (resp. unitaire) pour tous a, t si, et seulement si, l'adjoint A^* de A vérifie $A^*(t) = -A(t)$ pour tout t.

S'il en est ainsi le produit scalaire (resp. hermitien) de deux solutions de $x'(t) = A(t).x(t)$ ne dépend pas de t.

b) Si E est un espace réel symplectique (c'est-à-dire muni d'une forme bilinéaire alternée non dégénérée ω), $A(a, t)$ est symplectique pour tous a, t si, et seulement si,

$$\omega(A(t).x, y) + \omega(x, A(t).y) = 0$$

pour $x, y \in E$ et pour tout t. On dit alors que $A(t)$ est une application infinitésimale symplectique.

Théorème de Jacobi-Liouville. 2.15. — ***Supposons E de dimension nfinie et soit $A : I \to \mathrm{End}(E)$ une fonction réglée. Alors***

$$\textbf{dét } P_a^t(A) = \exp \int_a^t \mathrm{tr}\,(A(s)).\mathrm{d}s \quad \textit{pour } a, t \in I \quad [\text{dét} = \textit{déterminant}, \ \mathrm{tr} = \textit{trace}].$$

Preuve. — Si la fonction A est en escalier, en reprenant les notations de 1.1. on a

$$P_a^t(A) = \exp(\Delta t_n . A_n) \dots \exp(\Delta t_1 . A_1).$$

La formule résulte alors de (2.4. chap. 5) et du fait que le déterminant d'un produit est le produit des déterminants des facteurs.

Le cas général s'en déduit. On prend une suite de fonctions en escalier convergeant uniformément vers A sur $[a, t]$ et l'on observe que le déterminant et la trace sont des fonctions continues.

Exemple. — Rappelons qu'un volume v de E est une forme n-linéaire antisymétrique non nulle et que l'image réciproque $f * v = \text{dét}\,(f).v$. Pour que f préserve un volume v, c'est-à-dire pour que $f * v = v$, il faut donc, et c'est suffisant, que dét $(f) = 1$.

D'après 2.15, on voit donc que pour que $P_a^t(A)$ préserve un volume de E pour tous a, t, il faut et il suffit que tr $A(t) = 0$ pour tout t.

3. Equations différentielles linéaires avec second membre

Définition. 3.1. — Soient E un espace de Banach, I un intervalle, $A : I \to \mathrm{End}(E)$ et $B : I \to E$ deux fonctions continues. On dit que l'application différentiable $f : I \to E$ vérifie l'équation différentielle linéaire du premier ordre

$$x'(t) = A(t).x(t) + B(t)\,, \tag{3.2}$$

si $f'(t) = A(t).f(t) + B(t)$ pour tout $t \in I$.

La fonction B s'appelle le « second membre » de l'équation.

Unicité. 3.3. — Etant donnés $x_0 \in E$ et $t_0 \in I$, il existe au plus une solution f de (3.2) vérifiant $f(t_0) = x_0$.

PREUVE. — La différence h de deux solutions vérifie l'équation différentielle linéaire homogène $h'(t) = A(t).h(t)$ et $h(t_0) = 0$. D'après l'unicité (théorème 2.2.), $h(t) = 0$ pour $t \in I$. □

Existence. 3.4. — Soit $R(t_0, t)$ la résolvante de l'équation homogène associée $x'(t) = A(t).x(t)$. Nous allons chercher la solution f de (3.2), vérifiant $f(t_0) = x_0$, sous la forme $f(t) = R(t_0, t).g(t)$ où $g : I \to E$ est une fonction différentiable inconnue. Si le second membre B de (3.2) était nul, g se réduirait à la constante x_0. C'est pourquoi cette méthode, due à Lagrange, s'appelle méthode de la variation de constante.

En tenant compte de $\dfrac{\mathrm{d}}{\mathrm{d}t} R(t_0, t) = A(t).R(t_0, t)$, on obtient

$$f'(t) = A(t).R(t_0, t).g(t) + R(t_0, t).g'(t) = A(t).f(t) + R(t_0, t).g'(t)\,.$$

Portons cette expression dans (3.2), on trouve $R(t_0, t).g'(t) = B(t)$. On sait que la résolvante est inversible et que $R(t_0, t)^{-1} = R(t, t_0)$ (voir c. 1.4.), donc $g'(t) = R(t, t_0).B(t)$. En intégrant et en utilisant $g(t_0) = R(t_0, t_0).g(t_0) = f(t_0) = x_0$, on obtient $g(t) = x_0 + \displaystyle\int_{t_0}^{t} R(s, t_0).B(s).\mathrm{d}s$.

Finalement, en utilisant la propriété d de (6.4. chap. 2) et la relation de Chasles (cf. 1.4.) on obtient

$$f(t) = R(t_0, t).x_0 + \int_{t_0}^{t} R(s, t).B(s).\mathrm{d}s\,. \tag{3.5}$$

Notons que le second membre est la somme de la solution générale $R(t_0, t).x_0$ de l'équation homogène et de la solution $\displaystyle\int_{t_0}^{t} R(s, t).B(s).\mathrm{d}s$ de l'équation (3.2) qui s'annule pour $t = t_0$. Cette dernière dépend linéairement du second membre B.

Comme l'intégration au second membre de (3.5) n'est pas toujours aisée, on recourt souvent, dans la pratique, à des méthodes qui sortent du cadre de ce livre (transformation de Laplace).

Si la dimension de E est finie, on peut procéder ainsi pour intégrer l'équation (3.2) : une fois déterminé un système fondamental de solutions $x_1, \ldots, x_n$ de l'équation homogène, on cherche la solution de (3.2) sous la forme $f(t) = \sum_r f_r(t).x_r(t)$, où les f_r sont des fonctions différentiables à valeurs numériques. En écrivant que f vérifie (3.2) on obtient

$$\sum f_r'(t).x_r(t) = B(t)\,.$$

Cette équation détermine les f_r', car les $x_r(t)$ forment une base de E. On en déduit les f_r.

Application : Comparaison de solutions. 3.6. — Montrons comment la formule (3.5) permet d'améliorer, il est vrai sous des hypothèses plus fortes, le théorème de comparaison 2.5.

Soient I un intervalle, F un espace de Banach, $A : I \times F \to \mathrm{End}(E)$ une application continue en $t \in I$ et différentiable en $k \in F$.

Considérons l'équation différentielle linéaire homogène, dépendant d'un paramètre $k \in F$:

$$x'(t) = A(t, k).x(t)\,. \tag{3.7}$$

Soit $f(t, k)$ sa solution telle que $f(t_0, k) = x_0$, où $t_0 \in I$ et $x_0 \in E$ sont donnés. Cherchons comment elle dépend de k et, pour cela, calculons sa différentielle en k pour $k = k_0$. Puisque $D_1\, f(t, k) = A(t, k).f(t, k)$ on a, en utilisant le théorème de Schwarz (1.5. chap. 4),

$$D_1\, D_2\, f(t, k_0) = D_2\, D_1\, f(t, k_0) = A(t, k_0).D_2\, f(t, k_0) + D_2\, A(t, k_0).f(t, k_0)\,.$$

C'est une équation de la forme $x'(t) = A(t, k_0).x(t) + B(t)$, où $x(t) = D_2\, f(t, k_0)$, $B(t) = D_2\, A(t, k_0).f(t, k_0)$, avec la condition initiale $x(t_0) = D_2\, f(t_0, k_0) = 0$, car $f(t_0, k) = x_0$ pour tout k. D'après (3.5) on a donc

$$D_2\, f(t, k_0) = \int_{t_0}^{t} R(s, t, k_0).D_2\, A(s, k_0).f(s, k_0).\mathrm{d}s\,, \tag{3.8}$$

où $R(.,\, .,\, k_0)$ est la résolvante de (3.7).

En remplaçant $f(t, k_0)$ par sa valeur $R(t_0, t, k_0).x_0$ et en jouant sur l'arbitraire de x_0, on obtient une formule semblable pour $D_2\, R(t_0, t, k_0)$.

Exemple. — Le lecteur appliquera la formule (3.8) au calcul approché de la solution de l'équation de l'oscillateur harmonique perturbé par un terme de frottement $k.\varphi(t)$, où k est un petit paramètre réel et φ une fonction continue :

$$q' = p\,, \qquad p' = -\,\omega^2.q - k.\varphi.p\,.$$

4. Equations différentielles linéaires d'ordre n

Définition 4.1. — Donnons-nous $n + 1$ fonctions continues $a_1, \dots, a_2$, b, définies sur un intervalle I et à valeurs dans $\mathbf{C}$. On dit qu'une fonction n fois différentiable $f : I \to \mathbf{C}$ est solution de l'équation différentielle d'ordre n

$$y^{(n)} + a_1.y^{(n-1)} + \cdots + a_n.y = b\,, \tag{4.2}$$

si f et ses dérivées $f^{(k)}$ d'ordres $k = 1, \dots, n$ vérifient

$$f^{(n)}(t) + a_1(t).f^{(n-1)}(t) + \cdots + a_n(t).f(t) = b(t) \quad \text{pour } t \in I\,.$$

La fonction b est appelée le « second membre » de (4.2), et l'équation obtenue à partir de (4.2) en annulant b s'appelle l'équation homogène associée.

Nous allons ramener cette équation à une équation vectorielle du premier ordre par la méthode de réduction au premier ordre (6. chap. 5). Posons $y = x_1$, $y' = x_2, \dots, y^{(n-1)} = x_n$, on obtient

$$\left\{\begin{aligned} x_1' &= x_2\,; \\ &\,\vdots \\ x_{n-1}' &= x_n\,; \\ x_n' &= -\,a_n.x_1 - \cdots - a_1.x_n + b\,. \end{aligned}\right. \tag{4.3}$$

Si f est une solution de (4.2), alors $(f, f', \dots, f^{(n-1)})$ est une solution de (4.3). Réciproquement, la première composante f d'une solution de (4.3) est une solution de (4.2).

Introduisons les vecteurs $x(t)$ et $B(t)$ de composantes respectives $(x_1(t), \dots, x_n(t))$ et $(0, \dots, 0, b(t))$, et la matrice

$$(4.4) \qquad A(t) = \begin{bmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ & & & 1 \\ -a_n(t) & -a_{n-1}(t) & \dots & -a_1(t) \end{bmatrix},$$

(4.3) s'écrit encore $x'(t) = A(t).x(t) + B(t)$. En traduisant les résultats de (3.3) et (3.4) on en déduit :

Théorème. 4.5. — ***Etant donnés $t_0 \in I$ et $x_0, \dots, x_0^{(n-1)} \in \mathbf{C}$, l'équation (4.2) possède une solution $f : I \to \mathbf{C}$ et une seule vérifiant $f(t_0) = x_0, \dots, f^{(n-1)}(t_0) = x_0^{(n-1)}$.***

Théorème. 4.6. — ***L'intégrale générale de (4.2) s'obtient en ajoutant une solution particulière de cette équation à l'intégrale générale de l'équation homogène associée.***

Ce théorème ramène la résolution de l'équation (4.2) à celle de l'équation homogène associée et à la recherche d'une solution particulière.

Recherche d'une solution particulière. 4.7. — On appliquera la méthode de variation de constante de Lagrange.

Dans certains cas on pourra utiliser la remarque suivante : si le second membre b de (4.2) est la somme $\sum b_r$ de plusieurs fonctions, et si s_r est une solution particulière de l'équation obtenue à partir de (4.2) en remplaçant b par b_r, alors $\sum s_r$ est une solution particulière de (4.2).

Voici un cas encore plus particulier. Soit à résoudre l'équation $y^{(n)} = b$. Cela semble exiger n intégrations successives. Mais, d'après (4.5), on peut observer qu'il s'agit de déterminer une fonction f dont la dérivée n-ième est connue et pour laquelle $f(t_0), \dots, f^{(n-1)}(t_0)$ sont arbitraires. La formule de Taylor avec reste intégral (3.4. chap. 4) résout le problème :

$$f(t) = f(t_0) + (t - t_0) \cdot f'(t_0) + \cdots + \frac{(t - t_0)^{n-1}}{(n-1)!} \cdot f^{(n-1)}(t_0) + \\ + \frac{1}{(n-1)!} \int_{t_0}^{t} (t - s)^{n-1} \cdot b(s) \cdot \mathrm{d}s .$$

Résolution de l'équation homogène. 4.8. — Il n'existe pas, comme en (6.3. chap. 5), d'expression maniable de la solution générale. Même si $n = 2$.

Toutefois, supposons connue une solution particulière s non nulle de (4.2). Posons $y = s.u$ et calculons $y^{(k)}$ à l'aide de la formule de Leibniz. L'équation sera transformée en une équation du même type où le coefficient de u sera nul puisque s vérifie (4.2). On obtiendra donc une équation différentielle linéaire d'ordre $n - 1$ en u'. Si v est la solution générale de cette équation en u', la solution générale de (4.2) sera donnée par $s . \int v + C . s$, où C est une constante arbitraire.

Wronskien. 4.9. — Supposons connues n solutions $f_1, \dots, f_n$ de l'équation homogène associée à (4.2). On en déduit n solutions $x_r = (f_r, f_r', \dots, f_r^{(n-1)})$, $r = 1, \dots, n$, de l'équation $x'(t) = A(t).x(t)$. A partir de $(x_1, \dots, x_n)$ formons la matrice $X(t)$ introduite en (2.9) :

$$X(t) = \begin{bmatrix} f_1(t) & \dots & f_n(t) \\ \vdots & & \\ f_1^{(n-1)}(t) & \dots & f_n^{(n-1)}(t) \end{bmatrix}.$$

Le déterminant de $X(t)$ s'appelle le wronskien $w(t)$ des solutions $f_1, \dots, f_n$.

Si $R(t_0, t)$ est la matrice résolvante de $x'(t) = A(t).x(t)$, nous avons vu (2.9) que

$$R(t_0, t).X(t_0) = X(t).$$

Donc dét $R(t_0, t)$.dét $X(t_0)$ = dét $X(t)$. D'autre part nous savons (2.15) que

$$\text{dét } R(t_0, t) = \exp \int_{t_0}^{t} \operatorname{tr} A(s).\mathrm{d}s.$$

Comme (4.4) donne immédiatement tr $A(s) = -a_1(s)$, il en résulte :

$$w(t) = w(t_0).\exp\left(-\int_{t_0}^{t} a_1(s).\mathrm{d}s\right).$$

Cela montre (Liouville) que si $f_1, \ldots, f_n$ est une base de l'espace des solutions définies sur un intervalle I, leur wronskien ne s'annule pas sur I. A l'inverse, si $w(t_0) = 0$, alors $w(t) = 0$ pour tout $t \in I$.

Polynômes différentiels à coefficients variables. 4.10. — Cherchons à adapter la méthode de résolution proposée en (6.6. chap. 5) pour la résolution de l'équation homogène associée à (4.2).

On se heurte à une première difficulté concernant la construction d'un espace de solutions. Si les fonctions $a_1, \ldots, a_n$ sont de classe C^r, $r \geqslant 0$, au plus, alors f est de classe C^{n+r}, comme (4.2) le montre immédiatement par récurrence sur r. Par contre f n'est pas nécessairement de classe C^{n+r+1} et, a fortiori, n'appartient pas nécessairement à l'espace C^∞. L'opérateur $D = \dfrac{\mathrm{d}}{\mathrm{d}t}$ ne peut donc agir que sur un sous-espace de l'espace des solutions.

Supposons cette difficulté surmontée (par exemple, si les a_r sont de classe C^∞). Essayons d'écrire, dans le cas simple $n = 2$, l'équation $y'' + a_1.y' + a_2.y = 0$ sous la forme $(D - A_1)(D - A_2)\,y = 0$, où cette fois A_1 et A_2 ne sont plus des constantes. On obtient deux équations pour déterminer A_1 et A_2 :

$$A_1 + A_2 = -a_1 \quad \text{et} \quad A_1.A_2 - A_2' = a_2.$$

Il en résulte que A_2 doit être solution de l'équation $A_2' + a_1.A_2 + (A_2)^2 + a_2 = 0$. C'est une équation de Ricatti, c'est-à-dire une équation de la forme $z' + a. + b.z + c.z^2 = 0$; et Liouville a montré qu'en général elle ne s'intègre pas par quadratures.

La recherche des cas particuliers où elle s'intègre par quadratures relève de la théorie de Galois des corps différentiels et déborde le cadre de ce livre.

Retour aux équations homogènes à coefficients constants. 4.11. — Considérons l'équation linéaire et homogène à coefficients constants

$$y^{(n)} + a_1.y^{(n-1)} + \cdots + a_n.y = 0. \tag{4.12}$$

Supposons que l'équation caractéristique (voir 6.3. chap. 5) $p(k) = k^n + a_1.k^{n-1} + \cdots + a_n = 0$ admette une racine multiple c d'ordre r. On a donc $p(k) = (k - c)^r.q(k)$.

Substituons à p le polynôme $(k - c).(k - c - u)\ldots(k - c - (r - 1)\,u).q(k)$. C'est le polynôme caractéristique d'une équation de la forme

$$y^{(n)} + a_1(u).y^{(n-1)} + \cdots + a_n(u).y = 0. \tag{4.13}$$

D'après (6.3. chap. 5) cette équation admet pour solutions les fonctions $e^{ct}, \ldots, e^{(c+(r-1)u)t}$, donc aussi leurs combinaisons linéaires e^{ct}, $e^{c.t}\cdot\left(\dfrac{e^{ut}-1}{u}\right), \ldots, e^{ct}\cdot\left(\dfrac{e^{ut}-1}{u}\right)^{r-1}$, qui sont linéairement indépendantes. Lorsque u tend vers zéro ces fonctions ont pour limites

$$e^{ct}, e^{ct}.t, \ldots, e^{ct}.t^{r-1}. \tag{4.14}$$

Utilisons le théorème de comparaison 2.5. La matrice (4.4), formée avec $a_n(u)$, ..., $a_1(u)$, converge vers la matrice formée avec a_n, ..., a_1 si $u \to 0$. Il en résulte que les solutions de (4.13) convergent vers celles de (4.12). Donc (4.14) fournit r solutions linéairement indépendantes de (4.12). On a retrouvé, selon un procédé dû à d'Alembert, le résultat de (6.3. chap. 5).

Bibliographie. Il est impossible de donner une bibliographie, même sommaire, concernant les équations différentielles linéaires à coefficients variables d'ordre n (même si $n = 2$!). Renvoyons toutefois au traité déjà cité de Coddington et Levinson.

On trouvera aussi une étude des solutions périodiques à l'Appendice F.

7 CHAMPS DE VECTEURS ÉQUATIONS DIFFÉRENTIELLES

Ce chapitre est consacré aux équations différentielles, à l'existence de leurs solutions et à l'étude de la manière dont ces solutions dépendent des conditions initiales et d'éventuels paramètres.

1. Champs de vecteurs et équations différentielles autonomes

Champ de vecteurs. 1.1. — Un champ de vecteurs sur un ouvert U d'un espace de Banach E est une application $X : U \to E$ de classe C^r, $r \geqslant 0$.

Il est bon d'avoir présent à l'esprit l'interprétation suivante. Un fluide occupe l'intérieur U d'un récipient de l'espace usuel. Attachons à chaque point u de U le vecteur-vitesse $X(t, u)$ de la « molécule » du fluide passant en u à l'instant t. [Les mécaniciens le représentent par un vecteur « lié » d'origine u; c'est-à-dire, puisque l'espace usuel est un espace affine, par un bipoint d'origine u et d'extrémité $u + X(t, u)$.] Si $X(t, u)$ ne dépend pas de t [on dit que le fluide est en mouvement permanent], $u \leadsto X(u)$ est un champ de vecteur sur U, appelé le champ des vitesses du fluide.

Supposons que le mouvement soit permanent. Si $t \leadsto f(t) \in U$ est l'équation horaire d'une molécule, son vecteur-vitesse $f'(t)$ à l'instant t est encore $X(f(t))$. La connaissance des trajectoires de chaque molécule et de leur description horaire détermine donc le champ des vitesses.

C'est le problème inverse qui va nous occuper : connaissant le champ des vitesses d'un mouvement permanent, reconstituer l'équation horaire de chaque molécule.

Courbes intégrales. 1.2. — On appelle courbe intégrale du champ de vecteur X une application différentiable $f : I \to U$ d'un intervalle ouvert I dans U, telle que $f'(t) = X(f(t))$ pour $t \in I$.

Puisque X est continu, $X \circ f$ l'est aussi et f est *ipso facto* de classe C^1.

On dit encore que f est solution de l'équation différentielle du premier ordre $x' = X(x)$. Résoudre cette équation c'est, par définition, trouver toutes les courbes intégrales de X.

Connaissant la position x_0 à l'instant t_0 d'une molécule du fluide et connaissant le champ des vitesses, on peut espérer que la trajectoire de la molécule s'en trouve déterminée. Cela revient à chercher une solution $f : I \to U$ de $x' = X(x)$ vérifiant $f(t_0) = x_0$. C'est ce qu'on appelle le problème de Cauchy.

Elimination du temps. 1.3. — Soient J un intervalle ouvert, U un ouvert d'un espace de Banach E et $X : J \times U \to E$ une application de classe C^r, $r \geqslant 0$. Pour chaque $t \in J$ l'application $u \in U \leadsto X(t, u) \in E$ est un champ de vecteur sur U. Le mouvement du fluide précité n'est plus permanent et le champ des vitesses dépend du temps.

Définissons $\overline{X} : J \times U \to \mathbf{R} \times E$ par $\overline{X}(t, u) = (1, X(t, u))$ et considérons $\overline{X}$ comme un champ de vecteurs de l'ouvert $J \times U$ de l'espace de Banach $\mathbf{R} \oplus E$ (c'est l'espace-temps. du fluide). Les courbes intégrales $s \leadsto (u(s), f(s))$ de $\overline{X}$ vérifient donc $\dfrac{du}{ds} = 1$, $\dfrac{df}{ds} = X(u(s), f(s))$. Par conséquent $t = u(s) = s +$ constante, et les courbes intégrales de $\overline{X}$ telles que $u(0) = 0$ s'écrivent $t \leadsto (t, f(t))$; en sorte que $\dfrac{df}{dt} = X(t, f(t))$. La projection sur E de cette courbe intégrale $(t, f(t) \in \mathbf{R} \times E)$ est donc solution de l'équation différentielle $x' = X(t, x)$. On dit que cette équation est *non autonome*, par opposition aux équations $x' = X(x)$, dites *autonomes*, où t ne figure pas explicitement dans X.

Réciproquement, si $t \rightsquigarrow f(t)$ est une solution de $x' = X(t, x)$, alors $t \rightsquigarrow (t, f(t))$ est évidemment une courbe intégrale de $\overline{X}$. L'étude des champs de vecteurs $X(t, x)$ « dépendant du temps » est donc ramenée à celle des champs de vecteurs indépendants du temps.

Elimination des paramètres. 1.4. — Soient V un ouvert d'un espace de Banach P, U un ouvert d'un espace de Banach E, et $X : U \times V \to E$ une application de classe C^r, $r \geqslant 0$. Pour chaque $v \in V$ l'application $u \in U \rightsquigarrow X(u, v) \in E$ est un champ de vecteur dépendant du paramètre v.

Définissons $\overline{X} : U \times V \to E \times P$ par $\overline{X}(u, v) = (X(u, v), 0)$ et considérons $\overline{X}$ comme un champ de vecteur sur l'ouvert $U \times V$ de l'espace de Banach $E \oplus P$.

Les courbes intégrales $t \rightsquigarrow (f(t), v(t))$ de $\overline{X}$ vérifient donc $f'(t) = X(f(t), v(t))$, $v'(t) = 0$. Donc $v(t) = v$ (constante) et $f'(t) = X(f(t), v)$. La projection sur E de la courbe intégrale $(f(t), v) \in E \oplus P$ de $\overline{X}$ est donc solution de l'équation différentielle $x' = X(x, v)$. La réciproque est manifeste : si $t \rightsquigarrow f(t)$ est une solution de cette dernière équation, alors $t \rightsquigarrow (f(t), v)$ est une courbe intégrale de $\overline{X}$.

L'étude des équations dépendant de paramètres se ramène donc à celle des équations autonomes $x' = X(x)$.

Réduction au premier ordre. 1.5. — Soient E un espace de Banach, U un ouvert de l'espace de Banach $E \oplus E$ et $F : U \to E$ une application de classe C^r, $r \geqslant 0$. Cela définit un champ de vecteurs $(x, y) \in U \rightsquigarrow (y, F(x, y)) \in E \oplus E$, que nous noterons X. Une courbe intégrale $t \rightsquigarrow (f(t), g(t)) \in E \oplus E$ de X vérifie donc $f'(t) = g(t)$, $g'(t) = F(f(t), g(t))$. Cela montre que f est différentiable et que $f''(t) = F(f(t), f'(t))$. En d'autres termes $t \rightsquigarrow f(t)$ est solution de l'équation différentielle du second ordre $x'' = F(x, x')$.

Le procédé est général ; il a déjà été utilisé en (6. chap. 5) et (4. chap. 6). Par exemple, une équation différentielle d'ordre n :

$$\frac{d^n x}{dt^n} = F\left(x, \frac{dx}{dt}, \ldots, \frac{d^{n-1}x}{dt^{n-1}}\right) \quad \text{où } x \in E,$$

se ramène à un système autonome sur un ouvert de $E^n = E \oplus \cdots \oplus E$:

$$\frac{dx_1}{dt} = x_2, \ldots, \frac{dx_{n-1}}{dt} = x_n, \quad \frac{dx_n}{dt} = F(x_1, \ldots, x_n).$$

Remarque. — La réduction au premier ordre n'est pas canonique ; on peut ramener de bien des façons un système d'ordre n à un système du premier ordre. C'est ainsi que l'équation $x'' + \omega^2 . x = 0$ de l'oscillateur harmonique (6.1. chap. 5) s'écrit $x' = y$, $y' = -\omega^2 . x$; ou aussi bien $x' = \omega . y$, $y' = -\omega . x$.

Cette réduction est fondamentale en Mécanique, comme l'exemple qui suit le fait soupçonner.

Equations de Hamilton. — Sur un espace vectoriel réel euclidien E de dimension finie (l'espace de configuration) on se donne une fonction U de classe C^1, à valeurs réelles (le potentiel).

Rappelons (voir 1.3. chap. 1) que si $\langle \, , \, \rangle$ est le produit scalaire de E, le gradient grad $U(q)$ de U en $q \in E$ est défini par $dU(q) = \langle \operatorname{grad} U(q), . \rangle$.

Intéressons-nous à l'équation de Newton $\dfrac{d^2 q}{dt^2} = -\operatorname{grad} U(q)$, qui régit l'évolution au cours du temps t d'un système mécanique conservatif, d'énergie potentielle U. Il se ramène à un système de premier ordre $q' = p$, $p' = -\operatorname{grad} U(q)$ sur l'espace (des phases) $E \oplus E$ des positions q et des impulsions p.

Introduisons avec Lagrange la forme symplectique (bilinéaire, alternée et non dégénérée) ω, définie sur $E \oplus E$ par $\omega[(q_1, p_1), (q_2, p_2)] = \langle q_1, p_2 \rangle - \langle q_2, p_1 \rangle$. Elle définit un isomorphisme $a \in (E \oplus E)^* \to X_a \in E \oplus E$ selon $a(v) = \omega(X_a, v)$ pour $v \in E \oplus E$.

Introduisons, toujours avec Lagrange, la fonction « énergie totale » $H(q, p) = \frac{1}{2}\langle p, p \rangle + U(q)$ et cherchons l'image de sa différentielle dans l'isomorphisme ci-dessus. On trouve sans peine $X_{dH(q,p)} = (p, -\operatorname{grad} U(q))$. Par conséquent l'équation de Newton $q'' = -\operatorname{grad} U(q)$ se réduit à l'équation du premier ordre $m' = X_{dH(m)}$, où $m = (q, p)$. C'est la forme hamiltonienne des équations d'évolution, qui s'explicite dans une base orthonormée de E sous la forme familière aux mécaniciens : $q_i' = \dfrac{\partial H}{\partial p_i}$, $p_i' = -\dfrac{\partial H}{\partial q_i}$, $i = 1, \ldots, n$.

2. Existence et unicité des courbes intégrales

Applications lipschitziennes. 2.1. — Rappelons qu'une application $X : U \to F$ d'un ouvert U d'un e.v. normé E dans un e.v. normé F est dite K-lipschitzienne, s'il existe un nombre $K > 0$ tel que $\| X(x) - X(y) \|_F \leqslant K . \| x - y \|_E$ pour tous $x, y \in U$. Evidemment X est alors continue.

Le théorème de la moyenne (1. chap. 2) montre que, si X est de classe C^1, il existe pour chaque u de U un voisinage V dans U tel que la restriction de X à V soit lipschitzienne. On dit que X est localement lipschitzienne.

Théorème. 2.2. — ***Soient U un ouvert d'un espace de Banach E, $X : U \to E$ un champ de vecteurs K-lipschitzien. Donnons-nous un point x_0 de U et prenons $r > 0$ assez petit pour que la boule fermée $B_r(x_0) = \{ x \in E : \| x - x_0 \| \leqslant r \}$ soit dans U. Désignons par M une borne supérieure de $\| X(x) \|$ dans cette boule et posons $a = r/M$.***

Alors, pour chaque $t_0 \in \mathbf{R}$, il existe une application $f : [t_0 - a, t_0 + a] \to B_r(x_0)$ différentiable, et une seule telle que

(2.3) $$f'(t) = X(f(t)) \quad \textit{et} \quad f(t_0) = x_0 .$$

Preuve. — Vu la continuité de f', les conditions (2.3) se résument en

(2.4) $$f(t) = x_0 + \int_{t_0}^{t} X[f(s)] . ds .$$

Considérons l'espace des applications continues de $[t_0 - a, t_0 + a]$ dans $B_r(x_0)$ muni de la distance d de la convergence uniforme. Le sous-ensemble G de ses éléments u vérifiant $u(t_0) = x_0$ en est un sous-espace fermé ; G est donc un espace métrique complet.

Si $u \in G$, définissons Tu par $Tu(t) = x_0 + \displaystyle\int_{t_0}^{t} X[u(s)] . ds$, où $| t - t_0 | \leqslant a$. Evidemment Tu est continue et $Tu(t_0) = x_0$. D'autre part $Tu \in G$ car, d'après les définitions de M, a et r, $\| Tu(t) - x_0 \| \leqslant \left\| \displaystyle\int_{t_0}^{t} X[u(s)] . ds \right\| \leqslant M . | t - t_0 | \leqslant M . a \leqslant r$. Ainsi T applique G dans G.

En revenant à (2.4), on voit que f satisfait (2.3) si et seulement si f est un point fixe de T. Si nous montrons qu'une itérée de T est une contraction de G, le théorème sera démontré d'après (6. Appendice B).

Donnons-nous $u, v \in G$. On a

$$\| Tu(t) - Tv(t) \| = \left\| \int_{t_0}^{t} [X(u(s)) - X(v(s))] . ds \right\| \leqslant$$

$$\leqslant K . \left| \int_{t_0}^{t} \| u(s) - v(s) \| . ds \right| \leqslant K . | t - t_0 | . d(u, v).$$

Remplaçons respectivement u et v par Tu et Tv dans l'estimée précédente :

$$\| T^2 u(t) - T^2 v(t) \| \leqslant K \cdot \left| \int_{t_0}^{t} \| Tu(s) - Tv(s) \| \cdot \mathrm{d}s \right| \leqslant$$

$$\leqslant K^2 \cdot \mathrm{d}(u, v) \cdot \left| \int_{t_0}^{t} (s - t_0) \cdot \mathrm{d}s \right| \leqslant \frac{K^2 . |t - t_0|^2}{2} \cdot \mathrm{d}(u, v) .$$

De proche en proche on obtient $\| T^k u(t) - T^k v(t) \| \leqslant K^k \cdot \dfrac{|t - t_0|^k}{k!} \cdot \mathrm{d}(u, v)$. En particulier $\mathrm{d}(T^k u, T^k v) \leqslant \dfrac{(Ka)^k}{k!} \cdot \mathrm{d}(u, v)$. Si k est assez grand on a $(Ka)^k < k!$, et l'application T^k est une contraction.

Exemple. — Soit A un endomorphisme d'un espace de Banach E. Cherchons $f : \mathbf{R} \to \mathrm{End}(E)$ telle que $f'(t) = A . f(t), f(0) = 1$. Cela s'écrit

$$f(t) = 1 + A \cdot \int_0^t f(s) \cdot \mathrm{d}s .$$

En appliquant la méthode précédente, on trouve que les approximations successives $f_0(t) = 1, \ldots, f_n = T^n f_0$ de f s'expriment par $f_n(t) = 1 + tA + \cdots + t^n . A^n/n!$ Ainsi $f(t) = \mathrm{e}^{tA}$ et l'on a retrouvé le fait que $(\mathrm{e}^{tA})' = A . \mathrm{e}^{tA}$.

Corollaire. 2.5. — ***Soient U un ouvert d'un espace de Banach E, $X : U \to E$ un champ de vecteurs lipschitzien. Si $f_1 : I_1 \to U$ et $f_2 : I_2 \to U$ sont deux solutions de $x' = X(x)$ vérifiant la même condition initiale $f_1(t_0) = f_2(t_0) = x_0$, alors f_1 et f_2 coïncident sur $I_1 \cap I_2$.***

Preuve. — Conséquence immédiate de l'unicité. □

Solution maximale. 2.6. — Par définition, une solution $f : I \to U$ de $x' = X(x)$ prolonge une solution $f_1 : I_1 \to U$ si $I_1 \subset I$ et si $f(t_0) = f_1(t_0)$ pour un $t_0 \in I_1$. D'après le corollaire précédent f coïncide donc avec f_1 sur I_1, ce qui justifie la terminologie.

Par définition une solution est maximale si elle n'est pas prolongeable. Montrons qu'il en existe.

Théorème. 2.7. — ***Toute solution de $x' = X(x)$ est contenue dans une solution maximale unique.***

Preuve. — Considérons l'ensemble de toutes les solutions $f_a : I_a \to U$ qui prolongent une solution donnée $f_0 : I_0 \to U$. La réunion I des I_a est un intervalle, car c'est la réunion d'intervalles contenant I_0. Nous allons définir une solution $f : I \to U$ qui, évidemment, sera maximale.

Si $t \in I$, il existe I_a tel que $t \in I_a$. Posons $f(t) = f_a(t)$. Si t appartient à un autre intervalle I_b, puisque f_a et f_b coïncident sur I_0 on a $f_a(t) = f_b(t)$ d'après 2.4. Donc $f(t)$ ne dépend pas de f_a. On a défini une application $f : I \to U$. Elle vérifie bien $x' = X(x)$ pour chaque $t \in I$, puisqu'elle coïncide sur un intervalle I_a contenant t avec une solution f_a. □

Application. 2.8. — Le théorème précédent permet de vérifier qu'une liste de solutions est exhaustive.

Soit, par exemple, l'équation $x' = x^2, x \in \mathbf{R}$. Chacune des solutions $f_0(t) = 0$ pour $t \in \mathbf{R}$, $p_a(t) = (a - t)^{-1}$ pour $t > a$ et $n_a(t) = (a - t)^{-1}$ pour $t < a$ est maximale.

A chaque condition initiale (t_0, x_0) correspond une seule de ces solutions : pour $x_0 = 0$, c'est f_0 ; pour $x_0 > 0$ (resp. < 0), c'est n_a (resp. p_a), où $a = x_0^{-1} + t_0$.

Comme $X(x) = x^2$ est localement lipschitzienne, 2.6. s'applique et l'on a bien toutes les solutions maximales. On notera que ces solutions, à l'exclusion de f_0, ne sont pas définies sur $\mathbf{R}$ tout entier. Nous verrons ultérieurement quelles conditions imposer à X afin d'éviter cette pathologie.

Remarque. 2.9. — Si le champ de vecteur X est continu sans être localement lipschitzien, l'unicité des solutions maximales peut être en défaut.

Exemple : l'équation $x' = 3.(x^2)^{1/3}$ admet deux solutions maximales distinctes $f(t) = 0$ et $g(t) = (t - t_0)^3$ prenant la valeur zéro pour $t = t_0$. Raison : $X(x) = 3.x^{2/3}$ possède une dérivée « infinie » pour $x = 0$.

Néanmoins on a le résultat suivant, que nous ne démontrerons pas.

Théorème (Arzela). — ***Soient U un ouvert d'un espace de Banach E de* dimension finie, *$X : U \to E$ un champ de vecteurs* continu. *Alors, pour tout $t_0 \in \mathbf{R}$ et tout $x_0 \in U$, il existe au moins une solution $f : I \to U$ de l'équation $x' = X(x)$, telle que $f(t_0) = x_0$, $t_0 \in I$.***

Traduction des résultats précédents pour les équations différentielles d'ordre quelconque. 2.10. — Reprenons l'équation différentielle d'ordre n de 1.5. : $x^{(n)} = F(x, x', \ldots, x^{(n-1)})$. Supposons que l'application F soit lipschitzienne par rapport à l'ensemble des variables $x, \ldots, x^{(n-1)}$. Alors, étant donnés $t_0 \in \mathbf{R}$ et $x_0, x_1, \ldots, x_{(n-1)} \in E$, il existe une solution maximale f et une seule telle que $f(t_0) = x_0, f'(t_0) = x_1, \ldots, f^{(n-1)}(t_0) = x_{n-1}$.

Le lecteur traduira lui-même le théorème 2.2. en termes d'équations différentielles non autonomes ou dépendant de paramètres.

3. Dépendance des conditions initiales

Le théorème 2.2. résout le problème de Cauchy : étant donnés $t_0 \in \mathbf{R}$ et $x_0 \in U$, si $a > 0$ est assez petit, il existe une courbe intégrale de X unique, définie sur $[t_0 - a, t_0 + a]$ et prenant la valeur x_0 pour tout $t = t_0$. Désignons-la par $\varphi(t, x_0)$. Nous nous proposons d'étudier $x_0 \rightsquigarrow \varphi(t, x_0)$.

Lemme de Gronwall. 3.1. — ***Soient u et v deux applications continues de $[a, b]$, $a < b$, dans les réels positifs ou nuls. Supposons qu'il existe un nombre $A \geqslant 0$ tel que***

$$u(t) \leqslant A + \int_a^t u(s).v(s).\mathrm{d}s \quad \textit{pour } a \leqslant t \leqslant b. \tag{3.2}$$

Alors $u(t) \leqslant A.\exp\left[\int_a^t v(s).\mathrm{d}s\right]$ ***pour $a \leqslant t \leqslant b$.***

PREUVE. — Si $h(t)$ désigne le membre de droite de (3.2), on a $h'(t) = u(t).v(t)$. Comme, par hypothèse, $u(t) \leqslant h(t)$ et $v(t) \geqslant 0$, on a $h'(t) \leqslant h(t).v(t)$. Posons

$$C(t) = h(t).\exp\left[-\int_a^t v(s).\mathrm{d}s\right];$$

on en déduit $C'(t) \leqslant 0$, puis $C(t) \leqslant A$ pour $a \leqslant t \leqslant b$, car $C(a) = h(o) = A$. D'après la définition de C cela s'écrit $h(t) \leqslant A.\exp\left[\int_a^t v(s).\mathrm{d}s\right]$. Le lemme en résulte car $u(t) \leqslant h(t)$. □

Théorème. 3.3 (*Cas lipschitzien*). — ***Gardons les hypothèses du théorème 2.2. Alors, pour chaque $x \in B_{r/2}(x_0)$, il existe une courbe intégrale $\varphi(t, x)$ et une seule de $x' = X(x)$, définie sur $[t_0 - a/2, t_0 + a/2]$ et vérifiant $\varphi(t_0, x) = x$.***

En outre $\| \varphi(t, x) - \varphi(t, y) \| \leqslant e^{K.|t-t_0|}.\| x - y \|$ pour $x, y \in B_{r/2}(x_0)$ et $|t - t_0| \leqslant a/2$. En particulier $x \rightsquigarrow \varphi(t, x)$ est lipschitzienne sur $B_{r/2}(x_0)$ uniformément en t :

$$\| \varphi(t, x) - \varphi(t, y) \| \leqslant e^{K.a/2}.\| x - y \|.$$

PREUVE. — La première partie n'est qu'une reformulation de 2.2. Si $x \in B_{r/2}(x_0)$, on remplace dans 2.2. le point x_0 par le point x et $B_r(x_0)$ par $B_{r/2}(x)$ [qui appartient encore à l'ouvert U]. Enfin, on remplace dans 2.2. le nombre a par $a/2$ afin que $\varphi(t, x)$ demeure dans $B_{r/2}(x_0)$.

Si $x, y \in B_{r/2}(x_0)$, $\varphi(t, x)$ et $\varphi(t, y)$ vérifient (2.4), à des changements de notations près. Posons $u(t) = \| \varphi(t, x) - \varphi(t, y) \|$. Puisque X est K-lipschitzien on a

$$u(t) = \left\| x - y + \int_{t_0}^{t} [X(\varphi(s, x)) - X(\varphi(s, y))].\mathrm{d}s \right\| \leqslant$$

$$\leqslant \| x - y \| + K . \left| \int_{t_0}^{t} \| \varphi(s, x) - \varphi(s, y) \| . \mathrm{d}s \right| = \| x - y \| + K . \left| \int_{t_0}^{t} u(s).\mathrm{d}s \right| ,$$

et le théorème résulte du lemme de Gronwall si $t \geqslant t_0$.

Si $t_0 > t$, on se ramène au cas précédent en changeant t en $-t$ et X en $-X$. □

Théorème. 3.4. (classe C^1). — ***Gardons les hypothèses du théorème 2.2., à ceci près que nous supposons X de classe C^1. Nous savons alors (voir 2.1.) que X est localement lipschitzien. Quitte à réduire l'ouvert U, nous pouvons encore supposer que X est K-lipschitzien pour un certain K. Les conclusions du théorème précédent subsistent donc. Mais nous allons montrer, qu'en outre, φ est de classe C^1.***

PREUVE. — Si l'on était assuré que φ est suffisamment différentiable, le théorème de Schwarz (1. chap. 4) et $\frac{\mathrm{d}\varphi}{\mathrm{d}t} = X \circ \varphi$ entraîneraient

$$\frac{\mathrm{d}}{\mathrm{d}t} D_2\, \varphi(t, x) = DX[\varphi(t, x)] \circ D_2\, \varphi(t, x) ,$$

avec, puisque $\varphi(t_0, x) = x$, $D_2\, \varphi(t_0, x) = 1$.

Cela conduit à étudier la solution u de

$$\frac{\mathrm{d}}{\mathrm{d}t} u(t) = DX[\varphi(t, x)].u(t) \tag{3.5}$$

vérifiant $u(t_0) = 1\ (= \mathrm{id}_E)$.

C'est une équation différentielle linéaire en u dont le coefficient $DX[\varphi(\ )]$ est continu puisque X est de classe C^1. La solution cherchée existe donc, elle est unique et elle est définie pour tout t (2.2. chap. 6). Notons-la $\psi(t, x)$. □

Montrons que ψ est continue

D'après (2.2. chap. 6) l'application $t \rightsquigarrow \psi(t, x)$ est continue, et même de classe C^1 pour tout x. Si nous montrons que $x \rightsquigarrow \psi(t, x)$ est uniformément continue en t, le résultat sera établi.

Puisque DX et φ sont continues, on peut choisir le rayon r de la boule $B_r(x_0)$ et le nombre $a > 0$ assez petits pour que $\| DX[\varphi(t, x)] \|$ soit majoré par un nombre m pour tout $t \in [t_0 - a/2, t_0 + a/2]$ et tout $x \in B_{r/2}(x_0)$.

Ceci posé, comme φ est continue, l'ensemble $\{ \varphi(t, x) : | t - t_0 | \leqslant a/2, x \in B_{r/2}(x_0) \}$ est compact. La fonction continue $\varphi(t, x) \rightsquigarrow DX[\varphi(t, x)]$ est donc uniformément continue sur ce compact : à tout $\varepsilon > 0$ correspond un $\delta > 0$ tel que

$$\sup_{|t - t_0| \leqslant a/2} \| \varphi(t, x) - \varphi(t, y) \| \leqslant \delta \tag{3.6}$$

implique

$$\sup_{|t - t_0| \leqslant a/2} \| DX[\varphi(t, x)] - DX[\varphi(t, y)] \| \leqslant \varepsilon . \tag{3.7}$$

D'après 3.3., φ est lipschitzienne en x uniformément en t; il existe donc $\delta' > 0$ tel que $\| x - y \| \leqslant \delta'$ implique (3.6) et par conséquent (3.7).

Considérons alors les fonctions $\psi(t, x)$ et $\psi(t, y)$ définies par

$$(3.8)\qquad \begin{cases} \dfrac{d\psi}{dt}(t, x) = DX[\varphi(t, x)].\psi(t, x)\,, & \psi(t_0, x) = 1\,, \\[2ex] \dfrac{d\psi}{dt}(t, y) = DX[\varphi(t, y)].\psi(t, y)\,, & \psi(t_0, y) = 1\,. \end{cases}$$

Ce sont des solutions d'équations différentielles linéaires. D'après le théorème de comparaison (2.5. chap. 6) on a donc

$$\| \psi(t, x) - \psi(t, y) \| \leqslant \frac{a}{2} \cdot e^{am/2} \cdot \sup_{|t - t_0| \leqslant a/2} \| DX[\varphi(t, x)] - DX[\varphi(t, y)] \| \leqslant a \cdot e^{am/2} \cdot \varepsilon/2$$

pour tout $| t - t_0 | \leqslant a/2$ si $\| x - y \| \leqslant \delta'$. C'est l'uniforme continuité en x cherchée.

Montrons que $D_2\, \varphi(t, x)$ existe et est égale à $\psi(t, x)$.

Posons $\theta(t, h) = \varphi(t, x + h) - \varphi(t, x)$. Puisque $\dfrac{d}{dt}\varphi = X \circ \varphi$ et que

$$\theta(t_0, h) = \varphi(t_0, x + h) - \varphi(t_0, x) = (x + h) - x = h\,,$$

on a

$$\theta(t, h) = h + \int_{t_0}^{t} \{ X[\varphi(s, x + h)] - X[\varphi(s, x)] \}.ds\,.$$

D'autre part, d'après (3.8),

$$\psi(t, x).h = h + \int_{t_0}^{t} DX[\varphi(s, x)].\psi(s, h).h.ds\,.$$

On en déduit

$$(3.9)\qquad \theta(t, h) - \psi(t, x).h = \int_{t_0}^{t} DX[\varphi(s, x)].[\theta(s, h) - \psi(s, x).h].ds + \int_{t_0}^{t} \{\ \}.ds\,,$$

où $\{\ \} = X[\varphi(s, x + h)] - X[\varphi(s, x)] - DX[\varphi(s, x)].\theta(s, h)$.

Etudions $\{\ \}$. Puisque, d'après 3.3, $x \rightsquigarrow \varphi(., x)$ est lipschitzienne uniformément en t, il existe $C > 0$ tel que $\| \theta(s, h) \| = \| \varphi(s, x + h) - \varphi(s, x) \| \leqslant C.\| h \|$ pour $| s - t_0 | < a/2$. D'autre part, puisque X est de classe C^1, étant donné $\varepsilon > 0$ il existe un $\delta > 0$ tel que $\| \theta(s,h) \| < \delta$ entraîne $\| \{\ \} \| \leqslant \varepsilon.\| \theta(s, h) \|$. Si l'on choisit h pour que $C.\| h \| < \delta$ on a donc $\| \{\ \} \| \leqslant \varepsilon.C.\| h \|$, et le second terme du membre de droite de (3.9) est dominé par $| t - t_0 |.\varepsilon.C.\| h \|$, c'est-à-dire encore par $a.\varepsilon.C.\| h \|$.

Appliquons l'inégalité de Gronwall en prenant pour u la norme du membre de gauche de (3.9), pour A le nombre $a.\varepsilon.C.\| h \|$ et pour v le majorant m de $\| DX(\ ) \|$. On obtient

$$\| \varphi(t, x + h) - \varphi(t, x) - \psi(t, x).h \| \leqslant (\text{constante}).\varepsilon.\| h \|\,.$$

Il en résulte que $D_2\, \varphi(t, x) = \psi(t, x)$.

Puisque $\varphi : (t, x) \rightsquigarrow \varphi(t, x)$ possède des différentielles partielles $D_1\, \varphi = \dfrac{d}{dt}\varphi$ et $D_2\, \varphi$ continues, il résulte de (3. chap. 2) que φ est de classe C^1.

Le lecteur trouvera à l'Appendice G une preuve courte et astucieuse, mais conceptuellement plus difficile, due à J. Robbin.

Théorème. 3.10. (classe C^k). — ***Gardons les hypothèses du théorème 2.2., à ceci près que nous supposons X de classe C^k, $k \geqslant 1$. Alors φ est de classe C^k.***

PREUVE. — Le théorème vient d'être établi si $k = 1$. Supposons $k \geqslant 2$ et admettons que le théorème soit vrai jusqu'à l'ordre $k - 1$. Alors φ est de classe C^1 et de

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t, x) = X[\varphi(t, x)]$$

on en déduit

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t, x) = DX[\varphi(t, x)].X[\varphi(t, x)]$$

et

$$\frac{\mathrm{d}}{\mathrm{d}t}D_2\,\varphi(t, x) = DX[\varphi(t, x)].D_2\,\varphi(t, x)\,.$$

Ces trois relations forment un système différentiel en l'inconnue $\left(\varphi, \frac{\mathrm{d}\varphi}{\mathrm{d}t}, D_2\,\varphi\right)$. Puisque X est de classe C^k, les membres de droite sont de classe C^{k-1} par rapport à cette inconnue. D'après l'hypothèse de récurrence, φ, $\frac{\mathrm{d}\varphi}{\mathrm{d}t}$ et $D_2\,\varphi$ sont donc de classe C^{k-1}. Il en résulte que φ est de classe C^k.

Si, en particulier, X est C^∞ alors φ est C^∞. □

Application. 3.11. — Soit A un endomorphisme d'un espace de Banach E. Alors $A \rightsquigarrow \exp(A)$ est de classe C^∞.

PREUVE. (H. Poincaré). — Cherchons la solution du système différentiel en $x, y \in \mathrm{End}(E)$: $\frac{\mathrm{d}x}{\mathrm{d}t} = y \circ x$, $\frac{\mathrm{d}y}{\mathrm{d}t} = 0$ vérifiant les conditions initiales $x(0) = 1\ (= \mathrm{id}_E)$, $y(0) = A$. On a vu au chapitre 5 que cette solution est $x(t) = \exp(tA)$, $y(t) = A$. Comme le champ de vecteurs $(x, y) \rightsquigarrow (y \circ x, 0)$ est de classe C^∞, cette solution dépend de façon C^∞ de la condition initiale A. En particulier $A \rightsquigarrow \exp(1.A)$ est C^∞. □

On peut souhaiter expliciter la différentielle de $X \rightsquigarrow \exp(X)$. Voici, sans démonstration, ce qu'on obtient.

Etant donné $A \in \mathrm{End}(E)$, désignons par $ad(A)$ l'application de $\mathrm{End}(E)$ dans $\mathrm{End}(E)$ définie par $X \rightsquigarrow A \circ X - X \circ A$. C'est une application linéaire continue car

$$\| ad(A).X \| \leqslant 2.\| A \|.\| X \|\,.$$

La différentielle $D(\exp)(A)$ de exp en A s'exprime par

$$D(\exp)(A) = \mathrm{e}^A \cdot \sum_0^\infty \frac{(-1)^{n+1}}{(n+1)!} \cdot [ad(A)]^n\,.$$

Nous aurions pu raisonner directement. Si $h \in \mathrm{End}(E)$, la série

$$\sum_1^\infty (A^{n-1}.h + A^{n-2}.h.A + \cdots + h.A^{n-1})/n\,!$$

converge normalement sur tout compact de $\mathrm{End}(E)$, car le module de son terme général est dominé par $\| A \|^{n-1}.\| h \|/(n-1)\,!$. Si $L(h)$ désigne sa somme, on voit donc que L est linéaire et que $\| L \| \leqslant \mathrm{e}^{\| A \|}$. Comme on a

$$\| \mathrm{e}^{A+h} - \mathrm{e}^A - L(h) \| \leqslant \sum_0^\infty [(\| A \| + \| h \|)^n - (\| A \|)^n - n(\| A \|)^{n-1}.\| h \|]/n\,! =$$

$$= \mathrm{e}^{\| A \| + \| h \|} - \mathrm{e}^{\| A \|} - \| h \|.\mathrm{e}^{\| h \|} = o(\| h \|)\,,$$

cela montre que exp est différentiable et que $D(\exp)(A).h = L(h)$.

4. Champs de vecteurs complets

Définition. 4.1. — Etant donné un ouvert U d'un espace de Banach E et un champ de vecteurs localement lipschitzien X sur U, on dit que X est complet si chaque courbe intégrale maximale de X est définie sur $\mathbf{R}$ tout entier. Autrement dit, si, pour chaque $t_0 \in \mathbf{R}$ et chaque $x_0 \in U$, il existe une application différentiable $f : \mathbf{R} \to U$ vérifiant $f'(t) = X[f(t)]$ et $f(t_0) = x_0$.

Exemple. 4.2. — D'après les chapitres 4 et 5, les solutions maximales des équations différentielles linéaires sont définies sur $\mathbf{R}$ tout entier.

Exemples. 4.3. — Sur $E = \mathbf{R}$ le champ de vecteur constant $X = 1$ est complet. Par contre, le même champ, considéré sur l'ouvert $U = \mathbf{R} - \{ 0 \}$, n'est pas complet : la solution générale $f(t) = t - c$, où c est une constante, n'est pas définie pour $t = c$ puisqu'il est interdit à f de prendre la valeur nulle.

Nous avons déjà vu que le champ de vecteur $X(x) = (x)^2$ sur $\mathbf{R}$ n'est pas complet : les solutions $(c - t)^{-1}$, où c est une constante, ne sont pas définies pour $t = c$.

Dans le premier exemple le champ n'est pas complet parce qu'il manque un point à l'espace $\mathbf{R} - \{ 0 \}$ (qui est incomplet !). Dans le second exemple, le champ est incomplet parce que $X(x)$ croît trop vite avec x.

En Mécanique, les trajectoires des points matériels sont solutions d'équations différentielles. Il est essentiel que ces solutions existent pour toutes les valeurs du temps t. C'est pourquoi nous allons donner quelques critères de complétude.

Théorème. 4.4. — ***Soit X un champ de vecteurs localement lipschitzien défini sur un ouvert U d'un espace de Banach E. Soit $f :]T_-, T_+[\to U$ une courbe intégrale maximale de X. Alors, ou bien $T_+ = +\infty$, ou bien $T_+ < +\infty$ et, dans ce dernier cas, à tout compact K de U correspond un $\varepsilon > 0$ tel que $f(t) \notin K$ si $t > T_+ - \varepsilon$. En d'autres termes $f(t)$ finit par sortir de tout compact.***

La propriété correspondante vaut pour T_-.

Preuve. — Remarquons d'abord qu'une courbe intégrale maximale est définie sur un intervalle d'après 2.7. Soit $]T_-, T_+[$ cet intervalle.

Supposons $T_+ < +\infty$. Raisonnons par l'absurde et supposons que le ε de l'énoncé n'existe pas. On peut donc trouver une suite de nombres réels t_n convergeant vers T_+ et telle que $f(t_n)$ demeure dans K. Puisque K est compact, quitte à extraire une sous-suite de t_n, on peut supposer que $f(t_n)$ converge vers un point x_0 de K. D'après le théorème 3.3. il existe un voisinage $U = B_{r/2}(x_0)$ de x_0 et un nombre $a > 0$ tels que pour chaque $x \in U$ il existe une courbe intégrale $\varphi(t, x)$ définie sur $[t_0 - a/2, t_0 + a/2]$. Prenons n assez grand pour que $T_+ - t_n < a/2$ et $f(t_n) \in U$. On aboutit à la contradiction suivante : la courbe intégrale maximale passant par $f(t_n)$ à l'instant t_n n'est définie que pour $t < T_+$, et pourtant on peut la prolonger jusqu'à l'instant $t_n + \frac{a}{2} > T_+$. □

Intégrale première. 4.5. — Gardons les notations de 4.4. On dit qu'une fonction $h : U \to \mathbf{R}$ est une intégrale première du champ X si, quelle que soit la courbe intégrale $f : I \to U$ de X, la fonction $h \circ f$ définie sur I ne dépend pas de t [par contre elle peut dépendre de la solution f].

Voici un critère simple pour décider si une fonction h est une intégrale première, sans avoir à connaître explicitement les courbes intégrales f.

Critère. — Une fonction $h : U \to \mathbf{R}$ de classe C^1 est une intégrale première de X si, et seulement si, $Dh(x).X(x) = 0$ pour tout $x \in U$.

Preuve. — Soient h une intégrale première et $x \in U$. Si f est une courbe intégrale de X telle que $f(a) = x$, on a $h[f(t)] = h[f(a)] = h(x)$. En différentiant par rapport à t on obtient

$$Dh[f(t)].f'(t) = Dh[f(t)].X[f(t)] = 0 .$$

Prenons en particulier $t = a$, on a $Dh(x).X(x) = 0$.

Réciproquement, supposons que $Dh(x).X(x) = 0$ pour tout x de U. Le calcul précédent montre que la dérivée de $t \rightsquigarrow h[f(t)]$ est nulle. Comme t décrit un intervalle connexe, il résulte de (2. chap. 2) que $h[f(t)]$ est constante.

Revenons à la complétude des champs de vecteurs. □

Corollaire. 4.6. — ***Soit X un champ de vecteurs localement lipschitzien défini sur un ouvert U d'un espace de Banach E. Si X possède une intégrale première $h : U \to \mathbf{R}$ telle que $h^{-1}(r)$ soit compact pour chaque $r \in \mathbf{R}$, alors X est complet.***

PREUVE. — Chaque courbe intégrale passant par $x \in U$ reste sur le compact $h^{-1}[h(x)]$. Le théorème résulte alors de 4.3. □

Remarques. — *a*) Il se peut que la fonction h soit définie sur un sous-ensemble plus grand que U ; sur E tout entier, par exemple. S'il en est ainsi, l'hypothèse du théorème signifie que $h^{-1}(r) \cap U$ doit être un compact de U.

b) Si h est différentiable, il est facile de voir que les hypothèses du corollaire impliquent que E est de dimension finie.

EXEMPLE. — Le champ de vecteur $X = (y - z, z - x, x - y)$ défini sur $\mathbf{R}^3 = \{(x, y, z)\}$ admet pour intégrale première $h(x, y, z) = (x)^2 + (y)^2 + (z)^2$, car $Dh.X = 2\,x.(y - z) + 2\,y.(z - x) + 2\,z.(x - z) = 0$. Comme $h^{-1}(r)$ est vide si $r < 0$ et que c'est la sphère de centre O et de rayon $r^{1/2}$ si $r > 0$, le champ X est complet.

Théorème de majoration a priori. 4.7. — ***Reprenons les hypothèses de 4.6. Soit $f :]T_-, T_+[\to U$ une courbe intégrale maximale de X. Supposons que pour chaque nombre réel $T > 0$ il existe un compact K_T de U tel que $f(t) \in K_T$ pour $|t| \leqslant T$. Alors $T_- = -\infty$ et $T_+ = +\infty$.***

PREUVE. — Sur l'ouvert $\mathbf{R} \times U$ de l'espace de Banach $\mathbf{R} \oplus E$ considérons le champ de vecteur $(1, X)$, c'est-à-dire le système $\dfrac{\mathrm{d}t}{\mathrm{d}s} = 1$, $\dfrac{\mathrm{d}x}{\mathrm{d}t} = X(x)$. Supposons que $T_+ < +\infty$. Prenons $T > T_+$ et formons le compact $[-T, +T] \times K_T$ de $\mathbf{R} \times U$. D'après le théorème 4.4. appliqué au champ $(1, X)$, la courbe intégrale maximale $]T_-, T_+[\to \mathbf{R} \times U$, définie par $s \rightsquigarrow (s, f(s))$, sort de ce compact. Comme $f(s)$ reste dans K_T, c'est qu'il existe s tel que $s > T > T_+$: en contradiction avec le fait que $f(s)$ n'est pas défini pour $s \geqslant T_+$.

Même démonstration pour T_-.

Ainsi une majoration a priori $f(t) \subset K_T$ d'une solution f sur l'intervalle $|t| \leqslant T$ entraîne l'existence de cette solution sur tout l'intervalle. □

EXEMPLE. 4.8. — Un champ de vecteurs X localement lipschitzien, défini sur E tout entier et constant ($= X_0$) en dehors d'un compact est complet.

PREUVE. — Les solutions maximales f vérifient $\| f(t) \| \leqslant A + |t|.\| X_0 \|$, où A est une constante (le détailler). □

EXEMPLE. 4.9. (V. Arnold). — Reprenons l'équation de Newton $\dfrac{\mathrm{d}^2 q}{\mathrm{d}t^2} = -\operatorname{grad} U(q)$ de 1.5. et supposons l'énergie potentielle U partout positive. Alors chaque solution $q(t)$ est définie sur $\mathbf{R}$ tout entier.

PREUVE. — L'équation de Newton s'écrit $\dfrac{\mathrm{d}q}{\mathrm{d}t} = p$, $\dfrac{\mathrm{d}p}{\mathrm{d}t} = -\operatorname{grad} U(q)$, où $q, p \in E$.

Observons d'abord que l'énergie totale $H(q, p) = \dfrac{1}{2}\| p \|^2 + U(q)$ est une intégrale première d'après 4.5. : $DH.X = \langle \operatorname{grad} U(q), p \rangle - \langle p, \operatorname{grad} U(q) \rangle = 0$.

Considérons la solution maximale $(q(t), p(t))$ telle que $q(0) = q_0$, $p(0) = p_0$ et posons $H_0 = H(q_0, p_0)$. De $\frac{1}{2}\| p \|^2 + U(q) = H_0$ et $U(q) \geqslant 0$, on déduit $\left\| \frac{dq}{dt} \right\| = \| p \| \leqslant \sqrt{2.H_0}$, donc $\| q(t) - q_0 \| \leqslant \sqrt{2 H_0}.| t |$. Par conséquent, si $| t | \leqslant T$, la solution $(q(t), p(t))$ reste dans le compact $\{ (q, p) : \| q \| \leqslant \| q_0 \| + \sqrt{2 H_0}.T, \| p \| \leqslant \sqrt{2 H_0} \}$ et il suffit d'appliquer 4. 6.

Remarques. 4.9. — On peut montrer que s'il existe une constante $k \geqslant 0$ telle $U(q) \geqslant -k.\| q \|^2$, alors chaque solution est définie sur $\mathbf{R}$ tout entier.

Par contre, prenons $E = \{ q \} = \mathbf{R}$ et $U(q) = -q^4/2$. La solution $q(t) = (t-1)^{-1}$ ne peut être prolongée jusqu'à $t = 1$.

Le lecteur trouvera d'autres critères de complétude dans l'ouvrage de R. Abraham et J. Marsden (page 71).

5. Groupes à un paramètre de difféomorphismes

Les groupes à un paramètre d'automorphismes linéaires d'un espace de Banach et leur relation avec les équations différentielles linéaires homogènes à coefficients constants ont été étudiés en (3. chap. 5). Nous allons généraliser cette étude.

Groupes à un paramètre de difféomorphismes. 5.1. — On appelle groupe à un paramètre $t \in \mathbf{R}$ de difféomorphismes φ_t d'un ouvert U d'un espace de Banach E une application $\varphi : \mathbf{R} \times U \to U$ telle que :

a) φ soit de classe C^r, $r \geqslant 1$;

b) pour chaque $t \in \mathbf{R}$ l'application $\varphi_t : U \to U$, définie par $\varphi_t(x) = \varphi(t, x)$, soit un difféomorphisme ;

c) la famille φ_t, $t \in \mathbf{R}$, soit un groupe à un paramètre de transformations de U : $\varphi_0 = \mathrm{id}_U$, c'est-à-dire $\varphi_0(x) = x$ pour $x \in U$; $\varphi_s \circ \varphi_t = \varphi_{s+t}$ pour $s, t \in \mathbf{R}$, c'est-à-dire

$$\varphi(s, \varphi(t, x)) = \varphi(t + s, x)$$

pour $s, t \in \mathbf{R}$ et pour $x \in U$.

Observons que le difféomorphisme inverse de φ_t est φ_{-t} car $\varphi_t \circ \varphi_{-t} = \varphi_{t-t} = \varphi_0 = \mathrm{id}_U$.

Fixons un point x de U. On peut considérer la courbe $t \in \mathbf{R} \rightsquigarrow \varphi_t(x) = \varphi(t, x) \in U$ comme définissant un mouvement d'un point dont la position à l'instant initial 0 est $\varphi(0, x) = x$ et dont la position à l'instant t est $\varphi_t(x)$. L'orbite de ce point est donc $\{ \varphi_t(x) : t \in \mathbf{R} \}$.

Générateur du groupe 5.2. — On appelle vitesse $X(x)$ de φ_t au point x de U le vecteur-vitesse à l'instant $t = 0$ de $t \rightsquigarrow \varphi_t(x) = \varphi(t, x)$:

$$X(x) = \frac{d}{dt} \varphi(t, x) \Big|_{t=0} .$$

On voit donc que si φ est de classe C^r, $r \geqslant 1$, on a défini un champ de vecteur $X : U \to E$ de classe C^{r-1} dont les courbes intégrales sont $t \rightsquigarrow \varphi(t, x)$. On dit que X est le générateur (infinitésimal) du groupe à un paramètre φ (qu'on appelle aussi un flot).

Exemple. — $A \in \mathrm{End}(E)$ est le générateur du flot $\varphi_t(x) = e^{tA}.x$.

Théorème. 5.3. — ***Le vecteur-vitesse du mouvement d'un point à chaque instant est égal au vecteur de la vitesse du flot à l'endroit où se trouve le point à l'instant considéré. En d'autres termes :***

$$\frac{d}{dt} \varphi_t(x) \Big|_{t=s} = X[\varphi_s(x)] .$$

PREUVE.

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t, x)\Big|_{t=s} = \frac{\mathrm{d}}{\mathrm{d}t}\varphi(s + r, x)\Big|_{r=0} = \frac{\mathrm{d}}{\mathrm{d}t}\varphi[r, \varphi(s, x)]\Big|_{r=0} = X[\varphi(s, x)] .$$

Le générateur X d'un flot φ est un champ de vecteurs complet puisque les courbes intégrales $t \rightsquigarrow \varphi(t, x)$ sont définies pour tout $t \in \mathbf{R}$. Nous allons montrer que, réciproquement, un champ de vecteur complet engendre un flot. En résumé, tout comme en (3. chap. 5), on aura montré qu'il existe une correspondance biunivoque entre les flots sur U et les champs de vecteurs complets sur U.

Théorème. 5.4. — ***Soit X un champ de vecteurs complet de classe C^r, $r \geqslant 1$, sur un ouvert U d'un espace de Banach E. Pour chaque $x \in U$, désignons par $t \rightsquigarrow \varphi(t, x)$ la courbe intégrale maximale de X telle que $\varphi(0, x) = x$. Alors $\varphi : \mathbf{R} \times U \to U$ est un groupe à un paramètre de difféomorphismes de U.***

PREUVE. — D'après (3.10.) φ est de classe C^r ; la condition a de 5.1. est donc vérifiée.

Vérifions la condition c. D'abord, par définition, $\varphi(0, x) = x$ pour tout x. Ensuite, donnons-nous $t, s \in \mathbf{R}$ et posons $f(t) = \varphi(t + s, x)$, $r = t + s$. On a

$$\frac{\mathrm{d}}{\mathrm{d}t}f(t) = \frac{\mathrm{d}}{\mathrm{d}t}\varphi(t + s, x) = \frac{\mathrm{d}}{\mathrm{d}r}\varphi(r, x) = X[\varphi(r, x)] = X[\varphi(t + s, x)] = X[f(t)] .$$

Par conséquent f est la courbe intégrale maximale de X telle que $f(0) = \varphi(s, x)$. Il en va évidemment de même, par définition de φ, pour $t \rightsquigarrow \varphi(t, \varphi(s, x))$. D'après l'unicité

$$\varphi(t + s, x) = \varphi(t, \varphi(s, x)) .$$

Quant à la condition b de 5.1. elle est en fait une conséquence des conditions a et b. Puisque φ est de classe C^r, $r \geqslant 1$, l'application $\varphi_t : x \rightsquigarrow \varphi(t, x)$ est de classe C^r. Nous avons vu qu'elle est inversible et son inverse $(\varphi_t)^{-1} = \varphi_{-t}$ est de classe C^r. Il en résulte bien que φ_t est un difféomorphisme de classe C^r de U sur U. □

Remarque. 5.5. — Si X n'est pas complet, la démonstration précédente montre que pour chaque $x_0 \in U$, il existe un voisinage V de x_0 contenu dans U et un intervalle $]-a, a[$ tels que :

a) $\varphi :]-a, a[\times V \to U$ vérifie

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t, x) = X[\varphi(t, x)] \quad \text{pour tout } x \in V ;$$

b) pour chaque $t \in]-a, a[$, $\varphi_t : x \rightsquigarrow \varphi(t, x)$ soit un difféomorphisme de V sur $\varphi_t(V)$;

c) pour $t, s, t + s \in]-a, a[$ on ait $\varphi_s \circ \varphi_t = \varphi_{s+t}$.

On résume ces propriétés en disant que φ définit un germe de groupe à un paramètre de difféomorphismes dans un voisinage de x_0.

Notes. — Le présent chapitre ne fait qu'effleurer la question. Le lecteur trouvera des exposés plus complets dans les ouvrages déjà cités de V. Arnold, E. A. Coddington et N. Levinson, de N. Rouche et J. Mawhin. Il trouvera aussi une autre démonstration du théorème d'existence dans l'ouvrage de H. Cartan. Enfin, les ouvrages de R. Abraham et J. E. Marsden et de M. W. Hirsch et S. Smale contiennent des développements récents de l'étude globale des systèmes dynamiques.

8 CONJUGAISON ET COORDONNÉES LOCALES

Ce chapitre est consacré à la conjugaison dans le groupe des difféomorphismes. Cette notion est illustrée, entre autres, par le théorème du rang constant et par le théorème du redressement des champs de vecteurs.

Préliminaires

En géométrie euclidienne les figures sont classées en figures égales, c'est-à-dire superposables par déplacement. On reconnaît que deux figures sont égales en montrant qu'elles ont en commun un certain nombre de propriétés invariantes par déplacement (songez aux cas d'égalité des triangles). De même, deux transformations de l'espace sont considérées comme équivalentes si l'action de l'une est superposable à l'action de l'autre par un déplacement. C'est ainsi que deux rotations du plan euclidien sont équivalentes si leurs angles de rotation sont égaux.

On peut étendre ce point de vue à des groupes plus généraux que le groupe euclidien.

A titre de second exemple, examinons le groupe $\mathrm{GL}(E)$ des automorphismes d'un e.v. E. Deux endomorphismes A, $B \in \mathrm{End}(E)$ sont considérés comme équivalents (on dit « semblables ») s'il existe $S \in \mathrm{GL}(E)$ tel que $B = S.A.S^{-1}$.

Comment reconnaître que A et B sont semblables ? Si l'on suppose que $E = \mathbf{C}^n$ et si l'on se limite à l'ouvert dense U de $\mathrm{End}(E)$ formé par les endomorphismes à valeurs propres distinctes (voir l'Appendice E), $A \in U$ et $B \in U$ sont semblables si, et seulement si, ils ont le même spectre $\{ \lambda_1, \ldots, \lambda_n \}$. En effet, si S est l'automorphisme qui envoie la base de E formée par les vecteurs propres de A sur celle formée par les vecteurs propres correspondants de B, on a bien $B = S.A.S^{-1}$.

Cette classification offre une méthode pour effectuer des opérations sur $\mathrm{End}(E)$, dès que ces opérations commutent avec la conjugaison. C'est ainsi que l'endomorphisme A ci-dessus est semblable à l'endomorphisme B dont la matrice dans la base canonique de $\mathbf{C}^n$ est $\mathrm{diag}(\lambda_r)$. L'opération d'élévation à la puissance entière k s'en trouve notablement simplifiée :

$$A^k = S.[\mathrm{diag}(\lambda_r)]^k.S^{-1} = S.\mathrm{diag}(\lambda_r^k).S^{-1}\ .$$

Si les valeurs propres λ_r sont positives, cette formule fournit, en prime, la possibilité d'étendre l'opération d'élévation à la puissance k au cas où k est un réel quelconque.

Dans ce chapitre, nous allons nous intéresser au groupe des difféomorphismes de classe C^k d'un espace de Banach E. Nous introduirons la notion d'équivalence correspondante (C^k-conjugaison) et nous chercherons les invariants qui permettent de décider si deux morphismes de E sont équivalents.

1. C^k-conjugaison et coordonnées

C^k-conjugaison. 1.1. — Soient f une application de classe C^k, $k \geqslant 1$, d'un espace de Banach E dans un espace de Banach E', f_1 une application de classe C^k d'un espace de Banach E_1 dans un espace de Banach E_1'. On dit que f et f_1 sont C^k-conjuguées s'il existe des difféomorphismes $\varphi : E \to E_1$ et $\varphi' : E' \to E_1'$ de classe C^k tels que le diagramme suivant soit commutatif

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[d, "\varphi"'] & E' \arrow[d, "\varphi'"] \\ E_1 \arrow[r, "f_1"] & E'_1 \end{tikzcd}$$

En d'autres termes, si $f_1 = \varphi' \circ f \circ \varphi^{-1}$.

C^k-conjugaison locale. 1.2. — La propriété précédente peut n'être vérifiée que localement Gardons les notations précédentes. Soient $a \in E$, $a_1 \in E_1$; posons $a' = f(a)$, $a'_1 = f_1(a_1)$. On dit que f et f_1 sont localement C^k-conjuguées aux voisinages de a et a_1 s'il existe des voisinages ouverts de a, a_1, a', a'_1, notés respectivement U, U_1, U', U'_1 tels que

a) $f(U) \subset U', f_1(U_1) \subset U'_1$,

b) les restrictions de f à U et de f_1 à U_1 soient C^k-conjuguées :

$$\begin{tikzcd} U \arrow[r, "f/U"] \arrow[d, "\varphi"'] & U' \arrow[d, "\varphi'"] \\ U_1 \arrow[r, "f_1/U_1"] & U'_1 \end{tikzcd}$$

Le composé de difféomorphismes étant un difféomorphisme, la C^k-conjugaison (locale) est une relation d'équivalence.

EXEMPLE. — Le diagramme suivant montre qu'un C^k-difféomorphisme $f : E \to E'$ est C^k-conjugué à l'application identique :

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[d, "\mathrm{id}_E"'] & E' \arrow[d, "f^{-1}"] \\ E \arrow[r, "\mathrm{id}_E"] & E' \end{tikzcd}$$

Nous allons traduire ce qui précède en termes de coordonnées.

Coordonnées. 1.3. — Soit U un ouvert d'un e.v. E de dimension finie n sur K (nous prendrons souvent $K = \mathbf{R}$, pour fixer les notations). Si $\{ e_k \}$ est une base de E, chaque élément u de U s'écrit de façon unique $\sum u_k . e_k$. Notons x_k l'application $u \rightsquigarrow u_k$ de U dans K. Les fonctions $x_1, \ldots, x_n$ sont les composantes d'une application φ de U dans K^n, qui est la restriction à U d'un isomorphisme de E sur K^n :

$$\varphi : \; U \to K^n$$
$$u \to (u_1 = x_1(u), \ldots, u_n = x_n(u)) .$$

On dit que $(x_1, \ldots, x_n)$ est un système de coordonnées linéaires sur U.

Bien entendu $\varphi : U \to \varphi(U)$ est un difféomorphisme. Cela nous amène à généraliser la notion de coordonnées.

Nous appellerons *carte de domaine* U, le couple (U, φ) formé par un ouvert U et un difféomorphisme φ de U sur un ouvert de K^n. Les composantes $(x_1, \ldots, x_n)$ de φ s'appellent un *système de coordonnées* et $(x_1(u), \ldots, x_n(u))$ s'appellent les coordonnées de $u \in U$ dans la carte considérée. L'application inverse $\varphi^{-1} : \varphi(U) \subset K^n \to U$ s'appelle une *paramétrisation de* U.

EXEMPLE : COORDONNÉES POLAIRES. — Prenons $E = \mathbf{R}^2 = \{ (x, y) \}$, $U = \{ (x, y) : x > 0. y > 0 \}$.

Alors $\varphi : u = (x, y) \rightsquigarrow (\rho = [x^2 + y^2]^{1/2}, \theta = \text{Arctg}\, y/x) \in \mathbf{R}^2$ est une carte sur U. On appelle $(\rho(u), \theta(u))$ les coordonnées polaires de u.

On constate que l'image inverse par φ des droites ρ = constante de $\mathbf{R}^2$ sont des cercles. D'où le nom de coordonnées curvilignes parfois donné aux coordonnées quelconques.

L'image par φ d'une figure de U s'appelle la *lecture* de cette figure *dans la carte*. Les cercles de centre O de $E = \mathbf{R}^2$ se lisent donc comme des droites en coordonnées polaires.

Si $f : U \to K$ est une fonction, son expression dans les nouvelles coordonnées (x_k) [ou, si l'on préfère, sa lecture dans la carte (U, φ)] est, par définition, $f \circ \varphi^{-1}$.

Soit encore f une application de classe C^k d'un ouvert U d'un espace de Banach E de dimension n dans un ouvert U' d'un espace de Banach E' de dimension m. Soient $\varphi : U \to K^n$ une carte sur U (avec des coordonnées x_r), $\varphi' : U' \to K^m$ une carte sur U' (avec des coordonnées y_s). L'application $\varphi' \circ f \circ \varphi^{-1}$ s'appelle l'expression (ou la lecture) de f dans les cartes (U, φ). (U', φ'). C'est une application de classe C^k de l'ouvert $\varphi(U)$ de K^n dans l'ouvert $\varphi'(U')$ de K^m, qui, par construction, est C^k-conjuguée à l'application f. Elle fournit donc un *modèle* de f dans sa classe de C^k-conjugaison.

Difféomorphisme local. Coordonnées locales. 1.4. — Rappelons qu'une application f de classe C^k, $k \geqslant 1$, d'un ouvert U d'un espace de Banach E dans un espace de Banach F, est appelée un difféomorphisme local en $a \in U$ s'il existe un voisinage ouvert V de a dans U tel que la restriction de f à V soit un difféomorphisme de V sur $f(V)$.

Le théorème d'inversion locale (2. chap. 3) se reformule ainsi : f est un difféomorphisme local en a si, et seulement si, la différentielle $Df(a)$ est un isomorphisme de E sur F.

Ceci rappelé, « localisons » les notions de 1.3.

Définition. — Soient E un e.v. de dimension n sur K, U un ouvert de E, a un point de U. On appelle système de coordonnées locales en a la donnée de n fonctions $x_i : U \to K$ de classe C^k telles que $u \in U \rightsquigarrow (x_1(u), ..., x_n(u)) \in K^n$ soit un C^k-difféomorphisme local en a.

On déduit immédiatement du théorème d'inversion locale, sous la forme que nous venons de lui donner, le résultat suivant :

Théorème. 1.5. — ***$(x_1, ..., x_n)$ est un système de coordonnées locales en a si, et seulement si, le déterminant de la matrice $(D_i\, x_j(a))$ n'est pas nul. Ou encore, si, et seulement si, les formes linéaires $Dx_1(a), ..., Dx_n(a)$ forment une base de l'espace dual E^*.***

La notion de coordonnées locales permet d'exprimer commodément celle de C^k-conjugaison locale. Reprenons les notations de 1.2. et supposons que E et E' soient de dimensions finies. Alors f et f_1 sont localement C^k-conjuguées aux voisinages de a et a_1 si leurs expressions dans des cartes adéquates sont les mêmes.

EXEMPLE. — Soient E et F deux e.v. de même dimension n, U un ouvert de E, $f : U \to F$ un C^k-difféomorphisme local en $a \in U$. Choisissons des coordonnées locales $(y_1, ..., y_n)$ en $b = f(a)$. Alors $(x_1 = y_1 \circ f, ..., x_n = y_n \circ f)$ est un système de coordonnées locales en a, car le composé de deux difféomorphismes est un difféomorphisme. Il en résulte que f, lue dans les coordonnées (x_r), (y_s) est l'application identique (comparer à l'exemple de 1.2.).

Nous allons généraliser ce résultat, en ne supposant plus que $Df(a)$ est un isomorphisme de E sur F.

2. Représentation locale d'une application différentiable

Sauf mention expresse, les espaces de Banach E et F de ce paragraphe sont de dimensions finies n et m sur le même corps K.

Cas où $Df(a)$ est surjective. 2.1. — Soient U un ouvert de E, f une application de classe C^k, $k \geqslant 1$, de U dans F, telle que Df soit surjective en $a \in U$. Alors f est localement C^k-conjuguée aux voisinages de a et $f(a)$ à la projection canonique

$$K^n = K^m \times K^{n-m} \to K^m$$
$$(x_1, \dots, x_m, \dots, x_n) \to (x_1, \dots, x_m)\,.$$

En d'autres termes, le comportement local de f au voisinage de a est celui de sa différentielle $Df(a)$.

PREUVE. — Puisque $Df(a) \in \mathscr{L}(E; F)$ est surjective, $n = \dim E \geqslant \dim F = m$. Soit $(y_1, \dots, y_m)$ un système de coordonnées locales en $b = f(a)$. Posons $x_r = y_r \circ f$ pour $r = 1, \dots, m$. On a donc $Dx_r(a) = Dy_r(b) \circ Df(a)$. Puisque $Df(a)$ est surjective et que $\{ Dy_r(b) \}$ forme une base de F^* (théorème 1.5.), les formes linéaires $Dx_1(a), \dots, Dx_m(a)$ sont linéairement indépendantes. Complétons-les par $n - m$ formes linéaires $x_{m+1}, \dots, x_n$ de façon à former une base de E^*. Alors, d'après 1.5., $(x_1, \dots, x_n)$ est un système de coordonnées locales en a.

Par construction, l'expression de f dans ces coordonnées est

$$(x_1(u), \dots, x_n(u)) \rightsquigarrow (x_1(u), \dots, x_m(u))\,. \qquad \square$$

Remarque. — Le théorème vaut encore en dimension infinie, pourvu qu'il existe un sous-espace fermé E_2 de E tel que E soit la somme directe $\operatorname{Ker} Df(a) \oplus E_2$. Il a été démontré sous cette hypothèse au (5.4. chap. 3), dans la perspective d'un théorème d'existence plutôt que dans celle, envisagée ici, d'un modèle local.

Cas où $Df(a)$ est injective. 2.2. — Soient U un ouvert de E, f une application de classe C^k, $k \geqslant 1$, de U dans F. Supposons que Df soit injective en $a \in U$. Alors f est localement C^k-conjuguée aux voisinages de a et de $f(a)$ à l'injection canonique

$$K^n \quad \to \quad K^m$$
$$(x_1, \dots, x_n) \rightsquigarrow (x_1, \dots, x_n, 0, \dots, 0)\,.$$

Autrement dit, le comportement local de f au voisinage de a est encore celui de sa différentielle $Df(a)$.

PREUVE. — Puisque $Df(a) \in \mathscr{L}(E; F)$ est injective $n = \dim E \leqslant \dim F = m$.

Soit $(x_1, \dots, x_n)$ un système de coordonnées locales en a. Si S est un supplémentaire de $Df(a)$ dans F, prenons un système de coordonnées linéaires $(x'_{n+1}, \dots, x'_m)$ sur S.

Alors $(x_1, \dots, x_n, x'_{n+1}, \dots, x'_m)$ est un système de coordonnées locales sur $U \times S$ au voisinage de $(a, 0)$.

Définissons $\varphi : U \times S \to F$ par $\varphi(u, s) = f(u) + s$. La différentielle de φ au point $(a, 0)$, évaluée sur $(h, k) \in E \times S$, est $D\varphi(a, 0).(h, k) = Df(a).h + k$; c'est donc un isomorphisme de $E \times S$ sur F.

Définissons des fonctions $y_1, \dots, y_m$ au voisinage de $b = f(a) \in F$ par $y_i = x_i \circ \varphi^{-1}$ pour $i = 1, \dots, n$ et $y_i = x'_i \circ \varphi^{-1}$ pour $i = n + 1, \dots, m$. Puisque $D\varphi(a, 0)$ est un isomorphisme et que les formes linéaires $Dx_1(a), \dots, Dx_n(a), Dx'_{n+1}(a), \dots, Dx'_m(a)$ sont linéairement indépendantes, il en résulte que les $Dy_i(b)$ forment une base de F^*. D'après 1.5., $(y_1, \dots, y_m)$ est donc un système de coordonnées locales sur F au voisinage de $b = f(a)$.

Par construction, on a $y_i \circ f(u) = x_i(u)$ pour $i = 1, \dots, n$ et $y_i \circ f(u) = y_i \circ \varphi(u, 0) = x'_i(0) = 0$ pour $i = n + 1, \dots, m$. L'expression de f dans les coordonnées (x_i), (y_j) est donc

$$(x_1(u), \dots, x_n(u)) \rightsquigarrow (x_1(u), \dots, x_n(u), 0, \dots, 0)\,. \qquad \square$$

GÉNÉRALISATION. — Le théorème s'étend aux espaces de Banach de dimension infinie, pourvu que le sous-espace $Df(a)\,E$ de F soit fermé et qu'il admette un supplémentaire fermé S.

La preuve est la même : $\varphi : U \times S \to F = Df(a)\,E \oplus S$, définie par $\varphi(u, s) = f(u) + (0, s)$, est un difféomorphisme local au voisinage de $(a, 0)$. Si g est le difféomorphisme local inverse, défini au voisinage de $(f(a), 0)$, alors $g \circ f$ applique un voisinage ouvert V, de a dans U, dans $Df(a)\, E \times \{ 0 \}$ et induit un difféomorphisme de ce voisinage sur un voisinage ouvert de 0 dans $Df(a)\, E$.

Rang de $Df(a)$. 2.3.

Rappelons qu'on appelle rang d'une application linéaire $L : E \to F$ la dimension de l'image de E.

Reprenons les notations de 1.2. Si f et f_1 sont C^k-conjuguées, $f_1 = \varphi' \circ f \circ \varphi^{-1}$. Par conséquent $Df_1(a_1) = D\varphi'(a') \circ Df(a) \circ D\varphi^{-1}(a_1)$. Comme $D\varphi'(a')$ et $D\varphi^{-1}(a_1)$ sont des isomorphismes, le rang de Df en a est égal à celui de Df_1 au point $a_1 = f(a)$. En d'autres termes, le rang est invariant par C^k-conjugaison.

Il en résulte que f ne peut être localement C^k-conjuguée, au voisinage de a, à une application linéaire f_1 [dont le rang est évidemment constant] que si le rang de Df est constant au voisinage de a.

C'est ainsi que $f : t \in \mathbf{R} \rightsquigarrow t^3 \in \mathbf{R}$ ne peut être C^k-conjuguée à une application linéaire au voisinage de zéro, car $f'(0) = 0$, tandis que $f'(t) = 3\,t^2 \neq 0$ si $t \neq 0$.

Il faut donc bien que le rang de Df soit constant au voisinage de a dans les théorèmes 2.1. et 2.2. Cherchons-en la raison.

Bien que l'application $u \rightsquigarrow$ rang $Df(u)$ ne soit évidemment pas continue, elle est semi-continue inférieurement : si rang $Df(a) = p$, alors rang $Df(u) \geqslant p$ dans un voisinage de a. En effet, un mineur d'ordre p de la matrice jacobienne de f est non nul en a. Comme il dépend continûment de u, il demeure non nul dans un voisinage de a.

Lorsque le rang de $Df(a)$ est égal à la dimension de E(injectivité) ou de F (surjectivité), il prend sa valeur maximum sup $\{ \dim E, \dim F \}$ au point a. D'après la semi-continuité, il reste alors constant au voisinage de a. C'est pourquoi la généralisation naturelle des théorèmes 2.1. et 2.2. est la suivante :

Théorème du rang constant. 2.3. — ***Pour qu'une application $f : U \subset E \to F$ de classe C^k, $k \geqslant 1$, soit localement C^k-conjuguée au voisinage de $a \in U$ à une application linéaire, il faut et il suffit que le rang de $Df(u)$ soit égal à une constante r sur un voisinage de a.***

PREUVE. — La nécessité est manifeste. Une application linéaire étant égale à sa différentielle, son rang est constant. Et nous venons de voir (2.2.) que le rang est invariant par C^k-conjugaison.

Démontrons la suffisance. On peut supposer $a = 0$, $f(a) = 0$. Choisissons sur E et F des coordonnées linéaires (x'_i) et (y'_j) telles que $Df(a)$ s'exprime par $(x'_1, \ldots, x'_n) \rightsquigarrow (x'_1, \ldots, x'_r, 0, \ldots, 0)$ et identifions E à $\mathbf{R}^n$ et F à $\mathbf{R}^m$ via ces coordonnées.

Si $f_1, \ldots, f_m$ sont les composantes de f, considérons l'application $\varphi : U \to \mathbf{R}^n$ définie par $\varphi = (f_1, \ldots, f_r, x'_{r+1}, \ldots, x'_n)$. L'expression de $Df(0)$ montre que $D\varphi(0)$ est l'identité. D'après le théorème d'inversion locale (1.5), il en résulte que $x_1 = f_1, \ldots, x_r = f_r, x_{r+1} = x'_{r+1}, \ldots, x_n = x'_n$ est un système de coordonnées locales au voisinage de $a = 0$. Exprimons f dans les coordonnées (x_i), (y'_j) ; on trouve $(x_1, \ldots, x_n) \rightsquigarrow (x_1, \ldots, x_r, F_{r+1}(x), \ldots, F_m(x))$, où F_j est une fonction de classe C^k de $x = (x_1, \ldots, x_r)$. La matrice jacobienne de cette application contient donc comme bloc $r \times r$ supérieur gauche la matrice unité $r \times r$. Comme son rang au voisinage de $a = 0$ est égal à r, tout bloc $(r + 1) \times (r + 1)$ qu'on en extrait est de déterminant nul. Il en résulte $D_i\, F_j(x_1, \ldots, x_n) = 0$ pour $i, j > r$. Les fonctions F_j ne dépendent donc pas de $x_{r+1}, \ldots, x_n$.

Posons $y_1 = y'_1, \ldots, y_r = y'_r, y_{r+1} = y'_{r+} - F_{r+1}(y_1, \ldots, y_r), \ldots, y_m = y'_m - F_m(y_1, \ldots, y_r)$. On obtient un nouveau système de coordonnées locales (y_i) au voisinage de $f(a) = 0$, car ces formules sont inversibles : $y'_i = y_i + F_i(y_1, \ldots, y_r)$ pour $i > r$. Exprimons f dans les coordonnées (x_i), (y_j) ; on trouve $(x_1, \ldots, x_n) \rightsquigarrow (x_1, \ldots, x_r, 0, \ldots, 0)$. C'est ce qu'il fallait démontrer.

□

GÉNÉRALISATION. — A titre d'exercice, le lecteur pourra démontrer le résultat général suivant qui ne suppose plus le rang constant :

Soit f une application de classe C^k, $k \geqslant 1$, d'un ouvert de $\mathbf{R}^n$ dans $\mathbf{R}^m$. Supposons qu'en un point a de cet ouvert le rang de $Df(a)$ soit égal à r. Alors f est localement C^k-conjugué au voisinage de a à

$$\begin{array}{ccccc} \mathbf{R}^n = \mathbf{R}^r \times \mathbf{R}^{n-r} & \to & \mathbf{R}^m = \mathbf{R}^r \times \mathbf{R}^{m-r}, \\ (u\quad ,\quad v) & \to & (u, \quad h(u, v)) \end{array}$$

où $h : \mathbf{R}^n \to \mathbf{R}^{m-r}$ est de classe C^k et $Dh(0) = 0$. [Voir : R. Abraham et J. Robbin. Transversal mappings and flows. Benjamin. New York. 1967. Pages 4 et 5.]

3. Le lemme de Morse-Palais

Le théorème 2.1. montre qu'une fonction f définie sur un ouvert de $\mathbf{R}^n$, à valeurs réelles, et de classe C^k, est localement C^k-conjuguée à une application affine au voisinage d'un point $a \in U$ où $Df(a) \neq 0$.

Nous allons examiner ce qu'il en est au voisinage d'un point où la différentielle s'annule.

Point critique non dégénéré. 3.1. — Soient U un ouvert d'un espace de Banach E, $f : U \to \mathbf{R}$ une fonction de classe C^k, $k \geqslant 1$. On dit que $a \in U$ est un *point critique* de f si $Df(a) = 0$.

Supposons que $k \geqslant 2$. La différentielle seconde $D^2 f(a)$, appelée le *hessien* de f en a, est une forme bilinéaire continue symétrique sur E (voir 1.5. chap. 4). Mais $\mathscr{L}(E, E; \mathbf{R})$ est canoniquement isomorphe à $\mathscr{L}(E; \mathscr{L}(E; \mathbf{R})) = \mathscr{L}(E; E^*)$ d'après (4.1. Appendice A). Cet isomorphisme fait correspondre à $D^2 f(a)$ l'application linéaire continue $h \in E \rightsquigarrow D^2 f(a).(h, .) \in E^*$. Si cette application est un isomorphisme de E sur son dual E^*, on dit que $D^2 f(a)$ est *non singulière* et que le point critique a est *non dégénéré.*

EXEMPLE. — Si $E = \mathbf{R}$, le point a est un point critique si $f'(a) = 0$. Il est non dégénéré si $f''(a) \neq 0$.

Notons que si φ est un C^k-difféomorphisme d'un voisinage d'un point critique a de f dans E, on a $D[f \circ \varphi](\varphi^{-1} a) = Df(a) \circ D\varphi(\varphi^{-1} a) = 0$. Par conséquent $\varphi^{-1}(a)$ est un point critique de $f \circ \varphi$. Autrement dit, *la C^k-conjugaison conserve les points critiques.*

Supposons que a soit un point critique non dégénéré de f. Examinons le point critique $\varphi^{-1}(a)$ de $f \circ \varphi$. On peut supposer, pour simplifier les notations, que $a = \varphi(a) = 0$. Si $h, l \in E$ on a (2.7. chap. 4) :

$$D^2(f \circ \varphi)(a).(h, l) = Df[\varphi(a)].D^2 \varphi(a).(h, l) + D^2 f[\varphi(a)].(D\varphi(a).h, D\varphi(a).l).$$

Puisque $a = 0$ est un point critique, cela se réduit à

$$D^2(f \circ \varphi)(0).(h, l) = D^2 f(0).(D\varphi(0).h, D\varphi(0).l).$$

Comme $D\varphi(0)$ est un isomorphisme, cette formule montre que $D^2(f \circ \varphi)(0)$ est non singulière. Le point critique $\varphi^{-1}(a) = 0$ est donc non dégénéré. *La C^k-conjugaison conserve donc la non-dégénérescence.*

Ceci posé, nous allons montrer qu'une fonction $f : \mathbf{R}^n \to \mathbf{R}$ de classe C^3 se comporte, au voisinage d'un point critique non dégénéré a, comme la partie $f(a) + \dfrac{1}{2} D^2 f(a).(h, h)$ de son développement de Taylor $f(a + h) = f(a) + \dfrac{1}{2} D^2 f(a)(h, h) + o(\| h \|^2)$. En fait, nous allons le démontrer dans le cadre plus général des espaces de Hilbert réels. Car la preuve qui suit, due à R. Palais, est notablement plus simple que celle donnée initialement par M. Morse, et qui ne s'appliquait qu'à $\mathbf{R}^n$.

Nous pouvons supposer, sans nuire à la généralité, que le point critique a est l'origine de E et que $f(a) = 0$.

Enoncé du lemme de Morse-Palais. 3.2. — Soit f une fonction à valeurs réelles, de classe C^{k+2}, $k \geqslant 1$, définie sur un voisinage U de l'origine d'un espace de Hilbert E.

Supposons que $f(0) = 0$ et que 0 soit un point critique non dégénéré. Alors, il existe un C^k-difféomorphisme local φ au voisinage de 0, tel que $\varphi(0) = 0$ et $f(x) = \frac{1}{2} D^2 f(0).(\varphi(x), \varphi(x))$

Commençons par établir un lemme.

Lemme. 3.3. — ***Soit E un espace de Banach. Si $B \in \mathrm{End}(E)$ est tel que $\| 1 - B \| < 1$, alors il existe une suite de polynômes en B qui converge dans l'espace de Banach $\mathrm{End}(E)$ vers un endomorphisme R tel que $R^2 = B$, et que nous noterons $\sqrt{B}$.***

De plus, il existe un voisinage U de $1 = \mathrm{id}_E$ tel que $B \in U \leadsto \sqrt{B}$ soit un difféomorphisme de classe C^∞ de U sur son image.

PREUVE. — Une éventuelle racine carrée R de B doit satisfaire $S = \frac{1}{2}[(1 - B) + S^2]$, si l'on pose $R = 1 - S$. Afin de prouver son existence, nous allons utiliser une méthode d'approximations successives. Pour cela, définissons une suite $A_n \in \mathrm{End}(E)$ par récurrence :

$$A_0 = 0, \qquad A_{n+1} = [(1 - B) + A_n^2]/2, \qquad n = 0, 1, \dots .$$

Posons $a = \| 1 - B \|$.

a) Montrons que $\| A_n \| < \sqrt{a}$.

C'est évident pour $n = 0$. Supposons-le établi pour n et prouvons-le pour $n + 1$.

$$\| A_{n+1} \| \leqslant [\| 1 - B \| + \| A_n \|^2]/2 \leqslant (a + a)/2 = a,$$

qui est inférieur à $\sqrt{a}$ puisque $0 < a < 1$.

b) A_n est un polynôme en $(1 - B)$, donc en B.

C'est évident, par récurrence sur n, à partir de la définition de A_n.

c) La suite A_n converge dans $\mathrm{End}(E)$.

On a $A_{n+1} - A_n = (A_n^2 - A_{n-1}^2)/2$ et des relations semblables en remplaçant n par $n + 1, \dots, n + p - 1$. Ajoutons-les membres à membres : $A_{n+p} - A_n = (A_{n+p-1}^2 - A_{n-1}^2)/2$.

D'après *b*, les A_i commutent deux à deux. On peut donc écrire

$$A_{n+p} - A_n = (A_{n+p-1} - A_{n-1}).(A_{n+p-1} + A_{n-1})/2 .$$

Puisque $\| A_i \| < \sqrt{a}$, on en déduit $\| A_{n+p} - A_n \| < \sqrt{a}.\| A_{n+p-1} - A_{n-1} \|$. De proche en proche on obtient $\| A_{n+p} - A_n \| \leqslant (\sqrt{a})^n . \| A_p - A_0 \| < (\sqrt{a})^{n+1}$. Puisque $0 < \sqrt{a} < 1$, il en résulte que A_n est une suite de Cauchy, qui converge donc vers un élément S de l'espace complet $\mathrm{End}(E)$.

d) $(1 - S)^2 = B$.

Il suffit de faire $n \to +\infty$ dans la relation qui définit A_n par récurrence.

Notons que, puisque $\| A_n \| < \sqrt{a}$, on a $\| S \| < \sqrt{a}$. Autrement dit $R = 1 - S$ vérifie $\| 1 - R \| < \sqrt{a}$. Ainsi $R = \sqrt{B}$ reste dans la boule $B(1, \sqrt{r})$ si B est dans la boule $B(1, r)$ de centre 1 et de rayon $r < 1$ de $\mathrm{End}(E)$.

e) La différentielle de l'application $Q : \mathrm{End}(E) \to \mathrm{End}(E)$ de classe C^∞, définie par $Q(A) = A^2$, est $DQ(A).h = A.h + h.A$ pour $h \in \mathrm{End}(E)$. Il en résulte que $DQ(1)$ est inversible. D'après le théorème des fonctions inverses et (2.9. chap. 5), Q est donc un difféomorphisme de classe C^∞ d'un voisinage U de $1 = \mathrm{id}_E$ sur son image. Prenons $r < 1$ assez petit pour que $B(1, \sqrt{r}) \subset U$. Alors, d'après *d*, Q^{-1} restreint à $B(1, r)$ n'est autre que l'application $B \leadsto R = \sqrt{B}$ que nous avons construite. Cela montre l'unicité de la racine carrée et prouve que $B \leadsto \sqrt{B}$ est C^∞. □

PREUVE DU LEMME DE MORSE-PALAIS. 3.4. — On peut supposer que U est une boule ouverte de centre O. Cela permet d'appliquer le théorème fondamental du calcul intégral (6.6. chap. 2), d'abord à f, ensuite à Df :

$$f(x) = f(x) - f(0) = \int_0^1 Df(tx).x.\mathrm{d}t .$$

$$Df(tx) = Df(tx) - Df(0) = \int_0^1 D^2 f(stx).tx.\mathrm{d}s .$$

On en déduit

$$f(x) = \int_0^1 \int_0^1 D^2 f(s.t.x).(x, x).t.\mathrm{d}s.\mathrm{d}t .$$

D'après la propriété (d.6.4. chap. 2) on peut donc écrire $f(x) = g(x).(x, x)$, où

(3.5) $$g(x) = \int_0^1 \int_0^1 D^2 f(s.t.x).t\,\mathrm{d}s.\mathrm{d}t .$$

Ainsi g est une application de classe C^k de U dans l'espace de Banach $\mathscr{L}(E, E; \mathbf{R})$, qu'on sait être canoniquement isomorphe à $\mathscr{L}(E; \mathscr{L}(E; \mathbf{R})) = \mathscr{L}(E; E^*)$ [Appendice A]. D'autre part, le produit scalaire $\langle\ ,\ \rangle$ de E induit un isomorphisme canonique $h \in E \rightsquigarrow \langle h, .\rangle \in E^*$. En conséquence, $\mathscr{L}(E, E; \mathbf{R})$ est canoniquement isomorphe à $\mathscr{L}(E; E) = \mathrm{End}\,(E)$. L'image $A(x)$ de $g(x)$ dans cet isomorphisme est donnée par

(3.6) $$g(x).(h, l) = \langle A(x).h, l \rangle \quad \text{pour } h, l \in E .$$

En particulier

(3.7) $$\frac{1}{2} D^2 f(0).(x, x) = g(0).(x, x) = \langle A(0).x, x \rangle ,$$

qui montre que $A(0)$ est un isomorphisme, puisque 0 est un point critique non dégénéré. Ainsi

(3.8) $$f(x) = \langle A(x).x, x \rangle \quad \text{pour } x \in U .$$

Nous cherchons un difféomorphisme local φ tel que $\varphi(0) = 0$ et $f(x) = \frac{1}{2} Df(0).(\varphi(x), \varphi(x))$, c'est-à-dire, d'après (3.7) et (3.8),

(3.9) $$\langle A(x).x, x \rangle = \langle A(0).\varphi(x), \varphi(x) \rangle .$$

Pour l'exhiber, posons $B(x) = A(0)^{-1}.A(x)$. Puisque g est continu, A l'est aussi, d'après (3.6). Quitte à réduire U on peut donc supposer que $A(0)^{-1}.A(x)$ est voisin de $A(0)^{-1}.A(0) = 1$; c'est-à-dire que $\| 1 - B(x) \| \leqslant 1$ pour $x \in U$.

D'après le lemme précédent, $B(x)$ possède une racine carrée $R(x)$. Nous allons montrer que $\varphi(x) = R(x).x$ répond à la question.

a) φ est de classe C^k.

En effet, g étant de classe C^k, A et $B = A(0)^{-1}.A$ le sont aussi, d'après (3.6). Comme l'application « racine carrée » est C^∞, l'application composée $x \rightsquigarrow B(x) \rightsquigarrow R(x) = \sqrt{B(x)}$ est de classe C^k.

b) φ est un difféomorphisme local et $\varphi(0) = 0$.

D'après la règle de Leibniz on a $D\varphi(x).h = DR(x).(h, x) + R(x).h$ pour $h \in E$. En sorte que $D\varphi(0) = R(0) = \sqrt{B(0)} = 1$ est un isomorphisme. Il suffit alors d'appliquer le théorème d'inversion locale.

Bien sûr $\varphi(0) = R(0).0 = 0$.

c) Il reste à montrer que φ vérifie (3.9).

D'après le théorème de Schwarz (1.5. chap. 4), $D^2 f(\ )$ est symétrique. D'après (3.5), il en est donc de même de $g(x)$. Il résulte alors de (3.6) que $A(x)$ est égal à son adjoint $A(x)^*$. On en déduit $B(x)^* = [A(0)^{-1}.A(x)]^* = A(x).A(0)^{-1}$; ou encore $B(x)^*.A(0) = A(x) = A(0).B(x)$.

On peut remplacer, dans cette dernière relation, $B(x)$ par un polynôme quelconque en $B(x)$. Comme $R(x) = \sqrt{B(x)}$ est la limite d'une suite de tels polynômes (lemme précédent), on a encore $R(x)^*.A(0) = A(0).R(x)$.

Finalement $\langle A(0).\varphi(x), \varphi(x) \rangle = \langle A(0).R(x).x, R(x).x \rangle = \langle R(x)^*.A(0).R(x).x, x \rangle = \langle A(0).R^2(x).x, x \rangle = \langle A(0).B(x).x, x \rangle = \langle A(x).x, x \rangle$. □

Remarque. — Le lecteur trouvera dans (R. Abraham et J. Marsden, pages 175-176) une preuve encore plus courte, mais qui nécessite la mise en œuvre de notions étrangères à ce livre.

Corollaire. 3.10. — ***Les points critiques non dégénérés de f sont isolés.***

Corollaire. 3.11. — ***Soit $f : \mathbf{R}^n \to \mathbf{R}$ une fonction de classe C^{k+2}, $k \geqslant 1$. Supposons que $f(0) = 0$ et que 0 soit un point critique non dégénéré de f. Alors, il existe sur un voisinage de 0 des coordonnées locales (x_i) où f s'exprime par $(x_1)^2 + \cdots + (x_r)^2 - (x_{r+1})^2 - \cdots - (x_n)^2$.***

En particulier, la classe de C^k-conjugaison locale en 0 de f est caractérisée par le nombre de carrés affectés du signe $+$ qui interviennent dans la décomposition en somme de carrés de la forme quadratique $D^2 f(0).(h, h)$.

Note. — Le lecteur trouvera d'autres applications du lemme de Morse-Palais dans le chapitre « Calcul des variations ».

4. Linéarisation des champs de vecteurs

Conjugaison dans le groupe des difféomorphismes. 4.1. — Revenons à la C^k-conjugaison définie en 1.1. et aux notations que nous avions utilisées : $f : E \to E'$ est C^k-conjugué à $f_1 : E_1 \to E'_1$ s'il existe des C^k-difféomorphismes $\varphi : E \to E_1$ et $\varphi' : E' \to E'_1$ tels que $f_1 = \varphi' \circ f \circ \varphi^{-1}$.

Cela implique que tous les C^k-difféomorphismes $f : E \to E$ sont C^k-conjugués à l'application identique :

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[d, "\mathrm{id}_E"'] & E \arrow[d, "f^{-1}"] \\ E \arrow[r, "\mathrm{id}_E"] & E \end{tikzcd}$$

La C^k-conjugaison est donc une classification très grossière lorsque la source E et le but E' coïncident. Nous allons la raffiner.

Nous dirons que deux applications $f : E \to E$ et $f_1 : E \to E$ de classe C^k sont conjuguées dans le groupe $\mathrm{Diff}^k(E)$ des C^k-difféomorphismes de E s'il existe $\varphi \in \mathrm{Diff}^k(E)$ tel que $f_1 = \varphi \circ f \circ \varphi^{-1}$ [on dit, le plus souvent, « conjuguées » si le contexte évite les ambiguïtés].

De nouveaux invariants apparaissent. Si a est un point fixe de f, alors $\varphi(a)$ est un point fixe de f_1. Plus généralement, l'image par φ de chaque orbite $\{ f^n(a) : n \in \mathbf{N} \}$ de f est une orbite de f_1. Le nombre d'orbites de période P (c'est-à-dire comptant P éléments distincts) est donc un invariant.

Le lecteur définira sans peine la notion locale correspondante en s'inspirant de 1.2.

Conjugué d'un sous-groupe à un paramètre de difféomorphismes. 4.2.

Supposons que $F : \mathbf{R} \times E \to E$ soit un sous-groupe à un paramètre $t \in \mathbf{R}$ de difféomorphismes de E (voir 5. chap. 7). On vérifie immédiatement que son transformé $\varphi \circ F_t \circ \varphi^{-1}$ par $\varphi \in \mathrm{Diff}^k(E)$ en est également un.

Cherchons à exprimer son générateur, noté $\varphi * X$, à partir du générateur X de F. Le théorème de différentiation des applications composées entraîne

$$\frac{\mathrm{d}}{\mathrm{d}t}[\varphi \circ F_t \circ \varphi^{-1}(x)]\Big|_{t=0} = D\varphi(\varphi^{-1} x) \cdot \frac{\mathrm{d}}{\mathrm{d}t} F(t, \varphi^{-1} x)\Big|_{t=0}$$

Par conséquent, $(\varphi * X)(x) = D\varphi(\varphi^{-1} x).X(\varphi^{-1} x)$. Le champ de vecteurs $\varphi * X$ s'appelle l'*image de X par* φ.

Notons que, puisque $D\varphi$ est de classe C^{k-1}, $\varphi * X$ est de classe C^{k-1}. C'est un exemple d'un champ de vecteurs de classe C^{k-1} qui engendre pourtant un flot $\varphi \circ F_t \circ \varphi^{-1}$ de classe C^k. Ce qui, bien entendu, ne contredit pas le théorème (3.10. chap. 7).

Etant donné un sous-groupe à un paramètre de difféomorphismes F_t (ou son générateur X), cherchons s'il existe un difféomorphisme φ tel que $\varphi \circ F_t \circ \varphi^{-1}$ (ou $\varphi * X$) présente une forme plus simple. En particulier, peut-on choisir φ de façon que $\varphi \circ F_t \circ \varphi^{-1}$ soit un sous-groupe à un paramètre de translations de E ? c'est sûrement impossible en toute généralité : une translation n'a pas de point fixe, alors que F_t peut en posséder.

Si a est un point fixe de F_t, c'est-à-dire si $F(t, a) = a$ pour tout t, alors $X(a) = 0$ (et réciproquement). On dit que a est un *point critique* de X. Un point de E où X ne s'annule pas s'appelle un *point régulier* de X ; en conformité avec la terminologie introduite en 3.1.

L'existence de point critique de X est, nous venons de le voir, une obstruction à la transformation de F_t en groupe de translations. Nous allons voir que, *localement*, c'est la seule.

Théorème du redressement. 4.3. — *Soit a un point régulier d'un champ de vecteurs X de classe C^k, $k \geqslant 1$, défini sur un ouvert Ω d'un espace de Banach E. Alors, il existe un voisinage ouvert U de a et un C^k-difféomorphisme $\varphi : U \to \varphi(U)$ tels que $\varphi * X$ soit constant sur $\varphi(U)$.*

En particulier, si $E = \mathbf{R}^n$, il existe un système de coordonnées locales $(x_1, \ldots, x_n)$ en a dans lequel les composantes du champ X sont $X_1 = 1$, $X_2 = \cdots = X_n = 0$. Dans ce système de coordonnées, $\varphi \circ F_t \circ \varphi^{-1}$ s'écrit $(x_1, \ldots, x_n) \rightsquigarrow (x_1 + t, x_2, \ldots, x_n)$: c'est bien un groupe (local) de translations.

Preuve. — Sans perdre en généralité, on peut supposer que a est l'origine O de E. Soit E' un sous-espace fermé supplémentaire de l'espace engendré par $X(0) : E = \mathbf{R}.X(0) \oplus E'$. Alors $\Omega \cap E'$ est un voisinage ouvert de O dans E'.

Soit $F(t, x)$ le flot local défini par X [voir 5.5. chap. 7]. Puisque F est continue, on peut choisir un voisinage ouvert U de O dans $\Omega \cap E'$ et un intervalle $]-2r, 2r[$ tels que $F(t, u) \in \Omega$ si $-2r < t < 2r, u \in U$. Posons $I =]-r, r[$ et soit G la restriction de F à $I \times U$. Puisque X est de classe C^k, il en est de même de F et de G (3.10. chap. 7). Cherchons la différentielle de G en $(0, 0)$. Si $(h, k) \in \mathbf{R} \times E'$ on a $DG(0, 0).(h, k) = D_1 G(0, 0).h + D_2 G(0, 0).k$. Mais

$$D_1 G(s, u) = \frac{\mathrm{d}}{\mathrm{d}t} F(t, u)\Big|_{t=s} = X[F(s, u)];$$

donc $D_1 G(0, 0).h = h.X(0)$. D'autre part $F(0, u) = u$ entraîne $D_2 G(0, 0).k = k$. Par conséquent $DG(0, 0) : (h, k) \rightsquigarrow h.X(0) + k$ est un isomorphisme de $\mathbf{R} \oplus E'$ sur E. D'après le théorème d'inversion locale, et quitte à réduire r et U, G est donc un C^k-difféomorphisme de $I \times U$ sur $G(I \times U)$.

Posons $\varphi = G^{-1}$ et cherchons $\varphi \circ F_t \circ \varphi^{-1} = G^{-1} \circ G$ pour $-r < t < r$ (de façon à ne pas sortir de Ω). Puisque $F_t[F_s(u)] = F_{t+s}(u)$, on obtient

$$\varphi \circ F_t \circ \varphi^{-1} : (s, u) \in I \times U \rightsquigarrow (s + t, u) \in]-2r, 2r[\times U .$$

L'image de F_t par φ est donc une translation $(s, u) \rightsquigarrow (s + t, u)$ de $\mathbf{R} \oplus E' \sim E$. Son générateur $\varphi * X$ est le champ constant $\varphi * X(0) = (1, 0)$. □

Remarques. — *a*) On appelle ce théorème le théorème de redressement, ou de linéarisation, parce que les orbites de $\varphi \circ F_t \circ \varphi^{-1}$ sont des droites parallèles.

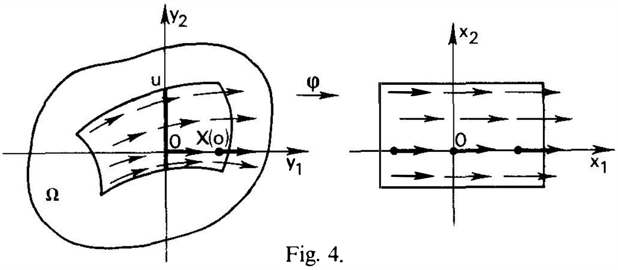


Fig. 4.

b) Le théorème de dépendance C^k par rapport à x de $F_t(x)$ se « lit » immédiatement sur son expression $\varphi \circ F_t \circ \varphi^{-1}$ dans la carte $(I \times U, \varphi)$.

c) Une autre preuve, fort ingénieuse, figure dans (E. Nelson).

Corollaire. 4.4. — ***Un champ de vecteurs X différentiable sur $\mathbf{R}^n$ possède $n - 1$ intégrales premières, dont les différentielles sont linéairement indépendantes en chaque point d'un voisinage d'un point régulier de X.***

PREUVE. — Soit O ce point régulier. Il existe au voisinage de O des coordonnées locales $(x_1, ..., x_n)$ où le champ $\varphi * X$ a pour composantes $(1, 0, ..., 0)$. Il en résulte que les fonctions coordonnées $x_2, ..., x_n$ sont des intégrales premières de $\varphi * X$ (voir 4. 5. chap. 7). Evidemment les différentielles de $x_2, ..., x_n$ sont linéairement indépendantes en chaque point.

Il en résulte que $x_i \circ \varphi$, $i = 2, ..., n$, constituent $n - 1$ intégrales premières de X, dont les différentielles sont linéairement indépendantes en chaque point. En effet, d'après le critère (4.5. chap. 7) et 4.2., si $x = \varphi(y)$, on a

$$D(x_i \circ \varphi)(y).X(y) = Dx_i(\varphi(y)).D\varphi(y).X(y) =$$
$$= Dx_i(x).D\varphi(\varphi^{-1} x).X(\varphi^{-1} x) = Dx_i(x).(\varphi * X)(x) = 0. \qquad \square$$

Remarque. — Il faudrait se garder d'en déduire qu'un champ de vecteurs, défini sur un ouvert U quelconque de $\mathbf{R}^n$ où il n'a pas de point critique, possède $n - 1$ intégrales premières définies sur U tout entier et dont les différentielles en chaque point sont linéairement indépendantes.

Voici un contre-exemple si $n = 4$:

Considérons dans $\mathbf{R}^4 = \{ m = (q_1, q_2, p_1, p_2) \}$ le champ de vecteurs X de composantes $(p_1, p_2 - q_1, -2 q_2)$. Il ne possède pas de point critique sur l'ouvert $U = \mathbf{R}^4 - \{ 0 \}$. Il est facile de voir (6.1. chap. 5) que $f(m) = (p_1)^2 + (q_1)^2$ et $g(m) = (p_2)^2 + 2(q_2)^2$ sont des intégrales premières dont les différentielles $Df(m)$ et $Dg(m)$ sont linéairement indépendantes. Par contre, il n'existe pas de troisième (ici $n - 1 = 3$) intégrale première h telle que $Df(m)$, $Dg(m)$ et $Dh(m)$ soient linéairement indépendantes en tout point de U. La raison en est que la courbe intégrale de X, issue de chaque point a de U, est partout dense sur le sous-ensemble $f^{-1}(a) \cap g^{-1}(a)$.

Le lecteur le prouvera en s'appuyant sur le théorème de Jacobi-Kronecker : chaque orbite de la rotation $z \rightsquigarrow e^{2\pi i r}.z$ du cercle $\{ z \in \mathbf{C} : | z | = 1 \}$ est partout dense sur le cercle si $r \in \mathbf{R}$ est irrationnel (dans le cas qui nous occupe $r = \sqrt{2}$).

En s'inspirant de ce qui précède, le lecteur pourra construire un contre-exemple sur $\mathbf{R}^3$: qu'il songe à une bobine torique sur laquelle s'enroule un fil figurant une orbite.

Par contre, on peut démontrer (W. Kaplan, 1940) que les courbes intégrales d'un champ de vecteurs différentiable et sans point critique de $\mathbf{R}^2$ sont les lignes de niveau $f^{-1}(r)$ d'une fonction $f : \mathbf{R}^2 \to \mathbf{R}$. Cette fonction est donc une intégrale première *globale*.

Notes. — Le problème de la classification par conjugaison dans le groupe des difféomorphismes motive quantité de recherches contemporaines. Le lecteur trouvera dans (V. Arnold) la classification des points critiques des champs de vecteurs linéaires : $X : x \rightsquigarrow A(x)$, où $A \in \mathrm{End}(E)$. Il trouvera dans (E. Nelson) celle des points critiques des champs de vecteurs quelconques. Enfin (R. Abraham et J. Marsden) et (M. Hirsch et S. Smale) fourmillent d'informations sur ce sujet.

9 SOUS-VARIÉTÉS DIFFÉRENTIABLES

Ce chapitre introduit les notions de sous-variété différentiable, d'espace tangent en un point d'une sous-variété et d'application différentiable d'une sous-variété dans une autre.

1. Sous-variétés différentiables

Considérons, dans l'esprit du chapitre précédent, que deux sous-ensembles d'un espace de Banach E sont équivalents s'ils se déduisent l'un de l'autre par un difféomorphisme $\varphi : E \to E$. A ce titre, les sous-ensembles les plus simples sont ceux qui sont équivalents à un sous-espace vectoriel, ou à un ouvert d'un tel sous-espace. Un morceau de courbe ou de surface de l'espace ordinaire est généralement de ce type. Par contre, un cercle ou une sphère ne sont que localement homéomorphes à une droite ou un plan. Si nous souhaitons les inclure dans notre étude, il faut « localiser » la notion d'équivalence par difféomorphisme.

Première définition des sous-variétés. 1.1. — Soit F un sous-espace de Banach d'un espace de Banach E. Une partie M de E est appelée *sous-variété de classe C^k de E modelée sur F*, si tout point $m \in M$ possède un voisinage ouvert U tel qu'il existe un C^k-difféomorphisme φ de U sur $\varphi(U)$ vérifiant $\varphi(U \cap M) = \varphi(U) \cap F$.

On peut dire encore de façon équivalente (voir 1. chap. 8), qu'au voisinage de chaque point $m \in M$, on peut trouver une carte (U, φ) telle que $\varphi(U \cap M) = \varphi(U) \cap F$.

Cette définition générale étant donnée, nous nous limiterons dans la suite au cas où E est un espace vectoriel réel de dimension n. Si le sous-espace F est de dimension d, on dira que M est une sous-variété de *E de dimension d* (ou de *codimension $n - d$*).

Nous omettrons le plus souvent le qualificatif « C^k » ; étant entendu que $k \geqslant 1$.

EXEMPLE. — Le cercle unité S^1 est une sous-variété C^∞ de dimension 1 de $\mathbf{R}^2 = \{ (x, y) \}$.

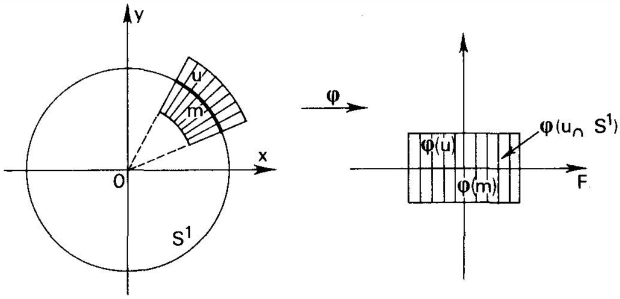


Fig. 5.

Ici, $\varphi : (x, y) \rightsquigarrow \left(\operatorname{Arctg} \dfrac{y}{x}, [x^2 + y^2]^{1/2} - 1 \right)$.

EXEMPLE. — Tout sous-espace affine de dimension d est une sous-variété de dimension d.

Conséquences. 1.2. — *a*) Si $d = n$, alors $F = E$ et l'on a $\varphi(U \cap M) = \varphi(U) \cap E = \varphi(U)$. Donc $U \cap M = U$ et $U \subset M$. Les sous-variétés de dimension maximale sont donc des ouverts de E (et réciproquement).

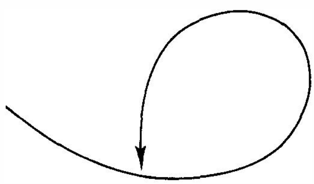
b) Si $d = 0$, alors $F = \{0\}$ et l'on a $\varphi(U \cap M) = \varphi(U) \cap F = \{0\}$. Donc $U \cap M$ se réduit au point $\varphi^{-1}(0)$. Les sous-variétés de dimension nulle sont donc des parties discrètes de E (et réciproquement).

c) L'image d'une sous-variété par un difféomorphisme $f : E \to E$ est une sous-variété de même dimension.

d) Puisque φ induit un homéomorphisme de $U \cap M$ sur un ouvert $\varphi(U) \cap F$ de F, une sous-variété de dimension d possède, localement, les mêmes propriétés topologiques que $\mathbf{R}^d$.

EXEMPLES. — Otons un point m à une sous-variété M de dimension 1. Alors un voisinage de $M - \{m\}$ est formé de deux parties connexes disjointes. C'est pourquoi la partie de $\mathbf{R}^2 = \{(x, y)\}$ d'équation $x . y = 0$ n'est pas une sous-variété : privée de l'origine, elle est formée de quatre composantes connexes.

Même raisonnement pour l'image A de la courbe $f : \mathbf{R} \to \mathbf{R}^2$ dont voici le dessin,



$f(0)$

Fig. 6.

et où la flèche indique que $f(t)$ tend vers $f(0)$ lorsque $t \to +\infty$. Au voisinage de $f(0)$, $A - \{f(0)\}$ est formée de trois composantes connexes.

Seconde définition des sous-variétés. 1.3. — Pour qu'une partie M d'un espace vectoriel E de dimension n soit une sous-variété de dimension d, il faut et il suffit, que tout point $m \in M$ possède un voisinage ouvert U sur lequel soient définies $n - d$ fonctions $f_i : U \to \mathbf{R}$ de classe C^k telles que

a) $U \cap M$ soit l'ensemble des zéros communs à ces fonctions ;

b) les formes linéaires $D_i f(u)$, $i = 1, \ldots, n - d$, soient linéairement indépendantes pour $u \in U$.

PREUVE. — *La condition est nécessaire.* — Identifions E à $\mathbf{R}^n = \mathbf{R}^d \times \mathbf{R}^{n-d}$ et F à $\mathbf{R}^d \times \{0\}$. Si (U, φ) est la carte de la définition 1.1. et si $(x_1, \ldots, x_n)$ sont les coordonnées canoniques de $E = \mathbf{R}^n$, les fonctions $f_i = x_i \circ \varphi$, $i = d + 1, \ldots, n$, vérifient trivialement a et b.

La condition est suffisante. — Soit $m \in M$. Complétons les formes linéaires $Df_i(m)$, $i = d+1, \ldots, n$, qui sont linéairement indépendantes, par des formes linéaires $f_1, \ldots, f_d$, de façon à former une base de E^*. D'après (1.5. chap. 8), $(f_1, \ldots, f_n)$ est un système de coordonnées locales sur un voisinage $V \subset U$ de m. L'application $\varphi : V \to \mathbf{R}^n = E$ de composantes $f_1, \ldots, f_n$ est un difféomorphisme de V sur $\varphi(V)$ vérifiant $\varphi(V \cap M) = \varphi(V) \cap (\mathbf{R}^d \times \{0\})$. □

Mise en garde. 1.4. — Supposons qu'une partie M de E soit l'ensemble des zéros communs à $n - d$ fonctions g_i. Il peut se faire que les différentielles des g_i ne soient pas linéairement indépendantes sur M et que, pourtant, M soit une sous-variété. Cela signifie simplement que les g_i sont mal choisies.

Exemple. — Soit M l'ensemble des zéros communs aux fonctions $g_1 : (x, y, z) \in \mathbf{R}^3 \rightsquigarrow x^2$ et $g_2 : (x, y, z) \in \mathbf{R}^3 \rightsquigarrow y^2$. Les différentielles Dg_1 et Dg_2 s'annulent sur M. Pourtant M est une sous-variété. Il fallait choisir $f_1(x, y, z) = x$ et $f_2(x, y, z) = y$.

Exemple. — Nous avons vu que l'ensemble A des zéros de $g_1 : (x, y) \in \mathbf{R}^2 \rightsquigarrow xy$ n'est pas une sous-variété. Pourtant, le fait que Dg_1 s'annule sur A ne suffit pas à le prouver. Rien ne dit, a priori, qu'une autre fonction f_1 n'aurait pas fait l'affaire.

Rappelons (voir 5. chap. 2) que si g est une application différentiable d'un ouvert U de l'espace de Banach E dans un espace de Banach E', un point $a \in U$ est dit point *critique* (resp. régulier) de g si $Dg(a)$ n'est pas (resp. est) surjective. Un point $b \in E'$ sera dit une *valeur régulière* de g si $g^{-1}(b)$ ne contient pas de point critique.

Corollaire. 1.5. — ***Soit g une application différentiable d'un ouvert U de E dans E'. Supposons que $b \in E'$ soit une valeur régulière et que $g^{-1}(b)$ ne soit pas vide. Alors $g^{-1}(b)$ est une sous-variété de E de codimension* dim *E'.***

Preuve. — Identifions E' à $\mathbf{R}^p$ par un choix de coordonnées linéaires. Soient $(g_1, ..., g_p)$ les composantes de g et $(b_1, ..., b_p)$ celles de b. Posons $f_i = g_i - b_i$. Alors $g^{-1}(b) = \{ x \in U : f_i(x) = 0, i = 1, ..., p \}$ et, puisque $Dg(x)$ est surjective, les $Df_i(x) = Dg_i(x)$ sont linéairement indépendantes pour $x \in U$. Il suffit maintenant d'appliquer 1.3. □

Remarque (troisième définition des sous-variétés). 1.6. — En fait, on a prouvé :

Pour que M soit une sous-variété de dimension d de $E = \mathbf{R}^n$, il faut et il suffit que chaque point $m \in M$ possède un voisinage ouvert U tel qu'il existe une application $g : U \to F = \mathbf{R}^{n-d}$ vérifiant

a) $Dg(u)$ est surjective pour $u \in U$;

b) $M \cap U = g^{-1}(b)$ pour $b \in F$.

La suffisance vient d'être montrée. La nécessité est contenue dans l'énoncé 1.3.

Exemple (sphères). 1.7. — La sphère S^n de $\mathbf{R}^{n+1} = \{ (x_1, ..., x_{n+1}) \}$ est une sous-variété de dimension n.

En effet $S^n = f^{-1}(1)$, où $f : (x_1, ..., x_{n+1}) \in \mathbf{R}^{n+1} \rightsquigarrow x_1^2 + \cdots + x_{n+1}^2 \in \mathbf{R}$. D'autre part $+1$ est valeur régulière de f, car $Df(x) = (2x_1, ..., 2x_{n+1})$ ne peut s'annuler sur S^n.

Exemple (groupe unimodulaire). 1.8. — Identifions $A \in \text{End}(\mathbf{R}^n)$ à sa matrice $n \times n$ dans la base canonique. Puis identifions cette matrice à un élément de $\mathbf{R}^{n^2}$, en lisant ses éléments de gauche à droite et de haut en bas (le lecteur arabisant fera la correction qui s'impose).

Le groupe GL (n) des automorphismes de $\mathbf{R}^n$ est un ouvert de $\mathbf{R}^{n^2}$ car c'est l'image inverse de l'ouvert $\mathbf{R} - \{ 0 \}$ par l'application continue $f : A \in \text{End}(\mathbf{R}^n) \rightsquigarrow$ déterminant (A).

Nous allons montrer que le groupe unimodulaire SL $(n) = f^{-1}(1)$, formé par les endomorphismes de déterminant $+1$, est une sous-variété de $\mathbf{R}^{n^2}$. L'application f est C^∞, car $f(A)$ est un polynôme en les coefficients de A. C'est un polynôme homogène de degré n, donc $f(t.A) = t^n.f(A)$ pour $t \in \mathbf{R}$. Il en résulte $Df(t.A).A = \frac{\mathrm{d}}{\mathrm{d}t} f(tA) = n.t^{n-1}.f(A)$. Faisons $t = 1 : Df(A).A = n.f(A)$. Par conséquent $Df(A) \neq 0$ sur SL $(n) = f^{-1}(1)$. Ainsi $+1$ est valeur régulière de $f : U \to \mathbf{R}$ et il suffit d'appliquer 1.5.

Exemple (groupe orthogonal). 1.9. — Une matrice $n \times n$ symétrique possède $n(n+1)/2$ éléments distincts. Cela permet d'identifier l'ensemble S des endomorphismes symétriques s (c'est-à-dire égaux à leur transposé ${}^t s$) à $\mathbf{R}^{n(n+1)/2}$, comme nous avons identifié End $(\mathbf{R}^n)$ à $\mathbf{R}^{n^2}$.

Considérons l'application f de l'ouvert $U = \text{GL}(n)$ de $\mathbf{R}^{n^2}$ dans $S = \mathbf{R}^{n(n+1)/2}$ définie par $f(A) = {}^tA.A$. C'est évidemment une application C^∞. D'après la règle de Leibniz, $Df(A).h = {}^th.A + {}^tA.h$ pour $h \in \text{End}(E)$. Montrons que $Df(A)$ est surjective. En effet, si $s \in S$, on a bien $Df(A).h = s$ en prenant $h = {}^t(A^{-1}).s/2$. Il en résulte que $\text{id}_{\mathbf{R}^n} \in S$ est une valeur régulière de f.

Comme le groupe orthogonal $O(n)$ de $\mathbf{R}^n$ n'est autre que $f^{-1}(\mathrm{id}_{\mathbf{R}^n})$, 1.5. montre que $O(n)$ est une sous-variété de $\mathbf{R}^{n^2}$ de dimension $n^2 - n(n+1)/2 = n(n-1)/2$.

La preuve de la généralisation suivante de 1.5. est laissée à la discrétion du lecteur :

Théorème. 1.10. — ***Soit f une application de classe C^k d'un ouvert U de E dans F. Supposons que $Df(u)$ soit de rang constant r pour $u \in U$. Alors l'image inverse $f^{-1}(b)$ (supposée non vide) d'un point $b \in F$ est une sous-variété de classe C^k de E, de dimension*** $\dim(E) - r$.

Graphe d'application. 1.11. — Soient E_1 et E_2 deux espaces vectoriels de dimension finie, U_1 un ouvert de E_1 et $g : U_1 \to E_2$ une application de classe C^k. Alors le graphe $\{ (u_1, g(u_1)) : u_1 \in U_1 \}$ de g est une sous-variété de $E = E_1 \oplus E_2$ de dimension égale à celle de E_1.

Preuve. — Identifions E_2 à $\mathbf{R}^r$ via des coordonnées linéaires, et soient $(g_1, \ldots, g_r)$ les composantes de g. Le graphe de g est l'ensemble des points $(u, x_1, \ldots, x_r)$ de l'ouvert $U_1 \times \mathbf{R}^r$ de E tels que $x_i - g_i(u) = 0$, $i = 1, \ldots, r$. Si l'on pose $f_i(u, x_1, \ldots, x_r) = x_i - g_i(u)$, c'est encore l'ensemble des zéros communs aux fonctions f_i. Comme les Df_i sont linéairement indépendantes en chaque point, 1.3. entraîne le résultat.

Nous allons montrer que chaque point m d'une sous-variété M possède un voisinage ouvert U tel que $U \cap M$ soit le graphe d'une application g comme ci-dessus. Nous aurons obtenu ainsi une quatrième caractérisation des sous-variétés de dimension d : *localement*, ce sont des graphes d'application d'un ouvert de $\mathbf{R}^d$ dans un $\mathbf{R}^r$. □

Proposition. 1.12. — ***Soit M une sous-variété de dimension d de*** $\mathbf{R}^n = \{ (x_1, \ldots, x_n) \}$. ***Quitte à permuter $x_1, \ldots, x_n$, tout point $m \in M$ possède un voisinage U tel que $U \cap M$ soit le graphe d'une application g d'un ouvert de*** $\mathbf{R}^d = \{ (x_1, \ldots, x_d, 0, \ldots, 0) \}$ ***dans***

$$\mathbf{R}^{n-d} = \{ (0, \ldots, 0, x_{d+1}, \ldots, x_n) \} .$$

Preuve. — D'après 1.3. il existe un voisinage U de m et $n - d$ fonctions $f_i : U \to \mathbf{R}$ telles que $U \cap M$ soit l'ensemble des zéros communs aux f_i, les différentielles Df_i étant linéairement indépendantes.

Soit $f : U \to \mathbf{R}^{n-d}$ l'application de composantes f_i. On peut extraire de la matrice jacobienne de f un mineur $(n-d) \times (n-d)$ de déterminant non nul. Autrement dit, quitte à permuter les coordonnées de $\mathbf{R}^n$, on peut écrire $\mathbf{R}^n = ((u_1, u_2) : u_1 \in \mathbf{R}^d, u_2 \in \mathbf{R}^{n-d})$ de sorte que $D_2 f \in \mathscr{L}(\mathbf{R}^{n-d}; \mathbf{R}^{n-d})$ soit un isomorphisme. D'après le théorème des fonctions implicites (5.2. chap. 3), il existe une application $g : u_1 \rightsquigarrow u_2$ d'un ouvert U_1 de $\mathbf{R}^d$ dans $\mathbf{R}^{n-d}$ telle que $f(u_1, g(u_1)) = 0$. L'ensemble des zéros de f est donc le graphe $\{ (u_1, g(u_1)) : u_1 \in U_1 \}$ de g.

Exemple. — Le cercle $S^1 = \{ (x, y) \in \mathbf{R}^2 : x^2 + y^2 = 1 \}$ est localement (mais non globalement) le graphe d'applications différentiables de $\mathbf{R}$ dans $\mathbf{R}$: S^1 est la réunion des $U_i \cap S^1$, où $U_1 = \{ |x| < 1, y > 0 \}$, $U_2 = \{ |x| < 1, y < 0 \}$, $U_3 = \{ x > 0, |y| < 1 \}$, $U_4 = \{ x < 0, |y| < 1 \}$. Les $U_i \cap S^1$, $i = 1, \ldots, 4$, étant les graphes respectifs des applications

$$x \rightsquigarrow y = \sqrt{1 - x^2}, \qquad x \rightsquigarrow y = -\sqrt{1 - x^2},$$
$$y \rightsquigarrow x = \sqrt{1 - y^2}, \qquad y \rightsquigarrow x = -\sqrt{1 - y^2}.$$

Paramétrisation. 1.13. — Soit M une sous-variété de $\mathbf{R}^n$ de dimension d. Nous venons d'établir que chaque point $m \in M$ possède un voisinage ouvert U tel que $U \cap M$ soit le graphe d'une application g d'un ouvert V de $\mathbf{R}^d$ dans $\mathbf{R}^{n-d}$. Autrement dit $U \cap M$ est l'image de l'application $p : x \in V \rightsquigarrow (x, g(x)) \in U \subset \mathbf{R}^d \times \mathbf{R}^{n-d} = \mathbf{R}^n$. La différentielle $Dp(x) = (\mathrm{id}_{\mathbf{R}^d}, Dg(x))$ est évidemment injective. D'autre part, p étant une application continue de V dans U, c'est encore une application continue de V sur $U \cap M$ muni de la topologie induite. Enfin, la restriction q, à $M \cap U$, de la projection canonique $\mathbf{R}^n = \mathbf{R}^d \times \mathbf{R}^{n-d} \to \mathbf{R}^d$ sur le premier facteur est continue et $q \circ p = \mathrm{id}_V$, $p \circ q = \mathrm{id}_{U \cap M}$; ainsi p est un homéomorphisme de V sur $U \cap M$.

Cela montre que la condition énoncée dans la cinquième caractérisation suivante des sous-variétés est nécessaire :

Cinquième définition des sous-variétés. — Pour que M soit une sous-variété de classe C^k et de dimension d de $\mathbf{R}^n$, il faut et il suffit que tout point $m \in M$ possède un voisinage ouvert U dans $\mathbf{R}^n$ tel que $U \cap M$ soit l'image d'un ouvert V de $\mathbf{R}^d$ par une application $p : V \to U$ de classe C^k vérifiant

a) $Dp(x)$ est injective pour $x \in V$;

b) p est un homéomorphisme de V sur $U \cap M$ muni de la topologie induite par $\mathbf{R}^n$.

On dit que (V, p) est une paramétrisation (ou une représentation paramétrique) de $U \cap M$.

PREUVE. — Il reste à montrer que la condition est suffisante.

Posons $a = p^{-1}(m)$. Puisque $Dp(a)$ est injective, elle est de rang d et l'on peut trouver d composantes de p dont les différentielles sont linéairement indépendantes en a. Supposons, pour fixer les idées, que ce soient les d premières. Alors, si $\pi_1 : \mathbf{R}^n = \mathbf{R}^d \times \mathbf{R}^{n-d} \to \mathbf{R}^d$ est la projection canonique sur le premier facteur, $D(\pi_1 \circ p)(a) = \pi_1 \circ Dp(a)$ est un isomorphisme et, d'après le théorème d'inversion locale, $\pi_1 \circ p$ est un difféomorphisme d'un voisinage de a sur son image dans $\mathbf{R}^d$. Notons encore V ce voisinage et $U_1 = \pi_1 \circ p(V)$ son image.

p applique continûment V dans l'ouvert U de $\mathbf{R}^n$. Quitte à réduire V, on peut supposer que U est le produit de l'ouvert U_1 de $\mathbf{R}^d \times \{0\}$ par un ouvert U_2 de $\{0\} \times \mathbf{R}^{n-d}$. Désignons par $\pi_2 :$ $\mathbf{R}^n = \mathbf{R}^d \times \mathbf{R}^{n-d} \to \mathbf{R}^{n-d} = \{0\} \times \mathbf{R}^{n-d}$ la projection canonique sur le second facteur. Si $x \in V$ on a $p(x) = (\pi_1 \circ p(x), \pi_2 \circ p(x))$. Comme $\pi_1 \circ p$ est un difféomorphisme de V sur U_1, en posant $y = \pi_1 \circ p(x)$ on peut écrire $p(x) = (y, \pi_2 \circ p_0 [\pi_1 \circ p]^{-1}(y))$. Cela montre que $U \cap M$ est le graphe de $\pi_2 \circ p \circ [\pi_1 \circ p]^{-1} : U_1 \to U_2$. Le théorème résulte alors de 1.11. □

EXEMPLE (TORE DE RÉVOLUTION). 1.14. — Le tore de révolution T^2, obtenu en faisant tourner un cercle de rayon r autour d'une droite D de son plan située à une distance $d > r$ de son centre, est une sous-variété de $\mathbf{R}^3$.

En effet, prenons pour axe des z la droite D et situons le centre du cercle dans le plan $z = 0$. Le tore est l'image de l'application

$$(\theta, \varphi) \in \mathbf{R}^2 \rightsquigarrow (x = (d + r.\cos\varphi).\cos\theta, y = (d + r.\cos\varphi).\sin\theta, z = r.\sin\varphi) \in \mathbf{R}^3 .$$

La restriction p de cette application au produit $I \times J$ de deux intervalles de longueur inférieure à 2π est une paramétrisation de T^2 (vérification immédiate).

EXEMPLE (RUBAN DE MÖBIUS). 1.15. — L'image de l'application

$$(\theta, t) \in \mathbf{R} \times]-1, 1[\to (x = (t.\cos\theta + 2).\cos 2\theta, y = (t.\cos\theta + 2).\sin 2\theta, z = t.\sin\theta) \in \mathbf{R}^3$$

est une sous-variété de $\mathbf{R}^3$.

La restriction de cette application au produit $I \times]-1, 1[$ d'un intervalle de longueur inférieur à π par $]-1, 1[$ est en effet une paramétrisation (le vérifier).

Mise en garde. 1.16. — Une courbe de $\mathbf{R}^n$, c'est-à-dire une application différentiable f d'un intervalle I de $\mathbf{R}$ dans $\mathbf{R}^n$, n'a pas nécessairement pour image une sous-variété ; même si f est un homéomorphisme de I sur $f(I)$.

Contre-exemple : $t \in \mathbf{R} \rightsquigarrow (t^2, t^3) \in \mathbf{R}^2$, à cause du point de rebroussement à l'origine.

2. Espace tangent

Définition. 2.1. — Soit M une sous-variété de l'espace vectoriel E. Un vecteur v de E est dit tangent en m à M s'il est le vecteur-vitesse en m d'une courbe tracée sur M et d'origine m. C'est-à-dire s'il existe un intervalle ouvert I contenant 0 et une application différentiable $\gamma :$ $I \to E$ tels que $\gamma(t) \in M$ pour $t \in I$, $\gamma(0) = m$ et $\gamma'(0) = v$.

L'ensemble des vecteurs tangents en m à M s'appelle l'espace tangent en m à M et se note $T_m M$.

Théorème. 2.2. — ***$T_m M$ est un espace vectoriel de même dimension que M.***

PREUVE. — D'après la définition 1.1., il existe un voisinage ouvert U de m et un difféomorphisme $\varphi : U \to E = \mathbf{R}^d \oplus \mathbf{R}^{n-d}$ tels que $\varphi(U \cap M) \subset \mathbf{R}^d \times \{ 0 \}$. On peut supposer, sans nuire à la généralité, que $\varphi(m) = 0$.

Une courbe γ, tracée sur M et d'origine m, a pour image $\varphi \circ \gamma$ une courbe tracée sur l'espace vectoriel $\mathbf{R}^d \times \{ 0 \}$ et d'origine $\varphi(m) = 0$ et réciproquement.

Il en résulte immédiatement que $D\varphi(m).T_m M = \mathbf{R}^d \times \{ 0 \}$; ce qui montre que $T_m M$ est l'espace vectoriel $D\varphi(m)^{-1}.(\mathbf{R}^d \times \{ 0 \})$ de dimension $d = \dim M$. □

On identifie le plus souvent l'espace *vectoriel* $T_m M$ à l'espace affine passant par m et parallèle à $T_m M$:

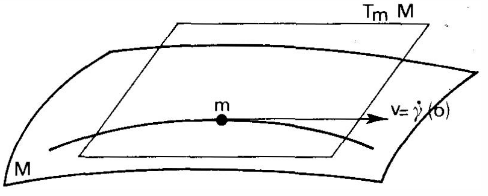


Fig. 7.

Les différentes définitions des sous-variétés conduisent à autant de caractérisations de l'espace tangent. Voici celle correspondant à la seconde définition :

Théorème. 2.3. — ***Reprenons les notations de* 1.3. *Alors $T_m M$ est l'intersection des noyaux des formes linéaires $Df_i(m)$.***

PREUVE. — Cette intersection est un espace vectoriel T de même dimension d que $T_m M$. Tout revient donc à prouver que $T_m M \subset T$.

Soit $\gamma : I \to E$ une courbe tracée sur M et d'origine O. Alors $f_i[\gamma(t)] = 0$ pour $t \in I$. En écrivant que la différentielle est nulle pour $t = 0$, on obtient $Df_i(m).\gamma'(0) = 0$, ce qui montre que $\gamma'(0) \in T$.

Voici la caractérisation correspondant à la troisième définition des sous-variétés.

Théorème. 2.4. — ***Reprenons les notations de* 1.5. *Alors l'espace tangent en m à la sous-variété de niveau $g^{-1}(b)$ est le noyau de $Dg(m)$.***

PREUVE. — Semblable à celle de 2.3. □

EXEMPLE. — La sphère S^n de $\mathbf{R}^{n+1} = \{ (x_1, \dots, x_{n+1}) \}$ est la sous-variété de niveau $g^{-1}(1)$, où $g(x_1, \dots, x_{n+1}) = x_1^2 + \cdots + x_{n+1}^2$.

$v \in T_m S^n$ est caractérisé par $Dg(m).v = 0$, c'est-à-dire $\langle m, v \rangle = 0$ (où $\langle\ ,\ \rangle$ est le produit scalaire usuel de $\mathbf{R}^{n+1}$). On retrouve le fait que $T_m S^n$ est orthogonal au « rayon » m de la sphère.

Point de vue des graphes. 2.5. — ***Reprenons les notations de* 1.11. *Alors l'espace tangent en $m = (u, g(u))$ au graphe de g est le graphe de $Dg(u)$.***

PREUVE. — Désignons par $\pi_1 : E_1 \oplus E_2 \to E_1$ et $\pi_2 : E_1 \oplus E_2 \to E_2$ les projections canoniques. On exprime qu'une courbe $\gamma : I \to E = E_1 \oplus E_2$ est tracée sur

$$M = \{ (x, g(x)) \in E_1 \times E_2 \}$$

en écrivant $\pi_2 \circ \gamma(t) = g[\pi_1 \circ \gamma(t)]$. Différentions et écrivons que $\gamma(0) = (u, g(u))$. On obtient $\pi_2 \circ \gamma'(0) = Dg(u) \circ \pi_1 . \gamma'(0)$ qui s'exprime bien que $\gamma'(0) = (\pi_1 \circ \gamma'(0), \pi_2 \circ \gamma'(0))$ appartient au graphe de $Dg(u)$ [et inversement]. $\square$

Point de vue des paramétrisations. 2.6. — ***Reprenons les notations de* 1.13. *Soient* (V, p) *une paramétrisation de* $U \cap M$ *et* $a \in V$. *Alors l'espace tangent en* $m = p(a)$ *à* M *est* $Dp(a).\mathbf{R}^d$.**

PREUVE. — Soit $s : I \to V$ une courbe différentiable d'origine a. Alors $p \circ s$ est une courbe tracée sur M et d'origine $p(a) = m$. Son vecteur-vitesse $D(p \circ s)(0) = Dp(a).s'(0)$ décrit un espace vectoriel T de dimension d lorsque $s'(0)$ décrit $\mathbf{R}^d$, car $Dp(a)$ est injective. Ainsi $T = Dp(a).\mathbf{R}^d$ est inclus dans $T_m M$ et ces deux espaces sont de dimension d ; ils coïncident donc. $\square$

EXEMPLE. 2.7. — Nous avons vu (1.2) qu'un ouvert M de E est une sous-variété de E de même dimension que E. Son espace tangent en $m \in M$ n'est autre que E. En effet, en reprenant les notations de 2.2., on peut choisir $U = M$ et $\varphi = \mathrm{id}_U$; alors $T_m M = D\varphi(m)^{-1}.E = E$.

C'est ainsi que l'espace tangent en un point quelconque à la sous-variété ouverte GL (n) de End $(E) = \mathbf{R}^n$ [voir 1.7.], n'est autre que End (E). Examinons ce qu'il en est pour quelques sous-groupes de GL (n) qui sont aussi des sous-variétés de End (E).

EXEMPLE. 2.8. — Nous avons vu (1.8) que le groupe unimodulaire SL (n), formé des endomorphismes de $\mathbf{R}^n$ de déterminant 1, est une sous-variété de dimension $n^2 - 1$ de End $(\mathbf{R}^n)$. Cherchons son espace tangent en $I = \mathrm{id}_{\mathbf{R}^n}$.

Soit N le sous-espace de End $(\mathbf{R}^n)$ formé des endomorphismes u de trace Tr (u) nulle. Si exp désigne l'application exponentielle, considérons la courbe $\gamma : t \in \mathbf{R} \rightsquigarrow \exp(t.u) \in$ GL (n), $u \in N$. C'est une courbe d'origine I qui, d'après (2.4. chap. 5), est tracée sur SL (n). D'après (2.6. chap. 5), son vecteur-vitesse en I est $\gamma'(0) = u$. L'espace T_I SL (n) contient donc N. Comme dim T_I SL (n) = dim SL $(n) = n^2 - 1$ = dim N, on voit que T_I SL $(n) = N$.

Nous apprendrons ultérieurement à déterminer l'espace tangent à SL (n) en un point quelconque.

EXEMPLE. 2.9. — Nous avons vu (1.9) que le groupe orthogonal O (n) est une sous-variété de dimension $n(n-1)/2$ de End $(\mathbf{R}^n)$.

Cherchons un espace tangent en $I = \mathrm{id}_{\mathbf{R}^n}$.

Soit A le sous-espace de End $(\mathbf{R}^n)$ formé des endomorphismes a antisymétriques : ${}^t a = -a$ [c'est-à-dire $\langle a(x), x \rangle = 0$ pour $x \in \mathbf{R}^n$, où $\langle\ ,\ \rangle$ est le produit scalaire usuel]. Considérons la courbe $\gamma : t \in \mathbf{R} \rightsquigarrow \exp(t.a)$, où $a \in A$. Si $F(t)$ désigne le carré scalaire de $\exp(t.a).x$, où $x \in \mathbf{R}^n$, la règle de Leibniz et (2.6. chap. 5) entraînent $F'(t) = 2 \langle a.\exp(t.a).x, \exp(ta).x \rangle$, qui est nul puisque a est antisymétrique. Par conséquent $F(t) = F(0)$; ce qui montre que la courbe γ d'origine I est tracée sur O (n). Toujours d'après (2.6. chap. 5), $\gamma'(0) = a$. L'espace T_I O (n) contient donc A. Comme dim T_I O (n) = dim O $(n) = n(n-1)/2$ = dim A, on voit que T_I O $(n) = A$.

EXEMPLE. 2.10. — Munissons $\mathbf{R}^{2n} = \{ (q, p) : q \in \mathbf{R}^n, p \in \mathbf{R}^n \}$ de la forme symplectique ω définie par (voir 2.12. chap. 6) : $\omega[(q, p), (q', p')] = \langle q, p' \rangle - \langle q', p \rangle$.

Le lecteur montrera que le sous-groupe Sp (n) de GL $(2n)$, formé des opérateurs s tels que $\omega[s.(\ ), s.(\ )'] = \omega[(\ ), (\ )']$, est une sous-variété de End $(\mathbf{R}^{2n})$.

Il montrera aussi que T_I Sp (n) est formé des applications symplectiques infinitésimales, c'est-à-dire des $\sigma \in$ End $(\mathbf{R}^{2n})$ tels que $\omega[\sigma.(\ ), (\ )'] + \omega[(\ ), \sigma.(\ )'] = 0$ pour $(\ ), (\ )' \in \mathbf{R}^{2n}$.

3. Applications différentiables

Définition. 3.1. — Soient M une sous-variété de l'espace vectoriel E, M' une sous-variété de l'espace vectoriel E', et $f : M \to M'$ une application. On dit que f est de classe C^k en $m \in M$ s'il existe un voisinage ouvert U de m et une application

$$\hat{f} : U \to \hat{f}(U) \text{ telle que } \hat{f}|_{U \cap M} = f_{U \cap M}.$$

Si f est de classe C^k en chaque point de M, on dit que f est de classe C^k.
Si M est un ouvert de E, on peut prendre $\hat{f} = f$, et l'on retrouve la définition usuelle.

Application linéaire tangente. 3.2. — Avec les notations précédentes, $D\hat{f}(m)$ applique $T_m M$ dans $T_{f(m)} M'$, et sa restriction à $T_m M$ ne dépend que de f, et non pas du choix de $\hat{f}$.

PREUVE. — Soit $\gamma : I \to U \cap M$ une courbe tracée sur M et d'origine m. Alors $\hat{f} \circ \gamma$ est une courbe tracée sur M' et d'origine $\hat{f}(m) = f(m)$. Comme $\hat{f} \circ \gamma(t) = f \circ \gamma(t)$ pour $t \in I$, on a $D\hat{f}(m).\gamma'(0) = (\hat{f} \circ \gamma)'(0) = (f \circ \gamma)'(0)$; ce qui montre que $D\hat{f}(m).\gamma'(0) \in T_{f(m)} M'$ ne dépend pas de $\hat{f}$.

La restriction de $D\hat{f}(m)$ à $T_m M$ est donc l'unique application linéaire $T_m f : T_m M \to T_{f(m)} M'$ telle que $T_m f.\gamma'(0) = (f \circ \gamma)'(0)$ pour toute courbe différentiable γ tracée sur M et d'origine m. On l'appelle *l'application linéaire tangente à f en m*. □

EXEMPLE. — Si M est un ouvert de E, alors $T_m f$ n'est rien d'autre que $Df(m)$, considérée comme application linéaire de E dans le sous-espace $Df(m).E$ de E'.

On peut encore interpréter $T_m f$ comme l'application $(m, \gamma'(0)) \rightsquigarrow (f(m), Df(m).\gamma'(0))$ que nous avions introduite en (2.6. chap. 1), précisément sous le nom d'application tangente.

Comme en (2.6. chap. 1), on a le résultat suivant, dont la preuve est immédiate :

Théorème de composition des applications. 3.3. — ***Si $f : M \to M'$ et $g : M' \to M''$ sont des applications de classe C^k de variétés de classe C^k, alors $g \circ f : M \to M''$ est une application de classe C^k et $T_m(g \circ f) = T_{f(m)} g \circ T_m f$.***

Difféomorphismes de sous-variétés. 3.4. — On dit que $f : M \to M'$ est un difféomorphisme de la sous-variété M sur la sous-variété M', si f est bijective et si f et f^{-1} sont différentiables.

On dit que f est un difféomorphisme local en $m \in M$, s'il existe un voisinage ouvert U_M de m sur M (pour la topologie induite par E) et un voisinage ouvert $V_{M'}$ de $f(m)$ sur M' (pour la topologie induite par E') telle que la restriction de f à U_M soit un difféomorphisme de U_M sur $V_{M'}$.

S'il en est ainsi, $T_m f$ est bijectif, car $f^{-1} \circ f = \mathrm{id}_{U_M}$ implique $T_{f(m)} f^{-1} \circ T_m f = \mathrm{id}_{T_m M}$.

EXEMPLE.— Si (V, p) est une paramétrisation d'un ouvert $U_M = U \cap M$ de M, alors p est un difféomorphisme de la sous-variété ouverte V de $\mathbf{R}^d$ sur U_M.

EXEMPLE. — Reprenons l'identification $\mathrm{End}(\mathbf{R}^n) = \mathbf{R}^{n^2}$.

Si $u \in \mathrm{GL}(n)$, l'application $L_u : \mathrm{End}(\mathbf{R}^n) \to \mathrm{End}(\mathbf{R}^n)$, définie par $L_u(g) = u \circ g$, est de classe C^∞. On l'appelle la translation à gauche par u. Elle possède un inverse, à savoir $L_{u^{-1}}$. C'est donc un difféomorphisme.

On définirait de la même manière une translation à droite $R_u : g \rightsquigarrow g \circ u$.

Supposons que $u \in \mathrm{SL}(n)$. Evidemment L_u laisse $\mathrm{SL}(n)$ invariante, puisque c'est un groupe. Par conséquent, la restriction de L_u à $\mathrm{SL}(n)$ est un difféomorphisme de $\mathrm{SL}(n)$ sur $\mathrm{SL}(n)$. Ce difféomorphisme transporte donc l'espace tangent N en $I = \mathrm{id}_{\mathbf{R}^n}$ sur l'espace tangent en u ; autrement dit $T_u \mathrm{SL}(n) = u \circ N$.

On a un résultat semblable pour $\mathrm{O}(n)$: $T_u \mathrm{O}(n) = u \circ A$.

Théorème d'inversion locale pour les sous-variétés. 3.5. — ***Avec les notations précédentes, si $T_m f$ est bijectif, f est un difféomorphisme local.***

PREUVE. — Choisissons des paramétrisations (p, V) et (p', V') de M et de M' aux voisinages respectifs de m et $f(m)$. Quitte à réduire V, on peut supposer que $f \circ p(V) \subset p'(V')$.

Soit $a = p^{-1}(m)$. Alors, d'après 3.3., $T_a(p'^{-1} \circ f \circ p) = T_{f(m)} p'^{-1} \circ T_m f \circ T_a f$ est bijectif, comme composé de bijections. Le théorème d'inversion locale pour les espaces vectoriels implique qu'il existe un voisinage ouvert de a dans V (notons-le encore V) et un voisinage ouvert de $p'^{-1}(f(m))$ dans V' (notons-le encore V') tels que $p'^{-1} \circ f \circ p : V \to V'$ soit un difféomorphisme. Il en résulte que le composé de difféomorphismes $f = p' \circ (p'^{-1} \circ f \circ p) \circ p^{-1}$ est un difféomorphisme. □

Notes. — Ce chapitre n'est qu'une introduction à des ouvrages plus complets, comme celui de P. Malliavin. Le lecteur trouvera aussi dans (J. Milnor) des applications directes de ce qui précède. En particulier une démonstration géométrique du théorème fondamental de l'algèbre : Tout polynôme $P : \mathbf{C} \to \mathbf{C}$ non constant possède au moins une racine.

On trouvera une application de 3.5. dans l'Appendice H.

10 CALCUL DES VARIATIONS

Ce chapitre est consacré à l'étude des extrema d'une fonction à valeurs réelles définie sur un espace vectoriel normé E. On s'attache particulièrement aux espaces E dont les éléments sont des courbes d'un espace vectoriel. Les conditions nécessaires de l'extremum s'expriment alors par une équation différentielle, l'équation d'Euler-Lagrange.

Le plus souvent ces questions ont pour origine des situations concrètes (Mécanique, Physique) qui présentent des symétries. L'invariance de l'équation d'Euler-Lagrange par ces symétries se traduit par l'existence d'intégrales premières : c'est le contenu du théorème de Noether.

1. Extrema libres. Extrema liés

Définition. 1.1. — Soient E un e.v. normé, A une partie de E, f une fonction définie sur A et à valeurs réelles. On dit que f présente un *minimum relatif* en $a \in A$ s'il existe un voisinage ouvert U de a dans E tel que $f(a) \leqslant f(x)$ pour $x \in U \cap A$.

On définirait de la même manière un *maximum relatif*. D'ailleurs f présente un maximum relatif en a si, et seulement si, $-f$ présente un minimum relatif en a.

On dit que f présente un *extremum relatif* en a si elle présente un minimum ou un maximum relatif en a.

Le résultat qui suit a été établi en (3.2, chap. 3) :

Théorème. 1.2. — ***Pour qu'une fonction différentiable $f : A \to \mathbf{R}$ présente un extremum relatif en un point a intérieur à A, il faut que $Df(a) = 0$.***

Remarques. 1.3. — *a*) La condition $Df(a) = 0$ est nécessaire. Elle n'est pas suffisante.

Contre-exemple. — Si $f : \mathbf{R} \to \mathbf{R}$ est définie par $f(t) = t^3$, alors $Df(0) = 0$. Mais 0 n'est pas un extremum relatif.

Contre-exemple. — Si $f : \mathbf{R}^2 \to \mathbf{R}$ est définie par $f(x, y) = x^2 - y^2$, on a $Df(0, 0) = 0$. Pourtant f prend des valeurs des deux signes dans un voisinage arbitraire de $(0, 0)$.

b) Si l'on utilise la condition $Df(a) = 0$ pour déterminer les extrema de f sur A, et si A n'est pas un ouvert, on est condamné à examiner séparément les valeurs que prend f sur la frontière de A. Mais cette frontière peut être A tout entier si, par exemple, A est une surface de $\mathbf{R}^3$. Nous allons voir comment se tirer d'affaire si la frontière de A est « assez régulière », plus précisément si c'est une sous-variété de E.

Extremum lié. 1.4. — Soient M une sous-variété différentiable de E, U un ouvert de E qui coupe M, f une fonction définie sur U et à valeurs réelles. On dit que f présente un extremum (relatif) lié en $m \in U \cap M$ si la restriction de f à $U \cap M$ présente un extremum relatif en m.

Exemple. — Si M est un sous-espace vectoriel de E et si f est différentiable, alors $f|_M$ est différentiable et $D(f|_M)(m) = Df(m)|_M$ (voir 2.4. chap. 1).

La condition d'extremum lié s'exprime donc par $Df(m)|_M = 0$. Cela n'implique évidemment pas $Df(m) = 0$.

Supposons, plus particulièrement encore, que E soit un espace euclidien, de produit scalaire $\langle\ ,\ \rangle$. La condition $Df(m).v = 0$, pour $v \in M$, s'écrit $\langle \operatorname{grad} f(m), v \rangle = 0$. Autrement dit, le gradient de f en m est orthogonal à M.

Nous allons généraliser ce résultat.

Théorème. 1.5. — ***Gardons les notations ci-dessus. Pour que la fonction différentiable f présente un extremum lié en m sur la sous-variété M, il faut que la restriction de Df(m) à l'espace tangent $T_m M$ soit nulle.***

PREUVE. — Soient $v \in T_m M$ et $\gamma : I \to E$ une courbe différentiable tracée sur M, d'origine m et telle que $\gamma'(0) = v$. Evidemment $(f|_{M \cap U}) \circ \gamma(t) = f \circ \gamma(t)$ pour $t \in I$, et cette fonction de t présente un extremum relatif pour $t = 0$. D'après le théorème de différentiation des applications composées (3.3. chap. 9) et 1.2., on a donc $T_m(f|_{U \cap M}).v = T_m f.v = Df(m).v = \frac{\mathrm{d}}{\mathrm{d}t}[f \circ \gamma](0) = 0$. □

A chacune des caractérisations de l'espace tangent correspond une traduction de ce résultat. Donnons la plus utile et, pour cela, commençons par un lemme.

Lemme. — ***Soient a, $b_1, \ldots, b_k$ des formes linéaires sur un espace vectoriel E de dimension n. Supposons que $b_1, \ldots, b_k$ soient linéairement indépendantes et que leurs noyaux soient contenus dans celui de a. Alors a est combinaison linéaire des b_i.***

PREUVE. — Complétons les b_i de façon à former une base $b_1, \ldots, b_n$ de E^*, et soit $e_1, \ldots, e_n$ la base duale. Il existe donc des constantes c_i telles que $a = \sum c_i.b_i$. D'autre part l'intersection des noyaux de $b_1, \ldots, b_k$ est l'espace engendré par $e_{k+1}, \ldots, e_n$. On a donc $0 = a(e_r) = \sum c_i.b_i(e_r) = c_r$ si $r = k+1, \ldots, n$. Par conséquent $a = c_1.b_1 + \cdots + c_k.b_k$. □

Théorème des multiplicateurs de Lagrange. 1.6. — ***Reprenons la seconde définition* (1.3. *chap.* 9) *des sous-variétés. Soient U un ouvert de $E = \mathbf{R}^n$, $g_1, \ldots, g_k$ des fonctions de classe C^1, définies sur U, à valeurs réelles, et dont les différentielles sont linéairement indépendantes en chaque point de U. On sait que l'ensemble des zéros communs aux g_i est une sous-variété M de $\mathbf{R}^n$.***

Donnons-nous une fonction différentiable $f : U \to \mathbf{R}$. Alors, pour que f présente un extremum lié en $m \in U$ sur M, il faut qu'il existe des constantes c_i telles que $Df(m) = c_1.Dg_1(m) + \cdots + c_k.Dg_k(m)$.

PREUVE. — D'après (2.3. chap. 9), $T_m M$ est l'intersection des noyaux des $Dg_i(m)$. D'après 1.5, $T_m M$ appartient au noyau de $Df(m)$. Il suffit maintenant d'appliquer le lemme précédent.

L'extremum de f sur M est parfois appelé l'extremum de f soumise aux contraintes $g_1 = \cdots = g_k = 0$ (d'où le nom d'extremum lié).

Afin de déterminer le point m, on peut résoudre le système $D_j f(m) = \sum c_i.D_j g_i(m)$, $j = 1, \ldots, n$, auquel on adjoint $g_1(m) = \cdots = g_k(m) = 0$. C'est un système de $n + k$ équations à $k + n$ inconnues $c_1, \ldots, c_k$, $m = (m_1, \ldots, m_n)$. Les inconnues auxiliaires c_i s'appellent *multiplicateurs de Lagrange*. □

Application. 1.7. — Etant donnée une surface M de $\mathbf{R}^3$, d'équation $g(x, y, z) = 0$, et un point $A \in \mathbf{R}^3$, les points $m \in M$ dont la distance à A est extremum sont donnés par la recherche des extrema de $f(m) = \langle m - A, m - A \rangle$ astreint à $g(m) = 0$.

$Df(m) = c.Dg(m)$ s'écrit $2(m - A) = c.\operatorname{grad} g(m)$. Géométriquement, cela signifie que la normale à M aux points m cherchés passe par A. Mais tous ces points ne donnent pas nécessairement lieu à un extremum : la sphère de centre A passant par m ne reste pas nécessairement du même côté de M dans un voisinage de m.

Application. 1.8. — Soit u en endomorphisme symétrique d'un espace euclidien $E = \mathbf{R}^n$: $\langle u(x), y \rangle = \langle x, u(y) \rangle$ pour $x, y \in E$.

Nous allons montrer que E possède une base de vecteurs propres de u.

La fonction $f : E \to \mathbf{R}$ définie par $f(x) = \langle u(x), x \rangle$ est différentiable. Elle atteint donc son maximum sur la sphère unité S (qui est compacte) en un point e_1. Pour trouver e_1, cherchons les extrema liés de f astreint à $g(x) = \langle x, x \rangle = 1$. D'après la règle de Leibniz, $Df(x).h = \langle u(x), h \rangle + \langle u(h), x \rangle = 2 \langle u(x), h \rangle$ et $Dg(x).h = 2 \langle x, h \rangle$ pour $h \in E$. La relation $Df(e_1) = c.Dg(e_1)$ s'écrit donc $u(e_1) = c.e_1$, où $f(e_1) = \langle u(e_1), e_1 \rangle = c$. Un point e_1 de $S = g^{-1}(1)$ où f est maximum est donc un vecteur propre de u de valeur propre maximum.

Si maintenant $\langle y, e_1 \rangle = 0$, on a $\langle u(y), e_1 \rangle = \langle y, u(e_1) \rangle = c.\langle y, e_1 \rangle = 0$. L'espace E' orthogonal à e_1 est donc invariant par u, et la restriction de u à E' est encore un endomorphisme symétrique. Comme $\dim E' = \dim E - 1$, on voit, par récurrence sur la dimension de E, que E possède une base de vecteurs propres de u qui sont orthogonaux deux à deux.

2. Conditions du second ordre pour un extremum

Revenons à une fonction différentiable f définie sur un ouvert U d'un espace vectoriel normé E, et à valeurs réelles. Nous avons vu que si f présente un extremum en a, alors $Df(a) = 0$. Nous allons obtenir une condition nécessaire supplémentaire en supposant que f est deux fois différentiable.

Théorème. 2.1. — ***Soit $f : U \to \mathbf{R}$ une fonction deux fois différentiable en $m \in U$. Si f admet un minimum relatif en m, on a $D^2 f(m).(X, X) \geqslant 0$ pour $X \in E$.***

Preuve. — Puisque $Df(m) = 0$, la formule de Taylor donne $f(m + h) - f(m) = \frac{1}{2} D^2 f(m).(h, h) + r(h)$, où $\lim_{h \to 0} | r(h) |/\| h \|^2 = 0$. Puisque f admet un minimum relatif en m, on a $f(m + h) - f(m) \geqslant 0$ pour $\| h \|$ assez petit, disons pour $\| h \| < \varepsilon$. Donc $D^2 f(m).(h, h) + 2.r(h) \geqslant 0$ pour $\| h \| < \varepsilon$. Donnons-nous $X \in E$ et prenons $t \in \mathbf{R}$ tel que $t \neq 0$, $\| t.X \| < \varepsilon$. Puisque $D^2 f(m)$ est bilinéaire, on obtient $t^2.D^2 f(m).(X, X) + 2\, r(t.X) \geqslant 0$. Divisons les deux membres par t^2 et faisons tendre t vers zéro. On obtient $D^2 f(m).(X, X) \geqslant 0$.

Si f admet un maximum relatif en m, on a évidemment $D^2 f(m).(X, X) \leqslant 0$ pour $X \in E$. □

Remarques. — *a*) Les conditions $Df(m) = 0$ et $D^2 f(m).(X, X) \geqslant 0$ pour $X \in E$ ne sont pas suffisantes pour que f présente un minimum relatif en m. Contre-exemple : si $f : \mathbf{R}^2 \to \mathbf{R}$ est définie par $f(x, y) = x^2 - 3.y^3$, on a $Df(0, 0) = 0$, $D^2 f(0, 0).(X, Y) = 2(X)^2 \geqslant 0$. Pourtant f prend des valeurs des deux signes dans un voisinage arbitraire de (0, 0).

b) Supposons que f soit trois fois différentiable. Si $Df(m) = 0$, $D^2 f(m) = 0$ tandis que $D^3 f(m) \neq 0$, alors f ne peut présenter un extremum local en m. En effet, la formule de Taylor donne $f(m + tX) - f(m) = \frac{1}{6} t^3.D^3 f(m).(X, X) + r(tX)$, où $\| r(tX) \|/t^3 \to 0$ si $t \to 0$. Cela montre que le membre de gauche se comporte (au signe près) comme t^3 au voisinage de m; il ne peut donc rester de signe constant.

Une condition suffisante pour un minimum relatif. 2.2. — Soit f une fonction à valeurs réelles, de classe C^3, définie sur un ouvert U d'un espace de *Hilbert* E. Supposons que $m \in U$ soit un point critique non dégénéré de f et que $D^2 f(m).(X, X) > 0$ pour $X \in E - \{ 0 \}$. Alors f admet un minimum relatif strict au point m; c'est-à-dire que $f(m) < f(x)$ pour x voisin de m et distinct de m.

Preuve. — D'après le théorème de Morse-Palais (3.2. chap. 8), il existe un C^1-difféomorphisme local φ au voisinage de m tel que $\varphi(m) = m$ et $f(m + h) - f(m) = \frac{1}{2} \cdot D^2 f(m) \cdot (\varphi(h), \varphi(h))$. Le théorème en résulte. □

En fait les hypothèses peuvent être affaiblies. On a le résultat suivant :

Théorème. 2.3. — ***Soient U un ouvert d'un espace de* Banach *E, $f : U \to \mathbf{R}$ une fonction* deux fois différentiable *en $m \in U$. Si m est un point critique, non dégénéré de f et si $D^2 f(m).(X, X) > 0$ pour $X \in E - \{0\}$, alors f admet un minimum strict en m.***

Commençons par un lemme.

Lemme. — ***Il existe une constante $c > 0$ telle que $D^2 f(m).(X, X) \geqslant c.\| X \|^2$ pour $X \in E$.***

Preuve. — Posons $L = D^2 f(m)$. Puisque L est non dégénérée, l'application

$$x \in E \rightsquigarrow L(x, .) \in E^*$$

est un isomorphisme. Si M est la norme de l'isomorphisme inverse, on a donc $\| x \|_E \leqslant M.\| L(x, .) \|_{E^*}$. Mais $\| L(x, .) \|_{E^*} = \sup\limits_{\| y \| = 1} | L(x, y) |$; x étant donné, il existe donc $y \in E$ de norme 1 tel que $2 | L(x, y) | \geqslant \| L(x, .) \|_{E^*}$.

Par conséquent

$$\| x \| \leqslant 2 M.| L(x, y) | . \tag{2.4}$$

Maintenant, $L = D^2 f(m)$ étant symétrique d'après le théorème de Schwarz, on a $0 < L(x + ty, x + ty) = t^2.L(y, y) + 2t.L(x, y) + L(x, x)$ pour $t \in \mathbf{R}$. Le membre de droite est donc un trinôme du second degré en t dont le discriminant est négatif; si $\| L \|$ est la norme de la forme bilinéaire continue L (voir 3.2. Appendice A) on a donc

$$[L(x, y)]^2 < L(x, x).L(y, y) \leqslant \| L \|.L(x, x) .$$

Avec (2.4) cela implique $\| x \|^2 \leqslant 4 M^2.\| L \|.L(x, x)$. En posant $c^{-1} = 4 M^2.\| L \|$ on a donc bien $c.\| x \|^2 \leqslant L(x, x)$. □

Preuve du théorème. — La formule de Taylor et le lemme précédent entraînent

$$f(m + h) - f(m) = \frac{1}{2} D^2 f(m).(h, h) + r(h) \geqslant \frac{c}{2} \cdot \| h \|^2 + r(h) ,$$

où $r(h) = o(\| h \|^2)$. Au nombre $c/2$ correspond $\delta > 0$ tel que $\| h \| < \delta$ entraîne $| r(h) | < \frac{c}{4} \cdot \| h \|^2$. Si $\| h \| < \delta$ on a donc $f(m + h) - f(m) \geqslant \frac{c}{4} \| h \|^2$. □

Note. 2.5. — Considérons une fonction $f : \mathbf{R} \to \mathbf{R}$ deux fois différentiable, qui ne possède qu'un *seul point critique* a non dégénéré où $f''(a) > 0$. Le théorème précédent montre que a est un minimum local. En fait, c'est un *minimum global*. Supposons en effet qu'il existe b tel que $f(b) \leqslant f(a)$. D'après le théorème de Rolle, il existerait un point c entre a et b où $f'(c) = 0$. En contradiction avec l'unicité du point critique a.

Le raisonnement qui précède est particulier aux fonctions d'une variable réelle. Pourtant, le résultat admet la généralisation suivante, que nous admettrons :

Soit f une fonction à valeurs réelles, de classe C^3, définie sur un espace de *Hilbert E*. Supposons que f ne possède qu'un seul point critique m, que ce point critique soit non dégénéré et que $D^2 f(m).(h, h) > 0$ pour $h \in E$. Alors $f(m) < f(x)$ pour tout $x \in E$ distinct de m.

Idée de la démonstration. — Les courbes intégrales du champ de vecteur grad f sont toutes issues du point m. Elles recouvrent E, et la restriction de f à l'une quelconque d'entre elles est une fonction croissante à partir du point a.

Dans ce qui suit nous allons particulariser l'espace E et la fonction f.

3. Espaces de courbes. Equations d'Euler-Lagrange

L'espace des courbes de classe C^1. 3.1. — Soient V un espace vectoriel normé sur $\mathbf{R}$, $I = [a, b]$, $a < b$, un segment. Une application $\gamma : I \to V$ de classe C^1 est appelée une C^1-courbe, paramétrée par $t \in I$, de l'espace V.

L'ensemble $C^1(I; V)$ de ces courbes est évidemment un e.v. sur $\mathbf{R}$: si $k \in \mathbf{R}$ et si $\gamma_1, \gamma_2 \in C^1(I; V)$, on définit encore des C^1-courbes $k.\gamma_1$ et $\gamma_1 + \gamma_2$ en posant $(k.\gamma_1)(t) = k.\gamma_1(t)$, $(\gamma_1 + \gamma_2)(t) = \gamma_1(t) + \gamma_2(t)$.

L'espace vectoriel réel $C^1(I; V)$ possède une norme naturelle. En effet, puisque I est compact et que γ et γ' sont continues, $\| \gamma(t) \|$ et $\| \gamma'(t) \|$ restent bornés sur I. On peut donc poser $\| \gamma \|_{C^1} = \sup_{t \in I} \| \gamma(t) \| + \sup_{t \in I} \| \gamma'(t) \|$. On vérifie sans peine que $\| \;\|_{C^1}$ est une norme. On l'appelle la C^1-norme.

Supposons que l'espace V soit un espace de Banach. Une adaptation immédiate de la preuve donnée au (c.1.5. Appendice A) montre que $C^1(I; V)$ est complet pour la C^1-norme ; c'est donc un espace de Banach.

Fonctions de courbes. 3.2. — On peut faire correspondre de bien des manières un nombre réel à une C^1-courbe γ. Supposons que V soit un espace euclidien ; on peut, par exemple, attacher à γ sa longueur euclidienne $\int_a^b \| \gamma'(t) \| . \mathrm{d}t$. On peut aussi, si l'on interprète $\gamma(t)$ comme la position dans V à l'instant t d'un point de masse $+ 1$, lui attacher l'action $\frac{1}{2} \int_a^b \| \gamma'(t) \|^2 . \mathrm{d}t$.

Ces exemples rentrent dans le cadre suivant, auquel nous nous limiterons :

Soit $L : \mathbf{R} \oplus V \oplus V \to \mathbf{R}$ une fonction de classe C^1 (appelée le *lagrangien*). Etant donné $\gamma \in C^1(I; V)$, alors $t \rightsquigarrow L(t, \gamma(t), \gamma'(t))$ est une fonction continue. Elle est donc intégrable, et l'on définit une fonction $\hat{L} : C^1(I; V) \to \mathbf{R}$ en posant

$$\hat{L}(\gamma) = \int_a^b L(t, \gamma(t), \gamma'(t)) . \mathrm{d}t . \tag{3.3}$$

Nous allons montrer que $\hat{L}$ est de classe C^1 et calculer sa différentielle.

Commençons par un lemme intéressant en soi.

Lemme de différentiation sous le signe somme. 3.4. — ***Soient U un ouvert d'un* e.v. *normé E, $I = [a, b]$, $a < b$, un segment, et f une application continue de $I \times U$ dans un* e.v. *normé F.***

Supposons que $D_2 f$ existe et soit continue. Posons $g(u) = \int_a^b f(t, u) . \mathrm{d}t$. Alors g est de classe C^1 sur U et $Dg(u) = \int_a^b D_2 f(t, u) . \mathrm{d}t$.

PREUVE. — Soit $u \in U$. Quitte à choisir un voisinage ouvert assez petit de u, noté encore U, on peut supposer que $D_2 f$ est bornée sur $I \times U$. Soit $l : E \to F$ l'application linéaire continue $h \in E \rightsquigarrow \int_a^b D_2 f(t, u) . h . \mathrm{d}t$. D'après le théorème fondamental du calcul intégral, on a

$$\begin{aligned} g(u + h) - g(u) - l(h) &= \int_a^b [f(t, u + h) - f(t, u) - D_2 f(t, u) . h] . \mathrm{d}t \\ &= \int_a^b \left[\int_0^1 D_2 f(t, u + s.h) . h . \mathrm{d}s - D_2 f(t, u) . h \right] \mathrm{d}t \\ &= \int_a^b \left\{ \int_0^1 [D_2 f(t, u + sh) - D_2 f(t, u)] . h . \mathrm{d}s \right\} . \mathrm{d}t . \end{aligned}$$

Il en résulte

$$\| g(u+h) - g(u) - l(h) \| \leqslant (b-a).\| h \|.\sup \| D_2 f(t, u+sh) - D_2 f(t,u) \|,$$

où la borne supérieure est prise pour $0 \leqslant s \leqslant 1$ et $a \leqslant t \leqslant b$.

Mais $D_2 f$ est uniformément continue sur le compact $I \times \{ u \}$. Etant donné $\varepsilon > 0$ il existe donc $\delta > 0$ tel que $\| h \| < \delta$ entraîne $\sup \| \quad \| < \varepsilon$. Cela prouve que $l = Dg(u)$.

Comme $u \rightsquigarrow \int_a^b D_2 f(t,u).\mathrm{d}t$ est continue, g est de classe C^1. □

Théorème. 3.5. — ***Si le lagrangien L est de classe C^1, alors l'application $\hat{L} : C^1(I; V) \to \mathbf{R}$, définie par* (3.3), *est de classe C^1 et $D\hat{L}(\gamma).h = \int_a^b [D_2 L(\ ).h(t) + D_3 L(\ ).h'(t)].\mathrm{d}t$, pour $h \in C^1(I; V)$ et où l'on a abrégé $(t, \gamma(t), \gamma'(t))$ en $(\ )$.***

PREUVE. — On va appliquer le lemme précédent en prenant $U = E = C^1(I; V)$, $F = \mathbf{R}$ et $f(t, \gamma) = L(t, \gamma(t), \gamma'(t))$.

L'application f est continue, comme composée d'applications manifestement continues :

$$I \times C^1(I; V) \xrightarrow{M} I \oplus V \oplus V \xrightarrow{L} \mathbf{R}$$
$$(t, \gamma) \rightsquigarrow (t, \gamma(t), \gamma'(t)) \rightsquigarrow L(t, \gamma(t), \gamma'(t)).$$

Montrons que $D_2 f$ existe et est continue. Puisque L est de classe C^1, il suffit de montrer que $D_2 M$ existe et est continue. Soient $\gamma, h \in C^1(I; V)$. On a $M(t, \gamma + h) - M(t, \gamma) = (0, h(t), h'(t))$. Mais l'application $h \in C^1(I; V) \rightsquigarrow (0, h(t), h'(t)) \in I \oplus V \oplus V$, qui est évidemment linéaire, est continue, car $\| (0, h(t), h'(t)) \| = \| h(t) \| + \| h'(t) \| \leqslant \| h \|_{C^1}$. Par conséquent $D_2 M$ existe et $D_2 M(t, \gamma).h = (0, h(t), h'(t))$, qui ne dépend pas de γ et qui est continue en t.

Appliquons le théorème de différentiation des applications composées à $f = L \circ M$: si $h \in C^1(I; V)$, on obtient

$$D_2 f(t, \gamma).h = DL(M(t, \gamma)).D_2 M(t, \gamma).h = DL(t, \gamma(t), \gamma'(t)).(0, h(t), h'(t)) ;$$

c'est-à-dire

$$D_2 f(t, \gamma).h = D_2 L(\ ).h(t) + D_3 L(\ ).h'(t), \tag{3.6}$$

où $(\ ) = (t, \gamma(t), \gamma'(t))$.

Puisque L est de classe C^1, on constate a posteriori que $D_2 f(t, \gamma)$ dépend continûment de (t, γ).

Avec notre choix de f, les hypothèses du lemme précédent sont vérifiées et l'on obtient $D\hat{L}(\gamma).h = \int_a^b D_2 f(t, \gamma).h.\mathrm{d}t$; ce qui, compte tenu de (3.6), démontre le théorème.

Un problème d'extremum lié. 3.7. — Gardons les notations ci-dessus. On peut se proposer de chercher les extrema de la fonction $\hat{L} : C^1(I; V) \to \mathbf{R}$, c'est-à-dire les courbes γ qui rendent minimum (ou maximum) $\hat{L}(\gamma) = \int_a^b L(t, \gamma(t), \gamma'(t)).\mathrm{d}t$.

Dans la pratique on rencontre plutôt des problèmes du type suivant : Trouver, par exemple, le plus court chemin de A à B dans l'espace euclidien V. Ce qui revient à chercher les C^1-courbes γ d'origine A et d'extrémité B rendant minimum $\int_a^b \| \gamma'(t) \|.\mathrm{d}t$.

Plus généralement, étant donnés deux points A et B de l'espace V où sont tracées les courbes, trouver une courbe $\gamma \in C^1(I; V)$ d'origine $\gamma(a) = A$ et d'extrémité $\gamma(b) = B$ rendant $\hat{L}(\gamma)$ extremum.

Nous sommes donc amenés à chercher les extrema de la restriction de $\hat{L}$ au sous-ensemble $\Gamma(A, B) = \{ \gamma \in C^1(I; V) : \gamma(a) = A, \ \gamma(b) = B \}$. On voit immédiatement que ce sous-ensemble est un sous-espace affine fermé de $C^1(I; V)$: si $0 \leqslant u \leqslant 1$ et si $\gamma_1, \gamma_2 \in \Gamma(A, B)$, on a $u.\gamma_1 + (1 - u).\gamma_2 \in \Gamma(A, B)$. Et si $\gamma_n \in \Gamma(A, B)$ converge vers $\gamma \in C^1(I; V)$ au sens de la C^1-norme, on a $\gamma(a) = \lim \gamma_n(a) = A$ et $\gamma(b) = \lim \gamma_n(b) = B$.

On peut encore observer que $\Gamma(0, 0)$ est un sous-espace vectoriel et que, si γ_0 est un élément fixé de $\Gamma(A, B)$, par exemple le segment $t \rightsquigarrow [(A - B).t + (aB - bA)]/(a - b)$, alors $\Gamma(A, B)$ se déduit de $\Gamma(0, 0)$ par la translation $\gamma \rightsquigarrow \gamma + \gamma_0$.

D'après le théorème de trace (2.4, chap. 1), la restriction de $\hat{L}$ à $\Gamma(0, 0)$ [donc à $\Gamma(A, B)$] est de classe C^1 et la différentielle de cette restriction est la restriction à $\Gamma(0, 0)$ de $D\hat{L}$. Il faut bien voir que la différentielle de la restriction de $\hat{L}$ à $\Gamma(A, B)$ est une application linéaire continue qui opère sur $\Gamma(0, 0)$: si $\gamma, \gamma + h \in \Gamma(A, B)$, leur différence $h \in \Gamma(0, 0)$ d'après l'observation ci-dessus. En sorte que

$$\hat{L}\,|_{\Gamma(A,B)}(\gamma + h) - \hat{L}\,|_{\Gamma(A,B)}(\gamma) = D\hat{L}(\gamma).h + o(h)\,.$$

En utilisant les théorèmes 1.2. et 3.5., on a donc prouvé le résultat suivant :

Théorème. 3.8. — ***Pour que $\hat{L} : \Gamma(A, B) \to \mathbf{R}$ présente un extremum en γ, il faut que***

$$\int_a^b [D_2\, L(t, \gamma(t), \gamma'(t)).h(t) + D_3\, L(\ ).h'(t)].\mathrm{d}t = 0\,,$$

pour tout $h \in \Gamma(0, 0)$, c'est-à-dire pour tout $h \in C^1(I; V)$ vérifiant $h(a) = h(b) = 0$.

L'usage qualifie une telle courbe γ du nom d'*extrémale à extrémités A et B fixes du lagrangien L*. Cette terminologie ne doit pas induire en erreur : la condition $D\hat{L}(\gamma) = 0$ est nécessaire pour l'extremum, mais elle n'est généralement pas suffisante.

Afin de déterminer les extrémales, nous allons donner une forme plus maniable au théorème précédent. Pour cela, nous allons nous borner aux courbes γ tracées sur un *espace de Hilbert V*. Cela permettra d'identifier V à son dual V^* selon $x \in V \rightsquigarrow \langle x, . \rangle \in V^*$.

Lemme de Du Bois Reymond. 3.9. — ***Soit $B : I = [a, b] \to V^*$ une fonction continue à valeurs dans le dual d'un espace de Hilbert V. Pour que $\int_a^b B(t).h'(t).\mathrm{d}t = 0$ quelle que soit $h \in C^1(I; V)$ vérifiant $h(a) = h(b) = 0$, il faut et il suffit que B soit constante.***

Preuve. — La condition est évidemment suffisante. Pour montrer qu'elle est nécessaire, observons que, si $C : I \to V^*$ est une constante, on a encore

$$\int_a^b [B(t) - C].h'(t).\mathrm{d}t = 0\,.$$

Choisissons $C = \dfrac{1}{b - a}\displaystyle\int_a^b B(t).\mathrm{d}t$, identifions V à son dual et prenons

$$h(t) = \int_a^t [B(s) - C].\mathrm{d}s\,.$$

On a bien $h(a) = h(b) = 0$; h est différentiable et $h'(t) = B(t) - C$ est continue. On doit donc avoir

$$0 = \int_a^b [\mathrm{B(t) - C}].h'(t).\mathrm{d}t = \int_a^b \| B(t) - C \|^2.\mathrm{d}t\,.$$

Ce qui exige $B(t) = C$ pour $t \in I$. □

Corollaire. 3.10. — ***Soient $A : I = [a, b] \to V^*$ et $B : I \to V^*$ deux fonctions continues. Pour que $\int_a^b [A(t).h(t) + B(t).h'(t)].\mathrm{d}t = 0$ quelle que soit $h \in C^1(I\,;\,V)$ vérifiant $h(a) = h(b) = 0$, il faut et il suffit que B soit différentiable et que $B' = A$.***

PREUVE. — La condition est suffisante, car $B'.h + B.h' = \langle B, h \rangle'$, donc

$$\int_a^b [A.h + B.h'].\mathrm{d}t = \int_a^b \langle B, h \rangle'.\mathrm{d}t = \langle B(b), h(b) \rangle - \langle B(a), h(a) \rangle = 0\,.$$

Pour montrer qu'elle est nécessaire, posons $C(t) = \int_a^t A(s).\mathrm{d}s$. On a $C'(t) = A(t)$, donc $C.h' + A.h = \langle C, h' \rangle + \langle C', h \rangle = \langle C, h \rangle'$. Intégrons en tenant compte de $h(a) = h(b) = 0$:

$$0 = \langle C(b), h(b) \rangle - \langle C(a), h(a) \rangle = \int_a^b \langle C, h \rangle'.\mathrm{d}t = \int_a^b [C(t).h'(t) + A(t).h(t)].\mathrm{d}t\,.$$

L'hypothèse entraîne donc $\int_a^b [B(t) - C(t)].h'(t)\,\mathrm{d}t = 0$ et, d'après le lemme précédent, $B(t) = C(t) +$ constante. Donc B' existe et est égal à $C' = A$. □

Théorème. 3.11. — ***Soit V un espace de Hilbert. Pour que $\gamma \in C^1(I\,;\,V)$ soit une extrémale à extrémités fixes $\gamma(a) = A$ et $\gamma(b) = B$ du lagrangien L, il faut et il suffit que $t \rightsquigarrow D_3\,L(t, \gamma(t), \gamma'(t))$ soit différentiable et que***

$$\frac{\mathrm{d}}{\mathrm{d}t} D_3\,L(t, \gamma(t), \gamma'(t)) = D_2\,L(t, \gamma(t), \gamma'(t)) \quad \textit{pour } t \in I\,. \tag{3.12}$$

PREUVE. — Conséquence immédiate du théorème 3.8. et du corollaire précédent, où l'on prend $A = D_2\,L$ et $B = D_3\,L$.

L'équation (3.12) porte le nom d'équation d'Euler-Lagrange. Elle détermine les extrémales. On observera qu'elle ne fait pas intervenir les valeurs de γ en a et b. □

Remarque. — On trouvera dans (H. Cartan, pages 294-296) une preuve du lemme de Du Bois Reymond, qui vaut non seulement pour les espaces de Hilbert, mais aussi pour les e.v. normés réels.

Le problème de l'extremum libre. 3.13. — Soit à chercher les extrema de $\hat{L} : C^1(I;\,V) \to \mathbf{R}$, sans souci d'imposer $\gamma(a) = A$ et $\gamma(b) = B$. Si $\hat{L}$ présente un extremum en γ, on a $D\hat{L}(\gamma) = 0$. La restriction de $D\hat{L}(\gamma)$ à tout sous-espace affine de $C^1(I;\,V)$ passant par γ est donc nulle. Prenons en particulier, des sous-espaces affines $\Gamma(A, B)$ et appliquons le théorème précédent. On voit que γ doit encore vérifier les équations d'Euler-Lagrange. C'est d'ailleurs évident d'après (3.5) : pour que $D\hat{L}(\gamma), h = 0$ pour *tout* $h \in C^1(I;\,V)$, il faut bien qu'il en soit ainsi pour les h vérifiant $h(a) = A$, $h(b) = B$.

Réciproquement, si γ vérifie les équations d'Euler-Lagrange, on peut se demander si si $D\hat{L}(\gamma) = 0$. Le lecteur établira que, pour tout $h \in C^1(I;\,V)$ on a

$$D\hat{L}(\gamma).h = D_3\,L(b, \gamma(b), \gamma'(b)) - D_3\,L(a, \gamma(a), \gamma'(a))\,,$$

et il conclura.

Remarque. 3.14. — Afin d'établir l'équation d'Euler-Lagrange, nous avons supposé que le lagrangien L était défini sur tout l'espace $I \oplus V \oplus V$. Mais il peut se faire que les images des courbes γ restent confinées dans une partie de l'espace V où elles sont tracées. Il se peut aussi que la vitesse $\gamma'(t)$ demeure comprise entre certaines limites. Il est donc souhaitable de restaurer le théorème 3.5. sous la seule hypothèse que L ne soit définie que sur $I \times O$, où O est un ouvert de $V \oplus V$. Il faut alors étudier $\gamma \rightsquigarrow \hat{L}(\gamma)$ lorsque $\gamma \in \Omega = \{ \gamma \in C^1(I; V) : (\gamma(t), \gamma'(t)) \in 0 \}$.

Nous allons montrer que *Ω est un ouvert de $C^1(I; V)$*. Cela aura donc un sens de chercher si $\gamma \in \Omega \rightsquigarrow \hat{L}(\gamma) \in \mathbf{R}$ est différentiable; et comme la différentiabilité est une propriété locale, tous les résultats de ce paragraphe subsisteront, en particulier les théorèmes 3.5. et 3.11.

Soit $\gamma_0 \in \Omega$. Donnons-nous $\varepsilon > 0$ et soit $\gamma \in C^1(I; V)$ tel que $\| \gamma - \gamma_0 \|_{C^1} < \varepsilon$. D'après la définition de la C^1-norme, on a donc $\| \gamma(t) - \gamma_0(t) \|_V < \varepsilon$ et $\| \gamma'(t) - \gamma_0'(t) \|_V < \varepsilon$ pour tout $t \in I$. D'après la définition de la norme sur $V \oplus V$, on a donc

$$\| (\gamma(t), \gamma'(t)) - (\gamma_0(t), \gamma_0'(t)) \|_{V \oplus V} = \| \gamma(t) - \gamma_0(t) \|_V + \| \gamma'(t) - \gamma_0'(t) \|_V < 2\,\varepsilon .$$

Il en résulte que l'ensemble $K = \{ (\gamma(t), \gamma'(t)) : t \in I \}$ reste à une distance inférieure à $2\,\varepsilon$ du compact $\{ (\gamma_0(t), \gamma_0'(t)) : t \in I \}$ de O. Comme O est un ouvert de $V \oplus V$, on peut choisir ε assez petit pour que $K \subset O$.

Si $\gamma_0 \in \Omega$, la boule de centre γ_0 et de rayon ε de $C^1(I; V)$ est donc encore dans Ω, qui est bien un ouvert.

Remarque. 3.15. — Il faudrait se garder de croire qu'un problème d'extremum, dont les données sont différentiables, possède nécessairement une solution différentiable. Voici, à titre récréatif, un contre-exemple.

Un homme, perdu en mer, sait qu'il est à une distance r d'un rivage rectiligne. Mais le brouillard est si épais qu'il ignore la direction de cette plage. Quelle est la route de longueur minimum qu'il doit suivre afin d'être sûr de toucher terre ? Mathématiquement parlant : trouver une courbe du plan d'origine O, de longueur minimum et qui coupe toute droite du plan située à distance r de l'origine.

Laissons la joie au lecteur de justifier la solution ci-contre :

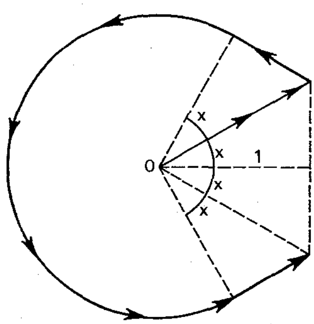


Fig. 8.

On constate que la courbe n'est pas différentiable. Prenons $r = 1$. Sa longueur, en fonction de l'angle x mesuré en radians, est $2\,\pi - 4.x + 2.\mathrm{tg}\,x + (\cos x)^{-1}$. En ajustant x de façon qu'elle soit minimum, on trouve qu'elle est égale à 7,28...

4. Nature de l'équation d'Euler-Lagrange

Revenons au lagrangien $L(t, q, v)$. Nous avons vu que si $\gamma : I \to V$ est une extrémale, alors $t \in I \rightsquigarrow D_3\, L[t, \gamma(t), \gamma'(t)]$ est dérivable. Comme c'est une fonction de t par l'intermédiaire de $\gamma(t)$ et de $\gamma'(t)$, on a envie d'appliquer le théorème de différentiation des fonctions composées. Pour que ce soit légitime, il faut introduire des hypothèses supplémentaires, en supposant que $D_3\, L$ et γ' sont différentiables. L'équation d'Euler-Lagrange s'écrit alors $D_{13}\, L(\ ) + D_{23}\, L(\ ).\gamma'(t) + D_{33}\, L(\ ).\gamma''(t) = D_2\, L(\ )$. C'est une équation différentielle du second ordre en γ.

Supposons, en outre, que $D_{33} L(\ ) \in \mathscr{L}(V^*; V)$ soit inversible dans le domaine de $(\ ) = (t, \gamma(t), \gamma'(t))$. En posant $\gamma' = v$, l'équation précédente s'écrit sous forme d'un système différentiel du premier ordre en (γ, v) :

$$(4.1) \qquad \begin{cases} \gamma'(t) = v(t) \\ v'(t) = [D_{33} L(\ )]^{-1}.[D_2 L(\ ) - D_{13} L(\ ) - D_{33} L(\ ).v(t)], \end{cases}$$

où $(\ ) = (t, \gamma(t), \gamma'(t))$.

Cela conduit à la notion suivante :

Lagrangien régulier. 4.2. — Soit O un ouvert de $V \oplus V$. On dit qu'un lagrangien $L : I \times O \to \mathbf{R}$ est régulier s'il est de classe C^2 et si $D_{33} L(t, q, v) \in \mathscr{L}(V^*; V)$ est un isomorphisme pour $t \in I, (q, v) \in 0$.

Théorème. 4.3. — ***Les extrémales $\gamma : I \to V$ d'un lagrangien régulier L sont de classe C^2.***

Etant donnés $t_0 \in I$ et $(q_0, v_0) \in \mathbf{0} \subset V \oplus V$, il existe un intervalle maximal $J \subset I$ contenant t_0 et une extrémale $\gamma : J \to V$ unique tels que $\gamma(t_0) = q_0$, $\gamma'(t_0) = v_0$.

PREUVE. — Posons $D_3 L(t, q, v) = p$, où $(q, v) \in 0$; et, t, q et p étant donnés, cherchons v. Puisque $D_{33} L(\ )$ est inversible, le théorème (5.1. chap. 3) des fonctions implicites montre que, localement, il existe une fonction G, de même classe C^1 que $D_3 L$, telle que $v = G(t, q, p)$.

Afin d'écrire l'équation d'Euler-Lagrange, exprimons que $v = \gamma'$ est la dérivée de $q = \gamma$ et que $\dfrac{\mathrm{d}}{\mathrm{d}t} D_3 L(\ ) = D_2 L(\ )$. On obtient le système différentiel du premier ordre

$$\begin{cases} \dfrac{\mathrm{d}q}{\mathrm{d}t} = G(t, q, p) ; \\[2ex] \dfrac{\mathrm{d}p}{\mathrm{d}t} = D_2 L[t, q, G(t, q, p)] . \end{cases}$$

Puisque les membres de droite sont de classe C^1, (3.4. chap. 3) montre que les solutions sont de classe C^1. En particulier, $\gamma' = \dfrac{\mathrm{d}q}{\mathrm{d}t}$ est de classe C^1 ; c'est-à-dire que γ est de classe C^2. D'autre part, $p = D_3 L(t, q, v)$ étant de classe C^1 et γ de classe C^2, il est maintenant légitime d'écrire les équations (4.1).

Quant à la seconde partie du théorème, c'est une conséquence immédiate du théorème d'existence et d'unicité (3.3. chap. 7). □

Remarque. 4.4. — Il n'est pas nécessaire que le lagrangien soit régulier pour que les conclusions du théorème 4.3. soient valides.

Supposons, par exemple, que $L(t, q, v)$ ne dépendent pas de v. Alors $D_3 L$ est identiquement nul et l'équation d'Euler-Lagrange se réduit à $D_2 L[t, \gamma(t)] = 0$, qui définit $\gamma(t)$ implicitement si $D_{22} L(\ ) \in \mathscr{L}(V^*; V)$ est inversible.

Voici un autre exemple. Supposons que $L(t, q, v)$ ne dépende pas de q. Alors $D_2 L$ est identiquement nul et l'équation d'Euler-Lagrange montre que $D_3 L[t, \gamma'(t)] =$ constante. Il se pourra que l'on puisse déterminer $\gamma'(t)$ à partir de cette équation sans pour autant que $D_{33} L(\ )$ soit inversible. C'est le cas si, par exemple, V est l'espace euclidien $\mathbf{R}^n$ et si $L(t, q, v) = \| v \|$. On obtient $L^{-1}.v = D_3 L(\ ) =$ constante; c'est-à-dire $\gamma'(t) =$ constante. Les extrémales $\gamma(t)$ sont donc des droites décrites à vitesse constante. Il fallait s'y attendre, puisque le problème d'extremum consiste à chercher les courbes de longueur $\displaystyle\int_a^b \| \gamma'(t) \| . \mathrm{d}t$ minimum. Pourtant le lagrangien n'est pas régulier, car on vérifie que $D_{33} L(v).v = 0$.

Remarque. 4.5. — Le théorème 4.3. résout le problème de Cauchy pour l'équation d'Euler-Lagrange, lorsque le lagrangien est régulier : $t_0 \in I$ étant fixé, par un point q_0 de V passe une extrémale γ et une seule de vitesse $\gamma'(t_0) = v_0$ donnée.

Ce théorème ne résout pas pour autant le problème que nous nous étions posé initialement : trouver une extrémale γ telle que $\gamma(a) = A$ et $\gamma(b) = B$ soient donnés.

Même si ce dernier problème est résolu, il reste encore à vérifier que γ réalise effectivement un extremum de $\hat{L}$. Ce sont là des problèmes qui débordent le cadre de ce livre.

Un exemple de lagrangien régulier. 4.6. — Soient N points matériels numérotés de 1 à N, de masses respectives $m_i > 0$, situés dans l'espace euclidien $\mathbf{R}^3$. Notons $q_i(t)$ et $v_i(t)$ leurs positions et leurs vitesses respectives à l'instant t.

Supposons qu'ils soient soumis à des forces dérivant d'un potentiel $U(q_1, \ldots, q_N)$ de classe C^1. Les équations de la dynamique de Newton s'écrivent $m_i \cdot \dfrac{\mathrm{d}v_i}{\mathrm{d}t} = -\dfrac{\partial U}{\partial q_i}$, où $\dfrac{\partial U}{\partial q_i} = D_i U.1$, $D_i U$ étant la différentielle partielle de U relativement à $q_i \in \mathbf{R}^3$.

On peut représenter ces N points par un point unique $q = (q_1, \ldots, q_N)$ de $\mathbf{R}^{3N}$, dont la vitesse est $v = (v_1, \ldots, v_N) \in \mathbf{R}^{3N}$. Introduisons la force vive $T = \dfrac{1}{2}\sum m_i . \| v_i \|^2$ et le lagrangien $L : \mathbf{R} \oplus \mathbf{R}^{3N} \oplus \mathbf{R}^{3N} \to \mathbf{R}$ défini par $L(q, v) = T - U$. Alors, les équations de Newton coïncident avec l'équation d'Euler-Lagrange. En effet, avec des notations évidentes,

$$\frac{\partial L}{\partial v_i} = \frac{\partial T}{\partial v_i} = m_i . v_i \quad \text{et} \quad \frac{\partial L}{\partial q_i} = -\frac{\partial U}{\partial q_i};$$

donc $0 = m_i \cdot \dfrac{\mathrm{d}q_i}{\mathrm{d}t} + \dfrac{\partial U}{\partial q_i} = \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial L}{\partial v_i}\right) - \dfrac{\partial L}{\partial q_i}$.

C'est le *principe de moindre action de Hamilton* : les mouvements du système sont donnés par les extrémales de l'intégrale d'action $\int (T - U).\mathrm{d}t$.

Le lagrangien est régulier, car $D_{33} L(\ )$ a pour matrice

$$\frac{\partial^2 L}{\partial v_i . \partial v_j} = \begin{cases} 0 & \text{si } i \neq j, \\ m_i & \text{si } i = j. \end{cases}$$

Si l'on pose $D_3 L = p$, on trouve que les composantes $p_1, \ldots, p_N$ de p sont $p_i = m_i . v_i$, qui n'est autre que l'impulsion de i-ième point matériel. La fonction G du théorème 4.3. est très simple : $v_i = p_i / m_i$. Le système du premier ordre écrit en 4.3. est ici

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = p_i / m_i, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial U}{\partial q_i}.$$

En posant $H(q, p) = \dfrac{1}{2}\sum \| p_i \|^2 / m_i + U(q)$, il s'écrit

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q_i}.$$

C'est la forme hamiltonienne des équations du mouvement (voir 1.5. chap. 7).

Supposons que le système se réduise à un point de masse $+1$ et qu'il ne soit soumis à aucune force extérieure. L'équation d'Euler-Lagrange (ou celle de Newton) montre que la vitesse de ce point reste constante. Les extrémales sont des droites décrites à vitesse constante.

Ainsi les droites ne sont pas seulement les extrémales de la longueur $\displaystyle\int_a^b \| \gamma'(t) \| . \mathrm{d}t$, elles sont également extrémales de l'action $\displaystyle\int_a^b \| \gamma'(t) \|^2 . \mathrm{d}t$. Voici une généralisation de ce résultat :

EXEMPLE. 4.7. — Soit $L : \mathbf{R} \oplus V \oplus V \to \mathbf{R}$ un lagrangien indépendant de $t \in \mathbf{R}$ et tel que, pour chaque $q \in V$, $v \in V \rightsquigarrow L(q, v)$ soit une forme quadratique non dégénérée et positive.

Il existe, autrement dit, une forme bilinéaire symétrique non dégénérée $b(q)$ telle que $L(q, v) = b(q).(v, v) > 0$ si $v \neq 0$; la fonction $q \rightsquigarrow b(q)$ étant de classe C^1.

On se propose de montrer que toute extrémale γ de $\int_a^b L.\mathrm{d}t$ est une extrémale de $\int_a^b \sqrt{L}.\mathrm{d}t$.

Commençons par montrer que, si γ est une extrémale de L, $L[\gamma(t), \gamma'(t)]$ ne dépend pas de t. En effet $\frac{\mathrm{d}L}{\mathrm{d}t} = D_2 L.\gamma' + D_3 L.\gamma'' = \left(D_2 L - \frac{\mathrm{d}}{\mathrm{d}t} D_3 L\right) \cdot \gamma' + \frac{\mathrm{d}}{\mathrm{d}t}(D_3 L.\gamma')$. Mais la parenthèse du membre de droite est nulle d'après l'équation d'Euler-Lagrange; d'autre part la règle de Leibniz entraîne $D_3 L.v = D_3[b(q).(v, v)].v = 2\, b(q).(v, v) = 2\, L$. Il reste donc $0 = \frac{\mathrm{d}}{\mathrm{d}t} L$.

Cette dernière relation et l'équation d'Euler-Lagrange pour L entraînent

$$\frac{\mathrm{d}}{\mathrm{d}t} D_3(\sqrt{L}) - D_2(\sqrt{L}) = \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{2} L^{-1/2}.D_3 L\right] - \frac{1}{2} L^{-1/2}.D_2 L =$$

$$= -\frac{1}{4} L^{-3/4} \cdot \frac{\mathrm{d}L}{\mathrm{d}t} + \frac{1}{2} L^{-1/2}\left[\frac{\mathrm{d}}{\mathrm{d}t} D_3 L - D_2 L\right] = 0.$$

C'est l'équation d'Euler-Lagrange pour le lagrangien $\sqrt{L}$; γ est donc une extrémale de $\sqrt{L}$ [voir (H. Cartan), pages 303-306].

5. Effet d'une application différentiable

Revenons à l'espace $C^1(I; V)$ des C^1-courbes $\gamma : I \to V$. Si $\varphi : V \to V$ est une application de classe C^k, $k \geqslant 1$, l'image $\varphi \circ \gamma$ d'une C^1-courbe est une C^1-courbe; donc φ induit une application $\Phi : C^1(I; V) \to C^1(I; V)$ définie par $(\Phi\gamma)(t) = \varphi[\gamma(t)]$.

Puisque $C^1(I; V)$ est un espace normé par la C^1-norme, il est légitime de se demander si Φ est différentiable.

Proposition. 5.1. — ***Si φ est de classe C^2, alors Φ est de classe C^1, et la valeur en $h \in C^1(I; V)$ de sa différentielle au point $\gamma \in C^1(I; V)$ est donnée par $D\Phi(\gamma).h : t \rightsquigarrow D\varphi[\gamma(t)].h(t)$ pour $t \in I$.***

PREUVE. — 1re *étape.* — Fixons $\gamma \in C^1(I; V)$. Si $h \in C^1(I; V)$, puisque $D\varphi$, γ et h sont de classe C^1, l'application $t \in I \rightsquigarrow D\varphi[\gamma(t)].h(t) \in V$ est de classe C^1. Cela définit donc une application L de $C^1(I; V)$ dans lui-même : $L(h) : t \in I \rightsquigarrow D\varphi[\gamma(t)].h(t)$, qui est évidemment linéaire. Montrons qu'elle est continue.

Il faut trouver une constante A telle que $\| L(h) \|_{C^1} \leqslant A.\| h \|_{C^1}$ pour tout h. Or, si l'on pose $A_1 = \sup_{t \in I} \| D\varphi[\gamma(t)] \|$, on a

$$\| D\varphi[\gamma(t)].h(t) \| \leqslant \| D\varphi[\gamma(t)] \|.\| h(t) \| \leqslant A_1.\| h \|_{C^1}.$$

D'autre part $D^2 \varphi[\gamma(t)].(\gamma'(t), .)$ est une forme linéaire continue sur V, qui dépend continûment de t. Si l'on désigne par A_2 la borne supérieure de sa norme lorsque $t \in I$, on a

$$\left\| \frac{\mathrm{d}}{\mathrm{d}t}\{ D\varphi[\gamma(t)].h(t) \} \right\| = \| D^2 \varphi[\gamma(t)].(\gamma'(t), h(t)) + D\varphi[\gamma(t)].h'(t) \| \leqslant$$

$$\leqslant \| D^2 \varphi[\gamma(t)].(\gamma'(t), .) \|.\| h(t) \| + \| D\varphi[\gamma(t)] \|.\| h'(t) \| \leqslant (A_2 + A_1).\| h \|_{C^1}.$$

En résumé, d'après la définition de la C^1-norme, on a $\| L(\gamma).h \|_{C^1} \leqslant (2\, A_1 + A_2).\| h \|_{C^1}$.

2e *étape.* — Il reste à montrer que $\| \Phi(\gamma + h) - \Phi(\gamma) - L(\gamma).h \|_{C^1}$ divisé par $\| h \|_{C^1}$, tend vers zéro avec $\| h \|_{C^1}$. D'après la définition de la C^1-norme, il faut montrer qu'à $\varepsilon > 0$ donné, on peut faire correspondre un $\delta > 0$ tel que $\| h \|_{C^1} < \delta$ entraîne $\| \Phi(\gamma + h)(t) - \Phi(\gamma)(t) -$

$L(\gamma).h(t) \| = \| \varphi[\gamma(t) + h(t)] - \varphi[\gamma(t)] - D\varphi[\gamma(t)].h(t) \| < \varepsilon. \| h \|_{C^1}$ pour $t \in I$, et une inégalité analogue pour la dérivée par rapport à t de $\Phi(\gamma + h) - \Phi(\gamma) - L(\gamma).h$.

Bornons-nous à montrer la première; la seconde se démontrant de la même manière. Donnons-nous γ, $h \in C^1(I; V)$. Puisque γ et h sont continues, l'ensemble $K = \{ \gamma(t) + s.h(t) : t \in I, 0 \leqslant s \leqslant 1 \}$ est un compact de V, qui contient les images de γ et $\gamma + h$. Puisque $D^2 \varphi$ est continue $\| D^2 \varphi[\gamma(t) + s.h(t)] \|$ est borné sur K par un nombre A. Appliquons la formule de Taylor (3.4. chap. 4) :

$$\| \varphi[\gamma(t) + h(t)] - \varphi[\gamma(t)] - D\varphi[\gamma(t)].h(t) \| \leqslant A. \| h(t) \|^2 \leqslant A.(\| h \|_{C^1})^2 .$$

Si $\delta < \varepsilon. A^{-1}$, on obtient le résultat promis. □

Corollaire. 5.2. — *Si φ est un difféomorphisme de classe C^2, alors Φ est un difféomorphisme de classe C^1.*

Maintenant, donnons-nous un lagrangien $L : I \oplus V \oplus V \to \mathbf{R}$ de classe C^1 et revenons à la recherche des extrema de $\hat{L}(\gamma) = \displaystyle\int_a^b L[t, \gamma(t), \gamma'(t)].\mathrm{d}t$.

Théorème. 5.3. — *La propriété d'une courbe $\gamma \in C^1(I; V)$ d'être extrémale de L ne dépend pas du système de coordonnées choisi sur l'espace V où elle est tracée.*

Preuve. — Identifions V à $\mathbf{R}^n$. Soit (U, φ) une carte dont le domaine U contient l'image de γ; φ est un difféomorphisme de U sur $\varphi(U)$. D'après 5.2., φ induit un difféomorphisme Φ de l'ensemble des courbes de $C^1(I; V)$, dont l'image est dans U, dans l'ensemble des courbes de $C^1(I; \mathbf{R}^n)$ dont l'image est dans $\varphi(U)$. D'après le théorème de différentiation des applications composées, en écrivant $\hat{L} = (\hat{L} \circ \Phi^{-1}) \circ \Phi$, on a $D\hat{L}(\gamma) = D[\hat{L} \circ \Phi^{-1}](\Phi\gamma) \circ D\Phi(\gamma)$. Puisque $D\Phi(\gamma)$ est un isomorphisme, cela montre que $D\hat{L}(\gamma) = 0$ entraîne $D[\hat{L}_0\ \Phi^{-1}].(\Phi\gamma) = 0$. Autrement dit, l'image $\Phi\gamma$ de l'extrémale γ dans la carte est extrémale de l'expression $\hat{L}_0\ \Phi^{-1}$ de $\hat{L}$ de cette carte. □

Ce théorème est fort important. En mécanique, par exemple, il permet d'écrire les équations du mouvement dans un système de coordonnées quelconques, en utilisant le principe de moindre action de Hamilton. Bornons-nous à un exemple simple :

Exemple : mouvement à accélération centrale. 5.4. — Un point matériel de masse $m > 0$, situé dans un plan $\mathbf{R}^2$, est soumis à un potentiel U qui ne dépend que de la distance ρ à l'origine.

Prenons des coordonnées polaires (ρ, θ) d'origine O. Le lagrangien $T - U$, introduit en 4.6., s'écrit

$$L(\rho, \theta, \rho', \theta') = \frac{1}{2} m[\rho'^2 + \rho^2.\theta'^2] - U(\rho) .$$

L'équation d'Euler-Lagrange se décompose en deux équations scalaires :

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \rho'}\right) = \frac{\partial L}{\partial \rho}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \theta'}\right) = \frac{\partial L}{\partial \theta} .$$

La première s'écrit $m.\rho'' = m.\rho.\theta'^2 - \dfrac{\mathrm{d}U}{\mathrm{d}\rho}$. Comme θ ne figure pas dans L, la seconde s'écrit $\dfrac{\partial L}{\partial \theta'}$ = constante; c'est-à-dire $m.\rho^2.\theta'$ = constante : c'est la loi de conservation du moment cinétique par rapport à O, appelée aussi « loi des aires ».

Nous verrons au paragraphe 7 comment de telles lois de conservation s'inscrivent dans un cadre plus général.

6. Invariance d'un lagrangien

Définition. 6.1. — Supposons que le lagrangien $L : I \oplus V \oplus V$ ne dépende pas de $t \in I$. C'est donc une fonction de classe C^1 définie sur $V \oplus V$. A ce titre, il est commode d'interpréter $V \oplus V$ comme l'ensemble $TV = \{ q \in V, v \in V \}$ des couples formés par un point q de V et par un vecteur v tangent à V en ce point (on considère V comme une sous-variété de V). Le lagrangien L est donc une fonction C^1 sur TV. Si $\varphi : V \to V$ est un C^k-difféomorphisme, il définit (voir 2.6. chap. 1) une application tangente $T\varphi : TV \to TV$ selon $(q, v) \rightsquigarrow (\varphi(q), D\varphi(q).v)$, qui est un C^{k-1}-difféomorphisme.

Nous dirons que L est invariant par φ si $L = L_0 \, T\varphi$; autrement dit si

$$L(q, v) = L[\varphi(q), D\varphi(q).v]$$

pour $q \in V$, $v \in V$.

Exemple. — Supposons que $V = \mathbf{R}^3 = \{ (q_1, q_2, q_3) \}$. Donnons-nous une fonction $U : \mathbf{R}^3 \to \mathbf{R}$ de classe C^1, ne dépendant pas de q_1, et considérons le lagrangien

$$L(q_1, q_2, q_3, v_1, v_2, v_3) = \frac{1}{2}[v_1^2 + v_2^2 + v_3^2] - U(q_2, q_3).$$

Alors la translation $(q_1, q_2, q_3) \rightsquigarrow (q_1 + s, q_2, q_3)$ laisse L invariant.

Théorème. 6.2. — ***Si $\gamma : I = [a, b] \to V$ est une extrémale de L, d'extrémités $\gamma(a) = A$ et $\gamma(b) = B$, et si le difféomorphisme $\varphi : V \to V$ laisse L invariant, alors $\varphi \circ \gamma$ est encore une extrémale de L, d'extrémités $\varphi(A)$, $\varphi(B)$.***

Preuve. — Avec les notations du paragraphe précédent, l'invariance se traduit par $\hat{L}(\Phi\gamma) = \hat{L}(\gamma)$. Il en résulte $D\hat{L}(\Phi\gamma) \circ D\Phi(\gamma) = D\hat{L}(\gamma)$. Si γ est une extrémale, $D\hat{L}(\gamma) = 0$. Puisque $D\Phi(\gamma)$ est inversible, il en résulte $D\hat{L}(\Phi\gamma) = 0$. Cela prouve que $\varphi \circ \gamma$ est une extrémale.

Remarque. — Beaucoup d'auteurs tiennent ce résultat pour évident. En voici probablement la raison. Supposons que γ réalise le minimum de $\hat{L}(\gamma)$. Alors $\varphi \circ \gamma$ le réalise aussi, puisque $\hat{L}(\Phi\gamma) = \hat{L}(\gamma)$; $\varphi \circ \gamma$ est donc une extrémale. Malheureusement la condition $D\hat{L}(\gamma) = 0$ n'est pas suffisante pour garantir l'extrémalité ; elle n'est que nécessaire.

Voici une application spectaculaire du théorème précédent, qui montre que des considérations d'invariance peuvent épargner bien des calculs.

Le demi-plan de Poincaré. 6.3. — L'espace V est le plan $\mathbf{R}^2 = \{ (x, y) \}$. Le lagrangien est la fonction $L(x, y, \dot{x}, \dot{y}) = (\dot{x}^2 + \dot{y}^2)/y^2$ définie sur le produit du demi-plan

$$P = \{ (x, y) : y > 0 \} \qquad \text{par} \qquad \mathbf{R}^2 = \{ (\dot{x}, \dot{y}) \}.$$

On se propose de chercher les extrémales γ dont l'image est dans P.

1re *étape.* — On voit immédiatement que le lagrangien est régulier. D'après le théorème 4.3., il en résulte qu'étant donnés $q_0 \in P$ et $v_0 \in \mathbf{R}^2$ il existe une extrémale γ unique telle que $\gamma(0) = q_0$, $\gamma'(0) = v_0$.

2e *étape.* — Complexifions P en associant au point (x, y) le nombre complexe $z = x + iy$. Avec un abus de notation évident, on peut écrire $L(z, \dot{z}) = |\dot{z}|^2/[\mathrm{Im}\,(z)]^2$, où Im = partie imaginaire.

Un calcul sans malice montre alors que L est invariant par la symétrie $s : (x, y) \rightsquigarrow (-x, y)$ et par les homographies $h : z \rightsquigarrow \dfrac{az + b}{cz + d}$, où les réels a, b, c, d vérifient $ad - bc = 1$.

3e *étape.* — Rappelons que ces homographies laissent globalement invariant P et qu'elles transforment la demi-droite $x = 0$, $y > 0$ en une demi-droite $x = x_0$, $y > 0$ ou en un demi-cercle centré sur l'axe des x.

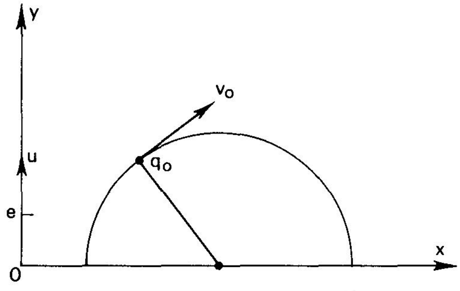
Rappelons enfin qu'étant donné $(q_0, v_0) \in TP$, où $L(q_0, v_0) = 1$, il existe une homographie h, du type ci-dessus, qui envoie le point $e = (1, 0)$ en q_0 et le vecteur $u = (0, 1)$ sur v_0 : $h(e) = q_0$, $Dh(e).u = v_0$.

4^e^ *étape.* — Cherchons l'extrémale γ telle que $\gamma(0) = e$, $\gamma'(0) = u$ [elle est unique d'après l'étape 1]. D'après la 2^e^ étape, $s \circ \gamma$ est encore une extrémale. Comme elle vérifie les mêmes conditions initiales $s \circ \gamma(0) = e$, $(s \circ \gamma)'(0) = s(\gamma'(0)) = s(u) = u$, que γ, elle coïncide avec γ. Puisque γ est invariante par la symétrie s, son image est donc la demi-droite $x = 0, y > 0$.

5^e^ *étape.* — Donnons-nous $q_0 \in P$ et $v_0 \in \mathbf{R}^2$ tels que $L(q_0, v_0) = 1$. D'après l'étape 1, il existe une extrémale π unique vérifiant $\pi(0) = q_0$, $\pi'(0) = v_0$.

D'autre part (3^e^ étape) il existe une homographie h telle que $h(e) = q_0$, $Dh(e).u = v_0$. D'après le théorème 6.2., si γ est l'extrémale de la 4^e^ étape, $h \circ \gamma$ est encore une extrémale. Comme elle vérifie les mêmes conditions initiales $(h \circ \gamma)(0) = q_0$ et $(h \circ \gamma)'(0) = Dh(e).u = v_0$ que π, on a donc $\pi = h \circ \gamma$ [unicité].

D'après l'étape 3, il en résulte que toute extrémale π a pour image une demi-droite $x = x_0$, $y > 0$ ou un demi-cercle de P centré sur l'axe des x.



Ceci posé, appelons longueur au sens de Lobatchevskii d'une courbe

$$l : t \in [a, b] \rightsquigarrow (x(t), y(t)) \in P$$

de classe C^1, le nombre $\displaystyle\int_a^b \frac{\sqrt{x'^2 + y'^2}}{y} \cdot dt$. Proposons-nous de chercher les courbes l de « longueur » minimum joignant deux points donnés de P. Cela amène à chercher les extrémales du lagrangien $\sqrt{L} = \dfrac{\sqrt{\dot{x}^2 + \dot{y}^2}}{y}$. En combinant les résultats précédents et celui de l'exemple 4.7., on voit que ces courbes ont pour image les demi-droites $x = x_0$, $y > 0$ ou les demi-cercles de P centrés sur l'axe des x.

Si l'on appelle distance (au sens de la géométrie de Lobatchevskii) de deux points A et B de P la longueur de l'extrémale qui les joint, ces demi-droites et ces demi-cercles jouent le rôle des droites de la géométrie euclidienne. On voit sans peine que la symétrie s et les homographies h jouent le rôle des déplacements de la géométrie euclidienne.

P s'appelle demi-plan de Poincaré, du nom du mathématicien qui a introduit le modèle que nous venons de décrire pour la géométrie de Lobatchevskii.

7. Le théorème d'Emmy Noether

Nous allons montrer comment l'invariance d'un lagrangien par un groupe à un paramètre de difféomorphismes produit des fonctions qui demeurent constantes le long des extrémales.

Définition 7.1. — On dit qu'un lagrangien $L : I \oplus V \oplus V \to \mathbf{R}$, indépendant de $t \in I$, est invariant par un groupe à un paramètre s de difféomorphismes φ_s de V si L est invariant par φ_s pour tout s.

Théorème de Noether. 7.2. — ***Soit L un lagrangien indépendant de $t \in I$ et invariant par un groupe à un paramètre s de difféomorphismes φ_s de V. Alors la fonction $\mathfrak{I} : V \oplus V \to \mathbf{R}$, définie par $(q, v) \rightsquigarrow D_3\, L(q, v) \cdot \left(\frac{\partial}{\partial s} \varphi_s(q)\Big|_{s=0}\right) \in \mathbf{R}$, où $D_3\, L$ est la différentielle partielle de L par rapport à v, est constante le long de chaque extrémale de L. On dit que c'est une intégrale première de l'équation d'Euler-Lagrange.***

PREUVE. — Soit $\gamma : I \to V$ une extrémale. Par hypothèse $L\left[\varphi_s(\gamma(t)), \frac{\mathrm{d}}{\mathrm{d}t}(\varphi_s \circ \gamma)(t)\right]$ ne dépend pas de s. Ecrivons que sa dérivée par rapport à s est nulle :

$$D_2\, L[\;] \cdot \frac{\partial}{\partial s}[(\varphi_s \circ \gamma)(t)] + D_3\, L[\;] \cdot \frac{\partial}{\partial s}\frac{\mathrm{d}}{\mathrm{d}t}(\varphi_s \circ \gamma)(t) = 0\,.$$

Mais, d'après le théorème de Schwarz, $\frac{\partial}{\partial s}\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial s}$. D'autre part l'équation d'Euler-Lagrange s'écrit $D_2\, L = \frac{\mathrm{d}}{\mathrm{d}t} D_3\, L$. On en déduit

$$\frac{\mathrm{d}}{\mathrm{d}t} D_3\, L[\;] \cdot \frac{\partial}{\partial s}[(\varphi_s \circ \gamma)(t)] + D_3\, L[\;] \cdot \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial s}(\varphi_s \circ \gamma)(t) = 0\,.$$

C'est-à-dire

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{ D_3\, L[\;] \cdot \frac{\partial}{\partial s}[(\varphi_s \circ \gamma)(t)] \right\} = 0\,.$$

Faisons $s = 0$; avec les notations du théorème on a

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{ D_3\, L[\gamma(t), \gamma'(t)] \cdot \frac{\partial}{\partial s}\varphi_s \circ \gamma(t)\Big|_{s=0} \right\} = 0\,. \qquad \square$$

Remarque. — $\frac{\partial}{\partial s}(\varphi_s \circ \gamma)(t)\Big|_{s=0}$ s'interprète comme le vecteur-vitesse pour $s = 0$ du point $\gamma(t)$ décrivant son orbite sous l'action du groupe φ_s :

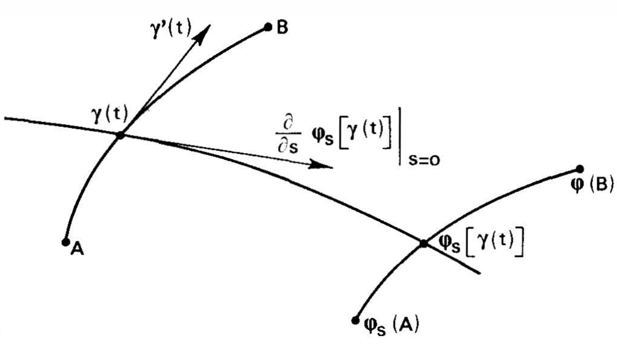


Fig. 10.

Premier exemple : invariance par translation. 7.3. — Supposons que L soit invariant par un groupe à un paramètre s de translations : $\varphi_s(q) = q + s.a$, où $a \in V$. Alors $\varphi_s[\gamma(t)] = \gamma(t) + s.a$ et $\frac{\partial}{\partial s}\varphi_s[\gamma(t)]\Big|_{s=0} = a$.

L'intégrale première correspondante est donc $D_3\, L[\gamma(t), \gamma'(t)].a$.

Supposons, de façon encore plus particulière, que L soit invariant par *toute* translation. On peut donc choisir a arbitrairement et le théorème de Noether signifie que la forme linéaire $D_3\, L[\gamma(t), \gamma'(t)]$ ne dépend pas de t. Ce fait est facile à établir directement. Puisque L est invariant par toutes les translations, $L(q, v)$ ne dépend pas de q. Donc $D_2\, L = 0$, et l'équation d'Euler-Lagrange $\frac{\mathrm{d}}{\mathrm{d}t} D_3\, L = D_2\, L$ entraîne $D_3\, L =$ constante.

EXEMPLE (CONSERVATION DE L'IMPULSION). — Reprenons l'exemple 4.6. Si le lagrangien $L = T - U$ est invariant par toutes les translations, U se réduit à une constante.

D'après ce qui précède $D_3\, L = D_3\, T$ demeure constant. Cela s'écrit $\sum m_i . v_i =$ constante. L'impulsion $\sum m_i . v_i$ se conserve donc au cours du temps.

SECOND EXEMPLE : INVARIANCE PAR ROTATION. 7.4. — Soit V l'espace euclidien $\mathbf{R}^3$ orienté canoniquement. Ce qui va permettre d'utiliser le produit vectoriel $\wedge$.

Exprimons que le lagrangien $L(q, v)$ est invariant par rotation autour d'un axe passant par l'origine $o \in \mathbf{R}^3$. Si ω est un vecteur porté par cet axe, on s'assure sans peine que $\left.\frac{\partial}{\partial s} \varphi_s(q)\right|_{s=0} = \omega \wedge q$. Le théorème de Noether s'écrit donc $D_3\, L[\gamma(t), \gamma'(t)].(\omega \wedge q) =$ constante.

En identifiant une forme linéaire à un vecteur par l'isomorphisme $x \in V \rightsquigarrow \langle v, . \rangle \in V^*$. cela s'écrit encore $\langle D_3\, L, \omega \wedge q \rangle =$ constant. Ou encore, compte tenu des propriétés du produit mixte, $\langle q \wedge D_3\, L[\ ], \omega \rangle =$ constant.

EXEMPLE (CONSERVATION DU MOMENT CINÉTIQUE). — Reprenons l'exemple 4.6. L'invariance par rotation se traduit par $\langle \sum m_i .(q_i \wedge v_i), \omega \rangle =$ constante.

Si, en particulier, le potentiel U ne dépend que de la distance à l'origine, $L = T - U$ est invariant par *toutes* les rotations autour de O. On peut donc choisir ω arbitrairement et l'on trouve que le moment cinétique $\sum m_i .(q_i \wedge v_i)$ par rapport à O est constant.

Généralisation du théorème de Noether au cas où L dépend de $t \in I$. 7.5. — Supposons que $L : I \oplus V \oplus V \to \mathbf{R}$ dépende de $t \in I$.

Nous allons nous ramener au cas où il n'en dépend pas un artifice.

Introduisons l'espace-temps $\overline{V} = \mathbf{R} \oplus V = \{(t, q) : t \in \mathbf{R}, q \in V\}$, et notons $\pi_2 : \overline{V} \to V$ la seconde projection. Si $\gamma : I \to V$ est une courbe de V, elle se « relève » en une courbe $\overline{\gamma} : t \in I \rightsquigarrow (t, \gamma(t)) \in \overline{V}$ et $\pi_2 \circ \overline{\gamma} = \gamma$.

Définissons un lagrangien L_1 sur l'espace tangent $T\overline{V} = \overline{V} \oplus \overline{V} = \{(t, q, u, v)\}$ par

$$L_1(t, q, u, v) = L(t, q, v/u).u \quad \text{pour} \quad u \neq 0\,. \tag{7.6}$$

Si $\overline{\gamma} : s \in I \rightsquigarrow (t(s), \gamma(s)) \in \overline{V}$ est une C^1-courbe de $\overline{V}$, on a

$$\begin{aligned} \hat{L}_1(\overline{\gamma}) &= \int_a^b L_1\left[t(s), \gamma(s), \frac{\mathrm{d}t}{\mathrm{d}s}, \frac{\mathrm{d}\gamma}{\mathrm{d}s}\right]\cdot \mathrm{d}s = \int_a^b L\left[t(s), \gamma(s), \frac{\mathrm{d}\gamma}{\mathrm{d}s}\Big/\frac{\mathrm{d}t}{\mathrm{d}s}\right] \cdot \frac{\mathrm{d}t}{\mathrm{d}s} \cdot \mathrm{d}s \\ &= \int_a^b L\left[t(s), \gamma(s), \frac{\mathrm{d}\gamma}{\mathrm{d}t}\right] \cdot \mathrm{d}t = \hat{L}(\pi_2 \circ \overline{\gamma})\,. \end{aligned}$$

Il en résulte que $D\hat{L}_1(\overline{\gamma}) = D\hat{L}(\pi_2\, \overline{\gamma}) \circ \pi_2 = D\hat{L}(\gamma) \circ \pi_2$. En sorte que si $\overline{\gamma}$ est une extrémale de L_1, sa projection $\gamma = \pi_2 \circ \overline{\gamma}$ est une extrémale de L et, qu'inversement, le relèvement $\overline{\gamma}$ d'une extrémale de γ est une extrémale de L_1 [à condition de prendre pour paramètre $s = t$].

Si φ_r est un sous-groupe à un paramètre de difféomorphismes de l'espace-temps $\overline{V}$, laissant L_1 invariant, il donne lieu, d'après le théorème de Noether, à une intégrale première $\mathfrak{I}_1(t, q, u, v)$. Si γ est une extrémale de L, son relèvement $\overline{\gamma}$, paramétré par $s = t$, est une extrémale de L_1.

Donc $\mathfrak{I}_1\left(t(s), \gamma(s), \dfrac{\mathrm{d}t}{\mathrm{d}s}, \dfrac{\mathrm{d}\gamma}{\mathrm{d}s}\right) = \mathfrak{I}_1\left(t, \gamma(t), 1, \dfrac{\mathrm{d}\gamma}{\mathrm{d}t}\right)$ ne dépend pas de $s = t$: c'est une intégrale première de L.

EXEMPLE (CONSERVATION DE L'ÉNERGIE). — Si L ne dépend pas de $t \in I$, alors L_1 est invariant par le groupe à un paramètre de translations temporelles $\varphi_s : (t, q) \rightsquigarrow (t + s, q)$. On a donc $\dfrac{\partial}{\partial s} \varphi_s(t, \gamma(t))\Big|_{s=0} = (1, 0) \in \mathbf{R} \oplus V$.

Le théorème de Noether entraîne donc $D_3\, L_1(\ ).(1, 0) =$ constante, où D_3 désigne la différentielle partielle relative à $(u, v) \in \overline{V}$. Cela s'écrit encore $\dfrac{\partial}{\partial u} L_1(\ ) =$ constante, ou, compte tenu de l'expression définissant L_1, $L(t, \gamma(t), \gamma'(t)) - D_3\, L((t, \gamma(t), \gamma'(t)).\gamma'(t) =$ constante. C'est ce qu'on appelle l'intégrale première de l'énergie. La raison en est que, si l'on reprend l'exemple 4. 6. et le lagrangien indépendant du temps $L(q, v) = T - U$, l'intégrale première correspondante est (au signe prés) $-L + D_3\, L.v = T + U$. C'est l'énergie totale.

Bibliographie. — On ne saurait trop recommander (L. C. Young). Outre qu'on est assuré d'y trouver une perle d'humour au détour de chaque paragraphe, l'étude détaillée du calcul des variations est complétée par celle de son renouveau moderne : le contrôle optimal. Le lecteur qui aime la géométrie et les promenades dans les jardins à la française sera comblé par le dernier chapitre de (P. Malliavin).

Appendice A

ESPACES DE BANACH ET APPLICATIONS MULTILINÉAIRES

Dans toute la suite le corps K sur lequel sont construits les espaces vectoriels (e.v., en abrégé) est celui des réels ou des complexes.

1. Espaces de Banach

Norme. 1.1. — On appelle norme sur un e.v. E une fonction $\| \ \| : E \to \mathbf{R}$ telle que $\| x \| \geqslant 0$, $\| x \| = 0$ si et seulement si $x = 0$, $\| x + y \| \leqslant \| x \| + \| y \|$, $\| k.x \| = | k |.\| x \|$ pour tout $k \in K$ et tous $x, y \in E$. Ces propriétés entraînent $| \| x \| - \| y \| | \leqslant \| x - y \|$.

Un e.v. doté d'une norme est dit espace normé.

EXEMPLES. 1.2.

a) On vérifie que chacune des applications suivantes est une norme dans $\mathbf{R}^n$ (ou $\mathbf{C}^n$) :

$$(x_1, \ldots, x_n) \rightsquigarrow \begin{cases} (| x_1 |^2 + \cdots + | x_n |^2)^{1/2}\,, \\ | x_1 | + \cdots + | x_n |\,, \\ \sup \{ | x_1 |, \ldots, | x_n | \}\,. \end{cases}$$

b) Soit $C_b^0(x)$ l'e.v. des fonctions définies sur un espace topologique X, à valeurs dans K. qui sont continues et bornées. Alors

$$\| f \|_{C^0} = \sup_{x \in X} | f(x) |$$

est une norme, appelée norme de la convergence uniforme, ou encore norme C^0.

c) Soit $C^1([a, b])$ l'e.v. des fonctions f définies sur un segment $[a, b]$, $a < b$, à valeurs réelles et possédant en tout point une dérivée continue. Alors

$$\| f \|_{C^1} = \sup_{x \in [a,b]} | f(x) | + \sup_{x \in [a,b]} | f'(x) |$$

est une norme appelée norme C^1. On observera que $\| f \|_{C^1} = \| f \|_{C^0} + \| f' \|_{C^0}$.

Distance déduite d'une norme. 1.3. — Soit E un e.v. normé. Pour $x, y \in E$ posons $d(x, y) = \| x - y \|$. On vérifie que d est une distance sur E, qui devient ainsi un espace métrique, donc un espace topologique.

Espaces de Banach. 1.4. — C'est un e.v. normé, qui est complet pour la distance déduite de sa norme.

EXEMPLES. 1.5.

a) $\mathbf{R}^n$ (ou $\mathbf{C}^n$) est un espace de Banach pour chacune des trois normes ci-dessus.

b) On sait que $C_b^0(X)$ est un espace de Banach pour la norme de la convergence uniforme

c) Montrons que l'espace normé $C^1([a, b])$ du 1.2. est un espace de Banach. Soit (f_n) une suite de Cauchy. Puisque $\| f \|_{C^1} = \| f \|_{C^0} + \| f' \|_{C^0}$, chacune des suites (f_n) et (f_n') est une suite de Cauchy de $C^0([a, b])$; comme cet espace est complet, f_n (resp. f_n') converge uniformément vers une fonction f (resp. g) de $C^0([a, b])$. Tenant compte de

$$f_n(x) - f_n(a) = \int_a^x f_n'(t).\mathrm{d}t , \qquad a \leqslant x \leqslant b ,$$

on a

$$\begin{aligned}\left|f(x) - f(a) - \int_a^x g(t).\mathrm{d}t\right| &= \left|f(x) - f_n(x) + f_n(a) - f(a) + \int_a^x [f_n'(t) - g(t)].\mathrm{d}t\right| \\ &\leqslant |f(x) - f_n(x)| + |f_n(a) - f(a)| + \int_a^x |f_n'(t) - g(t)|\,\mathrm{d}t \\ &\leqslant 2.\| f - f_n \|_{C^0} + (b - a).\| f_n' - g \|_{C^0} ,\end{aligned}$$

qui tend vers zéro si $n \to \infty$. Donc f' existe et $f' = g$, car $f'(x) = f(a) + \int_a^x g(t).\mathrm{d}t$.

Ainsi (f_n) converge vers f pour la norme C^1.

Somme directe d'espaces normés. 1.6.

Soient $E_1, \dots, E_n$ des espaces normés sur le même corps, de normes respectives $\| \ \|_1, \dots$. On vérifie que

a) leur produit cartésien devient un e.v. noté $E_1 \oplus \cdots \oplus E_n$ et appelé somme directe de $E_1, \dots, E_n$, pour les opérations

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n) ,$$
$$k.(x_1, \dots, x_n) = (k.x_1, \dots, k.x_n) , \qquad k \in K ;$$

b) $\| (x_1, \dots, x_n) \| = \| x_1 \|_1 + \cdots + \| x_n \|_n$ est une norme sur $E_1 \oplus \cdots \oplus E_n$;

c) $E_1 \oplus \cdots \oplus E_n$ est un espace de Banach si chaque espace $E_1, \dots, E_n$ est un espace de Banach.

2. Applications linéaires continues

Soient E et F deux espaces normés sur le même corps et $f : E \to F$ une application linéaire.

Théorème. 2.1. — ***Les conditions suivantes sont équivalentes :***

a) ***f est continue en tout point ;***

b) ***f est continue à l'origine ;***

c) $\sup_{\|x\| \leqslant 1} \| f(x) \| < \infty$ ***(autrement dit : l'image de la boule unité de E est bornée) ;***

d) $\sup_{\|x\| = 1} \| f(x) \| < \infty$ ***(autrement dit : l'image de la sphère unité de E est bornée) ;***

e) ***il existe une constante M telle que $\| f(x) \| \leqslant M.\| x \|$ pour tout $x \in E$.***

PREUVE. — Evidemment $a \Rightarrow b$.

Montrons que $b \Rightarrow c$: il existe $\delta > 0$ tel que $\| x \| \leqslant \delta \Rightarrow \| f(x) \| \leqslant 1$. Alors $\| x \| \leqslant 1 \Rightarrow \| \delta.x \| \leqslant \delta \Rightarrow \| f(\delta.x) \| \leqslant 1 \Rightarrow \| f(x) \| \leqslant \delta^{-1}$.

Evidemment $c \Rightarrow d$.

Montrons que $d \Rightarrow e$: soit $M = \sup_{\|x\| = 1} \| f(x) \|$.

Si $x \neq 0$, $x/\| x \|$ est de norme 1. donc $\| f(x)/\| x \| \| \leqslant M$ et $\| f(x) \| \leqslant M.\| x \|$; inégalité encore vérifiée si $x = 0$.

Enfin $e \Rightarrow a$ résulte de $\| f(x) - f(y) \| = \| f(x - y) \| \leqslant M.\| x - y \|$. □

Norme d'une application linéaire continue. 2.2. — *Le nombre* $\sup_{\|x\|=1} \|f(x)\|$ s'appelle norme de f et se note $\|f\|$. La preuve précédente de $d \Rightarrow e$ montre que $\|f(x)\| \leqslant \|f\|.\|x\|$ pour tout $x \in E$.

Montrons que les applications linéaires continues f de E dans F forment un e.v., noté $\mathscr{L}(E; F)$, et que $f \rightsquigarrow \|f\|$ est une norme (ce qui justifiera son appellation). Soient $f, g \in \mathscr{L}(E; F)$. Si $x \in E$, $\|x\| = 1$, on a $\|(f+g)(x)\| = \|f(x) + g(x)\| \leqslant \|f(x)\| + \|g(x)\| \leqslant \|f\| + \|g\|$; donc $f + g$ est continue et $\|f + g\| \leqslant \|f\| + \|g\|$. D'autre part, si $k \in K$, $\|(k.f)(x)\| = \|k.f(x)\| = |k|.\|f(x)\|$, d'où

$$\sup_{\|x\|=1} \|(k.f)(x)\| = |k|.\sup_{\|x\|=1} \|f(x)\|,$$

c'est-à-dire $\|k.f\| = |k|.\|f\|$. Enfin $\|f\| = 0 \Rightarrow f(x) = 0$ pour tout $x \Rightarrow f = 0$.

L'espace de Banach $\mathscr{L}$(E ; F). 2.3. — ***Si F est un espace de Banach, alors $\mathscr{L}$(E ; F) est complet.***

PREUVE. — Soit (f_n) une suite de Cauchy de $\mathscr{L}(E; F)$. Fixons $x \in E$. L'inégalité $\|f_p(x) - f_q(x)\| \leqslant \|f_p - f_q\|.\|x\|$ montre que $(f_n(x))$ est une suite de Cauchy dans F. Elle a donc une limite que l'on notera $f(x)$.

Montrons que f est linéaire :

$$f(x+y) = \lim_{n=\infty} f_n(x+y) = \lim_{n=\infty} [f_n(x) + f_n(y)] = \lim_{n=\infty} f_n(x) + \lim_{n=\infty} f_n(y) = f(x) + f(y).$$

De même $f(k.x) = k.f(x)$, $k \in K$.

Montrons que f est continue. A tout $\varepsilon > 0$ correspond N tel que $p, q \geqslant N \Rightarrow \|f_p - f_q\| \leqslant \varepsilon$, c'est-à-dire $\|(f_p - f_q)(x)\| = \|f_p(x) - f_q(x)\| \leqslant \varepsilon$ pour $\|x\| = 1$. Faisons tendre q vers ∞, on obtient $\|f_p(x) - f(x)\| \leqslant \varepsilon$ pour $\|x\| = 1$. On en déduit d'abord $\|f(x)\| \leqslant \|f(x) - f_p(x)\| + \|f_p(x)\| \leqslant \varepsilon + \|f_p\|$ pour $\|x\| = 1$; donc f est continue. On en déduit ensuite $\|f - f_p\| \leqslant \varepsilon$ pour $p \geqslant N$; donc (f_p) converge vers f. □

Norme d'un produit d'applications linéaires continues. 2.4. — Soient E, F et G trois espaces normés sur le même corps et $u : E \to F$, $v : F \to G$ des applications linéaires continues.

Leur composée $v_0 u : E \to G$ est évidemment linéaire et continue. Soit $x \in E$, $\|x\| = 1$, on a $\|v_0 u(x)\| = \|v[u(x)]\| \leqslant \|v\|.\|u(x)\| \leqslant \|v\|.\|u\|$. Par suite $\|v_0 u\| \leqslant \|v\|.\|u\|$.

Isomorphisme d'espaces normés. 2.5. — Soient E et F deux espaces normés sur le même corps. Une application $f : E \to F$ est appelée un isomorphisme (d'espaces normés) si $f \in \mathscr{L}(E; F)$, f est une bijection, et si la bijection inverse f^{-1} (qui est ipso facto linéaire) est continue.

Théorème. — ***Soit f : E → F une application linéaire* surjective. *Les conditions suivantes sont équivalentes :***

a) ***f est un isomorphisme ;***

b) ***il existe deux nombres $m > 0$ et $M > 0$ tels que $m.\|x\| \leqslant \|f(x)\| \leqslant M.\|x\|$ pour tout $x \in E$.***

PREUVE. — Montrons que $a \Rightarrow b$. Puisque f et f^{-1} sont continues, il existe deux constantes $A > 0$ et $B > 0$ telles que $\|f(x)\| \leqslant A.\|x\|$ pour tout $x \in E$ et $\|f^{-1}(y)\| \leqslant B.\|y\|$ pour tout $y \in F$. Si $x \in E$, posons $y = f(x)$, en sorte que $x = f^{-1}(y)$. Alors $\|x\| \leqslant B.\|f(x)\|$ et l'on a $B^{-1}.\|x\| \leqslant \|f(x)\| \leqslant A.\|x\|$.

Montrons que $b \Rightarrow a$. L'inégalité $\|f(x)\| \leqslant M.\|x\|$ prouve que f est continue. L'inégalité $m.\|x\| \leqslant \|f(x)\|$, où $m > 0$, montre que $\operatorname{Ker}(f) = \{0\}$; f est donc injective et, puisqu'elle est supposée surjective, c'est une bijection. Soit $y \in F$. Posons $x = f^{-1}(y)$; on a $y = f(x)$ et $m.\|x\| \leqslant \|f(x)\|$ s'écrit $\|f^{-1}(y)\| \leqslant m^{-1}.\|y\|$. Donc f^{-1} est continue.

Normes équivalentes. 2.6. — Deux normes sur un e.v. E sont dites équivalentes si elles définissent la même topologie.

Théorème. — *Les conditions suivantes sont équivalentes :*

a) ***les deux normes $\| \ \|_1$ et $\| \ \|_2$ sont équivalentes ;***

b) ***il existe des nombres $m > 0$ et $M > 0$ tels que $m.\| x \|_2 \leqslant \| x \|_1 \leqslant M.\| x \|_2$ pour tout $x \in E$;***

c) ***l'application identique de E (muni de $\| \ \|_1$) dans E (muni de $\| \ \|_2$) est un isomorphisme.***

PREUVE. — Evidemment $a \Leftrightarrow c$, et $b \Leftrightarrow c$ résulte du théorème précédent. □

Cas des espaces de dimension finie. 2.7.

Théorème. — *Sur un e.v. de dimension finie toutes les normes sont équivalentes.*

PREUVE. — Tout vecteur $x \in E$ s'écrit $x = x_1.e_1 + \cdots + x_n.e_n$ dans une base fixée (e_i) de E. Il est clair que $| x | = | x_1 | + \cdots + | x_n |$ définit une norme. Si $\| \ \|_1$ est une autre norme, $\| x \|_1 = \| \sum x_i.e_i \|_1 \leqslant \sum | x_i |.\| e_i \|_1 \leqslant M.(| x_1 | + \cdots + | x_n |) = M.\| x \|$, où $M = \sup \{ \| e_1 \|_1, \ldots, \| e_n \|_1 \}$. Il en résulte que $x \rightsquigarrow \| x \|_1$ est continue pour la topologie de $\| \ \|$, car (voir 1.1.) on a $| \| x \|_1 - \| y \|_1 | \leqslant \| x - y \|_1 \leqslant M.\| x - y \|$.

D'autre part la sphère unité S de E, normé par $\| \ \|$, est un ensemble fermé et borné, donc compact, pour la topologie usuelle. Comme $x \rightsquigarrow \| x \|_1$ est continue et positive sur S, on a $m = \inf_{x \in S} \| x \|_1 > 0$. En résumé, il existe bien des constantes $m > 0$ et $M > 0$ telles que $m \leqslant \| x \|_1 \leqslant M$ pour $\| x \| = 1$ et le théorème résulte de 2.6. □

Théorème. — *Soient E et F deux e.v. de dimensions finies sur le même corps. Alors toute application linéaire $f : E \to F$ est continue.*

PREUVE. — Soient (e_i) [resp. (f_j)] une base de E (resp. F). Choisissons sur E la norme $\| x \| = \sum | x_i |$ ci-dessus et sur F la norme analogue $\| y \| = \sum | y_j |$. On a $\| f(x) \| = \| \sum x_i f(e_i) \| \leqslant \sum | x_i |.\| f(e_i) \| \leqslant A.\sum | x_i | = A.\| x \|$, où $A = \sup \{ \| f(e_1) \|, \ldots, \| f(e_n) \| \}$. Ainsi f est continue pour les normes $\| \ \|$, donc pour toutes normes d'après le théorème précédent. □

3. Applications multilinéaires continues

Définition. 3.1. — Soient $E_1, \ldots, E_n$ et F des e.v. sur le même corps. Une application $f : E_1 \times \cdots \times E_n \to F$ est dite multilinéaire (bilinéaire si $n = 2$) si la fonction $f(x_1, \ldots, x_n)$ est linéaire par rapport à chacune des variables x_i [attention ! cela ne signifie pas que f est une application linéaire de $E_1 \oplus \cdots \oplus E_n$ dans F].

Supposons en outre, que $E_1, \ldots, E_n$ et F soient normés. Alors $E_1 \times \cdots \times E_n$ a une structure d'espace normé (la somme directe des E_i) et il est loisible de parler d'application multilinéaire $f : E_1 \times \cdots \times E_n \to F$ continue. A ce titre, voici une généralisation du théorème 2.1 :

Théorème. 3.2. — *Les conditions suivantes sont équivalentes :*

a) ***f est continue en tout point de $E_1 \times \cdots \times E_n$;***

b) ***f est continue à l'origine $(0, \ldots, 0)$ de $E_1 \times \cdots \times E_n$;***

c) ***$\| f(x_1, \ldots, x_n) \|$ est borné sur le produit des boules unités $\| x_i \| \leqslant 1$ des E_i ;***

d) ***$\| f(x_1, \ldots, x_n) \|$ est borné sur le produit des sphères unités $\| x_i \| = 1$ des E_i ;***

e) ***il existe une constante M telle que $\| f(x_1, \ldots, x_n) \| \leqslant M.\| x_1 \| \ldots \| x_n \|$ pour tous les $x_1 \in E_1, \ldots, x_n \in E_n$.***

PREUVE. — Evidemment $a \Rightarrow b$.

Montrons que $b \Rightarrow c$: il existe une constante $\delta > 0$ telle que $\| x_1 \| + \cdots + \| x_n \| \leqslant \delta \Rightarrow \| f(x_1, \ldots, x_n) \| \leqslant 1$. Alors $\| x_1 \| \leqslant 1, \ldots, \| x_n \| \leqslant 1 \Rightarrow \| \delta.x_1/n \| + \cdots + \| \delta.x_n/n \| \leqslant \delta \Rightarrow \| f(\delta.x_1/n, \ldots, \delta.x_n/n) \| \leqslant 1 \Rightarrow \| f(x_1, \ldots, x_n) \| \leqslant (n/\delta)^n$.

Evidemment $c \Rightarrow d$.

Montrons que $d \Rightarrow e$: soit M un majorant de $\| f(x_1, \dots, x_n) \|$ lorsque $\| x_1 \| = \cdots = \| x_n \| = 1$. Si les x_i sont tous non nuls, les $x_i / \| x_i \|$ sont de norme 1, donc $\| f(x_1 / \| x_1 \|, \dots) \| \leqslant M$ et par suite $\| f(x_1, \dots, x_n) \| \leqslant M . \| x_1 \| \dots \| x_n \|$; inégalité encore vérifiée si l'un des x_i est nul.

Enfin $e \Rightarrow a$ résulte de

$$\| f(x_1, \dots, x_n) - f(x'_1, \dots, x'_n) \| = \| f(x_1 - x'_1, x_2, \dots, x_n) + f(x'_1, x_2 - x'_2, x_3, \dots, x_n) + \\ + \cdots + f(x'_1, \dots, x'_{n-1}, x_n - x'_n) \| \leqslant M . [\| x_1 - x'_1 \| . \| x_2 \| \dots \| x_n \| + \cdots + \\ + \| x'_1 \| \dots \| x'_{n-1} \| . \| x_n - x'_n \|] .$$

L'espace $\mathscr{L}(E_1, \dots, E_n ; F)$. 3.3. — L'ensemble des applications multilinéaires continues $f : E_1 \times \cdots \times E_n \to F$, muni de façon naturelle d'une addition et du produit par un scalaire, est évidemmént un e.v. Nous le noterons $\mathscr{L}(E_1, \dots, E_n ; F)$. Si $E_1 = \cdots = E_n = E$, nous le noterons simplement $\mathscr{L}^n(E ; F)$.

Posons $\| f \| = \sup \{ \| f(x_1, \dots, x_n) \| \}$ lorsque $\| x_1 \| = \cdots = \| x_n \| = 1$. D'après la preuve de $d \Rightarrow e$ du théorème précédent on a $\| f(x_1, \dots, x_n) \| \leqslant \| f \| . \| x_1 \| \dots \| x_n \|$ pour tous $x_1, \dots, x_n$. On vérifie alors, comme en 2.2, que $f \rightsquigarrow \| f \|$ est bien une norme sur $\mathscr{L}(E_1, \dots, E_n ; F)$ et que, comme en 2.3, $\mathscr{L}(E_1, \dots, E_n ; F)$ ainsi normé est un espace de Banach si F est un espace de Banach.

EXEMPLE. 3.4. — Soient E, F et G trois e.v. normés sur le même corps. Si $f \in \mathscr{L}(E ; F)$ et $g \in \mathscr{L}(F ; G)$, notons $c(g, f)$ l'application linéaire composée $g \circ f : E \to G$. Evidemment $c(g, f)$ est continue. Ainsi c est une application de $\mathscr{L}(F ; G) \times \mathscr{L}(E ; F)$ dans $\mathscr{L}(E ; G)$. On sait qu'elle est bilinéaire et (voir 2.4.) que $\| g \circ f \| \leqslant \| g \| . \| f \|$; par conséquent elle est continue et de norme $\| c \| \leqslant 1$.

4. Isomorphismes canoniques

Théorème. 4.1. — *Soit E un* e.v. *normé construit sur le corps K. Alors $\mathscr{L}(K ; E)$ est canoniquement isomorphe à E.*

PREUVE. — Définissons $\mathscr{L}(K ; E) \to E$ par $f \rightsquigarrow \hat{f} = f(1)$. Evidemment $\hat{}$ est linéaire. Elle est injective car $\hat{f} = 0 \Rightarrow f(1) = 0 \Rightarrow f(t) = t.f(1) = 0$ pour tout $t \in K \Rightarrow f = 0$. Elle est surjective : si $x \in E$ posons $f(t) = t.x$, où $t \in K$; évidemment $f \in \mathscr{L}(K ; E)$ et $\hat{f} = x$. Enfin, elle est continue et $\| \hat{f}(t) \| = | t | . \| f \|$ montre que sa norme $\| \hat{} \| = 1$. □

Théorème. 4.2. — *$\mathscr{L}(E ; \mathscr{L}^k(E ; F))$ est canoniquement isomorphe à $\mathscr{L}^{k+1}(E ; F)$.*

PREUVE. — Si $f \in \mathscr{L}(E ; \mathscr{L}^k(E ; F))$ définissons $\hat{f}$ par $\hat{f}(x_1, \dots, x_{k+1}) = f(x_{k+1})(x_1, \dots, x_k)$. Evidemment $\hat{f} \in \mathscr{L}^{k+1}(E ; F)$ et $\hat{f}$ est une application de $\mathscr{L}(E ; \mathscr{L}^k(E ; F))$ dans $\mathscr{L}^{k+1}(E ; F)$. Il est clair que $\hat{f}$ est linéaire. Elle est injective, car $\hat{f} = 0 \to f(x_{k+1})(x_1, \dots, x_k) = 0$ pour tous les $x_1, \dots, x_{k+1} \Rightarrow f(x_{k+1}) = 0$ pour tout $x_{k+1} \Rightarrow f = 0$. Elle est surjective : si $g \in \mathscr{L}^{k+1}(E ; F)$ définissons $f : E \to \mathscr{L}^k(E ; F)$ par $x \rightsquigarrow g(x, \dots,)$; évidemment $f \in \mathscr{L}(E ; \mathscr{L}^k(E ; F))$ et $\hat{f} = g$. Enfin $\hat{f}$ est continue, car

$$\| \hat{f}(x_1, \dots, x_{k+1}) \| = \| f(x_{k+1})(x_1, \dots, x_k) \| \leqslant \| f(x_{k+1}) \| . \| x_1 \| \dots \| x_k \| \leqslant \\ \leqslant \| f \| . \| x_{k+1} \| \; \| x_1 \| \dots \| x_k \|$$

montre que $\| \hat{f} \| \leqslant \| f \|$, donc $\| \hat{} \| \leqslant 1$. [A titre d'exercice on montrera que $\| \hat{} \| = 1$.]
□

Appendice B

THÉORÈME DU POINT FIXE DE BANACH

Point fixe. 1. — On appelle point fixe d'une application f d'un ensemble E dans lui-même, un point $a \in E$ tel que $f(a) = a$.

Voici l'une des raisons de l'importance de cette notion. Si E est un e.v. et g une application de E dans E, résoudre en x l'équation $g(x) = b$, où $b \in E$ est donné, revient à trouver le (ou les) point(s) fixes(s) de $f(x) = g(x) + x - b$.

Contraction. 2. — Soit E un espace métrique, dont la distance est notée d. Une application f d'une partie U de E dans E est appelée une contraction s'il existe un nombre réel k, $0 \leqslant k < 1$ tel que $d[f(x), f(y)] \leqslant k.d(x, y)$ pour tous $x, y \in U$. On appelle k le rapport de contraction.

Notons qu'une contraction est uniformément continue.

Théorème de Banach. 3. [En anglais : « contracting mapping principle. »]

Soient E un espace métrique complet et f : U → E une application d'un ouvert U de E dans E. Faisons les hypothèses suivantes :

a) ***f est une contraction de rapport k ;***

b) ***il existe u ∈ U dont la distance*** $\mathbf{d}[u, E \setminus U]$ ***au complémentaire de U excède un nombre fixe M > o ;***

c) $\mathbf{d}[u, f(u)] < M.(1 - k)$.

Preuve. — Evidemment $f(u) \in U$, car $d[u, f(u)] < M.(1 - k) < M < d[u, E \setminus U]$. Définissons u_n par récurrence : $u_0 = u, \ldots, u_{n+1} = f(u_n)$. Pour que cela ait un sens il faut montrer que $u_n \in U$. Nous le savons déjà pour u_0 et u_1. Supposons-le établi pour $u_0, u_1, \ldots, u_n$. On a $d[u, f(u_n)] \leqslant d[u, u_1] + \cdots + d[u_n, f(u_n)]$. Mais il est clair que

$$d[u_p, f(u_p)] = d[f^p(u), f^p(u_1)] < k^p.M.(1 - k) \quad \text{pour} \quad 0 \leqslant p \leqslant n.$$

Donc $d[u, f(u_n)] < M.(1 - k).[1 + \cdots + k^n] = M.(1 - k^n) < M$.
Ainsi $f(u_n) = u_{n+1} \in U$, et la propriété est établie par récurrence.

Le même raisonnement montre encore que $d[u_p, u_{p+q}] \leqslant M.(1 - k).[k^p + \cdots + k^{p+q}] < M.k^p$. La suite u_n est donc une suite de Cauchy de U, qui converge vers une limite a, puisque E est complet. Cette limite est dans U, car l'inégalité précédente entraîne, en faisant $q \to +\infty$, $d(u_p, a) < M.k^p < M$. D'autre part, en faisant $p \to +\infty$, on obtient $d(u, a) < M$.

Montrons que a est un point fixe. En effet, puisque f est continue, $u_{n+1} = f(u_n)$ entraîne $a = f(a)$ en faisant $n \to +\infty$.

Soit b un second point fixe éventuel, $d(a, b) = d[f(a), f(b)] < k.d(a, b)$ montre que $d(a, b) = 0$. Donc $b = a$. □

Remarque. 4. — Si $U = E$, les hypothèses b et c sont superflues et n'importe quel point u fait l'affaire dans le théorème précédent.

Donnons deux corollaires utiles :

Dépendance d'un paramètre. 5. — ***Gardons les notations du théorème précédent. Soient T un espace topologique et f : U × T → E une application satisfaisant aux hypothèses suivantes :***

a) ***chaque application partielle*** $f_t : x \rightsquigarrow f(x, t)$, ***où t ∈ T, de U dans E vérifie les hypothèses du théorème de Banach ;***

b) ***pour chaque $x \in U$, l'application partielle $t \rightsquigarrow f(x, t)$ est continue.***
Alors, si a_t est l'unique point fixe de f_t, l'application $t \rightsquigarrow a_t$ est continue.

PREUVE. — Soient $t, s \in T$. On a

$$\mathrm{d}(a_t, a_s) = \mathrm{d}[f_t(a_t), f_s(a_s)] \leqslant \mathrm{d}[f_s(a_s), f_s(a_t)] + \mathrm{d}[f_s(a_t), f_t(a_t)] \leqslant$$
$$\leqslant k.\mathrm{d}(a_t, a_s) + \mathrm{d}[f_s(a_t), f_t(a_t)] .$$

Donc $\mathrm{d}(a_t, a_s) \leqslant (1 - k)^{-1}.\mathrm{d}[(f_s(a_t), f_t(a_t)]$, qui tend vers zéro si s tend vers t. □

Une itérée est une contraction. 6. — ***Si f est une application continue d'un espace métrique complet dans lui-même, dont une itérée f^r est une contraction de rapport k, alors f possède un point fixe unique.***

PREUVE. — D'après le théorème de Banach, f^r possède un point fixe $a = \lim\limits_{n=\infty} (f^r)^n (u)$. Puisque f est continue, $f(a) = f[\lim...] = \lim\limits_{n=\infty} f^{r.n+1}(u)$. Mais $\mathrm{d}[(f^r)^n (f(u)), (f^r)^n (u)] \leqslant$ $\leqslant k^n.\mathrm{d}[f(u), u]$ montre que $\lim\limits_{n=\infty} f^{r.n+1}(u) = \lim\limits_{n=\infty} f^{r.n}(u) = a$.

Enfin, si $f(b) = b$, on a $f^r(b) = b$; et $b = a$, puisque f^r ne possède qu'un point fixe. □

Caractère constructif du théorème de Banach. 7. — Comme nous l'avons vu à propos du théorème d'inversion locale (chap. 3), le théorème de Banach permet de prouver l'existence et l'unicité de solution d'équations. Mais il fournit en outre un procédé effectif pour construire cette solution par approximations successives. La preuve du théorème de Banach montre que la suite $u_0 = u$, $u_1 = f(u), ..., u_n = f^n(u)$ converge vers la solution a de $f(x) = x$. Elle montre aussi que la « *vitesse d'approximation* » *de cette solution par la solution approchée u_n est polynomiale en n* : $\mathrm{d}(a, u_n) \leqslant M.k^n$.

Prenons, par exemple, une fonction $f : [a, b] \to \mathbf{R}$, dérivable et dont la dérivée vérifie $|f'(x)| \leqslant k < 1$ pour un nombre k fixé et pour tout $x \in [a, b]$. Supposons enfin que f applique $[a, b]$ dans lui-même. Le théorème des accroissements finis (1.6. chap. 2) montre que $|f(x) - f(y)| < k.|x - y|$ pour tous $x, y \in [a, b]$. La fonction f est donc une contraction de $[a, b]$ dans $[a, b]$ et les figures ci-contre montrent les approximations successives x_n de la solution de $f(x) = x$ dans les cas respectifs $0 < f'(x) < 1$ et $-1 < f'(x) < 0$.

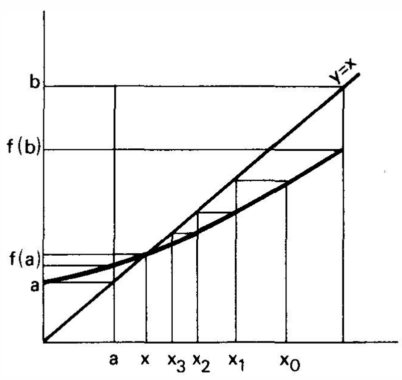


Fig. 11.

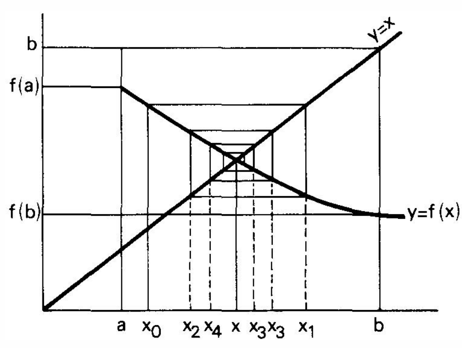


Fig. 12.

La portée de la méthode est plus générale qu'il n'y paraît. Soit à résoudre une équation de la forme $F(x) = 0$, où $F : [a, b] \to \mathbf{R}$ est une fonction dérivable telle que $F(a) < 0$, $F(b) > 0$, et telle qu'il existe deux constantes positives k_1 et k_2 satisfaisant $k_1 \leqslant F'(x) \leqslant k_2$ pour tout $x \in [a, b]$.

En posant $f(x) = x - t.F(x)$, $t \neq 0$, nous sommes ramenés à chercher les solutions de $f(x) = x$. Puisque $f'(x) = 1 - t.F'(x)$, on a $1 - t.k_2 \leqslant f'(x) \leqslant 1 - t.k_1$. La procédure antérieure s'appliquera, en ajustant t pour que $| 1 - t.k_2 | < 1$, $| 1 - t.k_1 | < 1$, car alors f sera une contraction.

Remarque. 8. — Donnons-nous un e.v. E et une application $f : E \to E$. Proposons-nous de résoudre l'équation $f(x) = x$. S'il existe sur E une norme pour laquelle : *a*) E est un espace de Banach, *b*) f est une contraction, alors le théorème de Banach fournit une solution.

Il se peut qu'il existe plusieurs normes vérifiant la condition *a*); c'est en particulier le cas si E est de dimension finie. S'il en est ainsi, les conditions sous lesquelles f est une contraction dépendront de la norme choisie. Cela invite à chercher, dans chaque cas particulier, la norme la « plus avantageuse ».

Illustrons cela par un exemple : soit $x = (x_1, \ldots, x_n) \in \mathbf{R}^n \rightsquigarrow f(x) = L.x + b \in \mathbf{R}^n$ une application affine $[L \in \mathscr{L}(\mathbf{R}^n ; \mathbf{R}^n)$ et $b \in \mathbf{R}^n]$. Sous quelles conditions f est-elle une contraction ?

Prenons d'abord la norme $\| x \|_1 = \sup_i \{ | x_i | \}$. On voit sans peine que

$$\| f(x) - f(y) \|_1 \leqslant \sup_i \sum_{j=1}^{n} | L_{ij} | . \| x - y \|_1 ,$$

où (L_{ij}) est la matrice de L dans la base canonique. Une condition suffisante pour que f soit une contraction est donc $\sum_{j=1}^{n} | L_{ij} | < 1$ pour $i = 1, \ldots, n$.

Prenons maintenant la norme $\| x \|_2 = \sum_{i=1}^{n} | x_i |$. Alors

$$\| f(x) - f(y) \|_2 \leqslant \sup_j \sum_{i=1}^{n} | L_{ij} | . \| x - y \|_2 ,$$

et une condition suffisante est

$$\sum_{i=1}^{n} | L_{ij} | < 1 \quad \text{pour} \quad j = 1, \ldots, n .$$

Prenons enfin la norme euclidienne $\| x \| = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$. L'inégalité de Schwarz entraîne $(\| f(x) - f(y) \|)^2 = \sum_i \left[\sum_j L_{ij}(x_j - y_j) \right]^2 \leqslant \sum_i \sum_j (L_{ij})^2 . \| x - y \|^2$. La condition de contraction s'exprime par $\sum_{i,j} (L_{ij})^2 < 1$.

Chacune des conditions précédentes est *suffisante* pour que f soit une contraction. Mais aucune n'est nécessaire pour assurer la convergence de la suite $x, f(x), \ldots, f^n(x)$. Le lecteur construira des exemples où chacune des trois conditions précédentes est remplie, tandis que les deux autres ne le sont pas.

Appendice C

LA MÉTHODE DE NEWTON

Le théorème de Banach de l'appendice B donne un algorithme pour trouver les racines d'une équation $f(x) = 0$. Si l'on pose $g(x) = f(x) + x$, nous avons vu que, sous certaines conditions, la suite u_0, $u_1 = g(u_0), \ldots, u_{n+1} = g(u_n)$ converge vers l'unique racine a et que *la vitesse d'approximation est polynomiale* : $\| a - u_n \| \leqslant M.k^n$, où M et $0 < k < 1$ sont des constantes.

La méthode de Newton, pour trouver une racine de l'équation $f(x) = 0$, consiste à substituer à la courbe $y = f(x)$ sa tangente en un point, dont l'abscisse x_0 est une approximation de la racine a cherchée. Si $| a - x_0 | < \varepsilon$, l'écart entre la courbe et sa tangente est d'ordre ε^2. L'approximation x_1, définie par l'équation linéarisée $f(x_0) + Df(x_0).(x_1 - x_0) = 0$, est donc voisine de a à ε^2 près (voir figure).

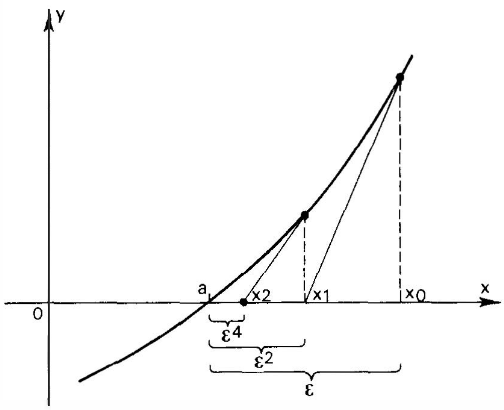


Fig. 13.

En itérant ce procédé, on obtient une suite d'approximations *convergeant exponentiellement* : $| x_{n+1} - x_n | < C.| x_n - x_{n-1} |^2$, C = constante. A l'issue de la $(n + 1)$-ième approximation $| a - x_n |$ est donc de l'ordre ε^{2^n}.

C'est cette méthode que nous allons étendre aux espaces de Banach.

Théorème. — ***Soient E et F deux espaces de Banach, B une boule de rayon $\delta > 0$ de E, $f : B \to F$ une application de classe C^1.***

Faisons les hypothèses suivantes :

a) ***il existe une constante M telle que $\| Df(x) - Df(y) \| \leqslant M.\| x - y \|$ pour tous $x, y \in B$;***

b) ***il existe $x_0 \in B$ tel que $Df(x_0)$ soit un isomorphisme de E sur F et tel que $\| f(x_0) \|$ soit assez petit.***

Alors $x_{n+1} = x_n - [Df(x_n)]^{-1}.f(x_n)$ existe, et la suite x_n converge vers l'unique solution a de $f(x) = 0$. En outre $\| a - x_n \|$ est de l'ordre de ε^{2^n}, où $0 < \varepsilon < 1$ est une constante.

PREUVE. — Soit $x_0 \in B$ une approximation de la racine cherchée a. Posons $y_0 = f(x_0)$; tout revient à résoudre en r l'équation $f(x_0 + r) = 0$, qui s'écrit encore $f(x_0 + r) - f(x_0) = - y_0$.

Supposons que l'on sache que $x_0 + r \in B$, alors $f(x_0) = f(x_0 + r) - Df(x_0 + r).r + R(x_0, r)$, où $R = 0(\| r \|)$. Ainsi r doit vérifier l'équation $-y_0 = Df(x_0 + r).r - R(x_0, r)$.

Puisque Df est continue et que $Df(x_0)$ est un isomorphisme, $Df(x_0 + r)$ est encore un isomorphisme (4.1. chap. 3). Ainsi l'opérateur $\varphi : r \rightsquigarrow -[Df(x_0 + r)]^{-1}.[y_0 - R(x_0, r)]$ est bien défini si $\| r \|$ est assez petit.

Nous allons montrer que :

1) φ est une contraction de la boule $B(x_0, \delta')$ dans elle-même, si son rayon δ' est assez petit ;

2) l'algorithme de construction du point fixe de cette contraction, tel qu'il est donné dans le théorème de Banach, coïncide avec celui de l'énoncé.

Commençons par estimer $g(r) = [Df(x_0 + r)]^{-1}$. L'identité $A^{-1} - B^{-1} = A^{-1}.(B - A).B^{-1}$ entre isomorphismes linéaires et l'hypothèse a du théorème entraînent $\| g(r) - g(r') \| \leqslant M.\| r - r' \|.\| g(r) \|.\| g(r') \|$. Posons $C = \| g(0) \|$, il en résulte $\| g(r) \| \leqslant \| g(r) - g(0) \| + \| g(0) \| \leqslant C.M.\| r \|.\| g(r) \| + C \leqslant C.[1 - C.M.\| r \|]^{-1}$.

Par conséquent, si $\| r \| \leqslant \delta' < \delta$ et $(2\, MC)^{-1}$, alors $g(r)$ existe et $\| g(r) \| \leqslant 2\, C$.

D'autre part, si $r, r' \in B(x_0, \delta')$ on a $\| \varphi(r) - \varphi(r') \| \leqslant \| g(r) \|.\| R(x_0, r) - R(x_0, r') \| + \| g(r) - g(r') \|.\| y_0 - R(x_0, r') \|$. Mais, d'après le corollaire 1.7. (chap. 2) et l'hypothèse a) du théorème

$$\begin{aligned}\| R(x_0, r) - R(x_0, r') \| &= \| f(x_0 + r') - f(x_0, r) + Df(x_0 + r).r - Df(x_0 + r').r' \| \leqslant \\ &\leqslant \| f(x_0 + r') - f(x_0 + r) - Df(x_0 + r).(r' - r) \| + \| [Df(x_0 + r) - Df(x_0, r')].r' \| \leqslant \\ &\leqslant M.[\| r' - r \|^2 + \| r' \|.\| r' - r \|].\end{aligned}$$

En utilisant les résultats précédents et l'hypothèse a) du théorème, on obtient

$$\begin{aligned}\| \varphi(r) - \varphi(r') \| &\leqslant 2.C.M.[\| r' - r \|^2 + \| r' \|.\| r' - r \|] + \\ &+ 4\, C^2.M.\left(\| y_0 \| + \frac{1}{2} M.\| r \|\right)^2.\| r' - r \|,\end{aligned}$$

c'est-à-dire $\| \varphi(r) - \varphi(r') \| \leqslant K.\| r - r' \|$, où K sera une constante positive, indépendante de r et r', et inférieure à 1 si δ' et $\| y_0 \|$ sont assez petits.

Par ailleurs $\| \varphi(r) \| \leqslant \| \varphi(r) - \varphi(0) \| + \| \varphi(0) \| \leqslant K.\| r \| + C.\| y_0 \|$ implique $\| \varphi(r) \| \leqslant \delta'$ si $C.\| y_0 \| \leqslant (1 - K).\delta'$.

En résumé, en choisissant $\| y_0 \| = \| f(x_0) \|$, et par conséquent δ', assez petits, φ est une contraction de $B(x_0, \delta')$ dans elle-même. D'après le théorème de Banach, φ possède un unique point fixe $a \in B(x_0, \delta')$, qui est la limite de la suite définie par $r_0 = 0$,

$$\begin{aligned}r_{n+1} &= \varphi(r_n) = -[Df(x_0 + r_n)]^{-1}.[y_0 - R(x_0, r_n)] = \\ &-[Df(x_0 + r_n)]^{-1}.[f(x_0 + r_n) - Df(x_0 + r_n).r_n] = r_n - [Df(x_0 + r_n)]^{-1}.f(x_0 + r_n).\end{aligned}$$

Posons $x_n = x_0 + r_n$, nous obtenons la relation de récurrence annoncée dans le théorème : $x_{n+1} = x_n - [Df(x_n)]^{-1}.f(x_n)$.

Evaluons la vitesse de convergence. D'après 1.7 (chap. 2) et l'hypothèse a) du théorème, on a $\| f(x_n) - f(x_{n-1}) - Df(x_n).(x_n - x_{n-1}) \| \leqslant M.\| x_n - x_{n-1} \|^2$. Ainsi $\| x_{n+1} - x_n \| \leqslant D.\| x_n - x_{n-1} \|^2$, où D est un majorant des $M.\| [Df(x_n)]^{-1} \|$.

Il en résulte

$$\begin{aligned}\| x_{n+1} - x_n \| &\leqslant D.\| x_n - x_{n-1} \|^2 \leqslant \cdots \leqslant D^{1+2+\ldots+2^{n-1}}.\| x_1 - x_0 \|^{2^n} = \\ &= D^{-1}.(D.\| x_1 - x_0 \|)^{2^n}.\end{aligned}$$

Par conséquent, si l'on pose $\varepsilon = D.\| x_1 - x_0 \| = D.\| [Df(x_0)]^{-1} \|.\| f(x_0) \|$ et $a = \lim_{n = \infty} x_n$, on obtient

$$\| a - x_n \| \leqslant \sum_{j=0}^{\infty} \| x_{n+j+1} - x_{n+j} \| \leqslant D.[\varepsilon^{2^n} + \varepsilon^{2^{n+1}} + \cdots] = 0(\varepsilon^{2^n}). \qquad \square$$

Appendice D

THÉORÈMES D'INVERSION GLOBALE

Une application $f: \mathbf{R} \to \mathbf{R}$ de classe C^1, et dont la dérivée ne s'annule pas, est un difféomorphisme d'un voisinage de chaque point sur son image. Mais, à moins d'imposer une condition supplémentaire, ce n'est pas nécessairement un difféomorphisme de $\mathbf{R}$ sur $\mathbf{R}$ (songer à Arctg x). Une telle condition peut être de caractère global : par exemple : $\lim_{|t| \to \infty} |f(t)| = \infty$. Elle peut aussi être de nature locale ; par exemple : il existe un nombre $k > 0$ tel que $f'(x) > k$ pour tout x. C'est cette dernière condition que nous allons adapter aux espaces de Banach. Remarquons qu'elle peut s'écrire de deux façons, dont chacune va donner lieu à un théorème global.

a) il existe un nombre $k > 0$ tel que $[f(y) - f(x)].[y - x] \geqslant k.(y - x)^2$;

b) il existe un nombre $k > 0$ tel que $|[Df(x)]^{-1}| < 1/k$ pour tout x.

1. Applications strictement monotones

Dans ce paragraphe, H est un espace de Hilbert réel, de produit scalaire $\langle\ ,\ \rangle$. Avec des modifications mineures, laissées à la discrétion du lecteur, les résultats qui suivent s'étendent aux espaces de Hilbert complexes.

Définition. 1.1. — Généralisons la condition *a* de l'introduction.

Une application $f: H \to H$ est dite strictement monotone s'il existe un nombre $k > 0$ tel que $\langle f(y) - f(x), y - x \rangle \geqslant k . \| y - x \|^2$ pour tous $x, y \in H$.

Si H était un espace de Hilbert complexe, le membre de gauche pourrait être remplacé par sa norme, ou par sa partie réelle.

Notons que l'inégalité de Schwarz entraîne

(1.2)
$$\| f(y) - f(x) \| \geqslant k. \| y - x \|$$

Théorème. 1.3. — ***Une application strictement monotone de classe C^1 est un difféomorphisme de classe C^1 de H sur H.***

La preuve va résulter des lemmes suivants.

Lemme. 1.4. — ***La différentielle de f vérifie***

(1.5)
$$\langle Df(x).y, y \rangle \geqslant k. \| y \|^2 \quad \textit{pour tous } x, y \in H .$$

PREUVE. — Pour tout t réel et pour tous $x, y \in H$ on a $\langle f(x + ty) - f(x), ty \rangle \geqslant k. t^2. \| y \|^2$. Divisons les deux membres par $t^2 \neq 0$ et faisons tendre t vers zéro, en tenant compte de la continuité de $u \rightsquigarrow \langle u, y \rangle$. On obtient (1.5). □

Notons que, réciproquement, (1.5) entraîne la stricte monotonicité, dont on a ainsi un critère local. En effet, on a $\langle Df(x + ty).y, y \rangle \geqslant k. \| y \|^2$ pour tout $t \in [0, 1]$ et tous $x, y \in H$. Intégrons les deux membres entre 0 et 1, en tenant compte du théorème fondamental du calcul intégral (6.6. du chap. 2) et de la propriété *d* de (7.4. chap. 2) :

$$\langle f(x+y)-f(x), y\rangle = \langle \int_0^1 Df(x+ty).y.\mathrm{d}t, y\rangle =$$

$$= \int_0^1 \langle Df(x+t.y).y, y\rangle.\mathrm{d}t \geqslant k.\| y\|^2$$

Lemme de Lax-Milgram. 1.6. — ***Soient H un espace de Hilbert réel et $L \in \mathscr{L}(H\,;\,H)$. Supposons qu'il existe un nombre $k > 0$ tel que $\langle Lx, x\rangle \geqslant k.\| x\|^2$ pour tout $x \in H$. Alors L est un isomorphisme et $\| L^{-1}\| \leqslant 1/k$.***

PREUVE. — L est injective, $Lx = 0$ entraînant évidemment $x = 0$.

Montrons que $L(H)$ est un sous-espace fermé. Soit $y_n = L(x_n)$ une suite convergeant vers y. L'inégalité de Schwarz montre que $\| Lx \| \geqslant k.\| x \|$ pour tout x, donc

$$\| y_p - y_q \| = \| L(x_p - x_q)\| \geqslant k.\| x_p - x_q \|$$

et x_n est une suite de Cauchy. Puisque H est complet et L continue, x_n converge vers un élément x et $L(x) = y$.

Montrons que L est surjective. Le théorème de la projection montre qu'il suffit de prouver que l'orthocomplément A de l'e.v. fermé $L(H)$ est réduit à $\{ o \}$. Si $a \in A$, alors $o = \langle La, a\rangle \geqslant k.\| a\|^2$, et l'on a bien $a = 0$.

L'application linéaire L possède donc une inverse L^{-1}. En posant $Lx = y$, l'inégalité $\| Lx\| \geqslant k.\| x\|$ s'écrit $\| y \| \geqslant k.\| L^{-1}(y)\|$. Ainsi L^{-1} est bornée et $\| L^{-1}\| \leqslant 1/k$. □

PREUVE DU THÉORÈME. 1.7. — La relation (1.2) montre que f est injective. Les lemmes 1.4. et 1.6. montrent que $Df(x)$ est un isomorphisme pour tout x. D'après le théorème d'invariance du domaine (4.4. chap. 3), f est donc un difféomorphisme de classe C^1 de E sur l'ouvert $F = f(E)$.

Il nous reste à montrer que f est surjective. Pour cela il suffit de montrer que $f(E)$ est fermé, ou encore que tout point frontière b de $f(E)$ appartient à $f(E)$.

Il existe un chemin, c'est-à-dire une courbe continue $\gamma : [o, 1] \to H$, de classe C^1 sur $[o, 1[$, d'extrémité $\gamma(1) = b$, telle que $\gamma(t) \in f(H)$ pour $0 \leqslant t < 1$. Puisque $f : E \to f(E)$ est un difféomorphisme de classe C^1, $\tilde{\gamma} = f^{-1} \circ \gamma|_{[0,1[} : [o, 1[\to H$ est de classe C^1 et $f \circ \tilde{\gamma}(t) = \gamma(t)$ pour $0 \leqslant t < 1$. Si t_n est une suite croissante de nombres tendant vers 1, (1.2) entraîne

$$\| \gamma(t_p) - \gamma(t_q)\| = \| f[\tilde{\gamma}(t_p)] - f[\tilde{\gamma}(t_q)]\| \geqslant k.\| \tilde{\gamma}(t_p) - \tilde{\gamma}(t_q)\|\,;$$

ainsi $\tilde{\gamma}(t_n)$ est une suite de Cauchy, qui converge vers un point x. La continuité de f entraîne $f(x) = f(x) = \lim_{n=\infty} f[\tilde{\gamma}(t_n)] = \gamma(1) = b$ et b appartient bien à $f(H)$. □

2. Le théorème de Hadamard-Levy

En généralisant la condition b de l'introduction, nous allons obtenir un résultat dû à Jacques Hadamard (1906) et Paul Lévy (1920)

Théorème. 2.1. — ***Soient E et F deux espaces de Banach, $f : E \to F$ une application de classe C^1. Supposons que, pour tout $x \in E$, $Df(x) \in \mathscr{L}(E\,;\,F)$ soit un isomorphisme et qu'il existe un nombre $A > 0$ tel que $\| [Df(x)]^{-1}\| < A$. Alors f est un difféomorphisme de E sur F.***

La preuve exige plusieurs étapes.

Lemme d'unicité des relèvements. 2.2. — ***Soient $x \in E$, $y = f(x)$, $\varphi : Z \to F$ une application continue d'un espace topologique connexe Z dans F. Supposons qu'il existe un point $z \in Z$ et une application continue $\varphi' : Z \to E$ tels que $\varphi(z) = y$, $\varphi'(z) = x$, $f \circ \varphi' = \varphi$*** **[on dit que φ' est un relèvement de φ].** ***Alors φ' est unique.***

PREUVE. — Soit φ'' une application possédant les mêmes propriétés que φ'. Z est l'union disjointe de l'ensemble A des points où φ' et φ'' prennent la même valeur et de son complémentaire B. Nous allons montrer que A et B sont ouverts. Puisque $z \in A$ et que Z est connexe, il en résultera $Z = A$.

Si $a \in A$, $\varphi'(a) = \varphi''(a)$ appartient à un ouvert U dont l'image par f est un voisinage ouvert de $\varphi(a) = f[\varphi'(a)]$, car f est un difféomorphisme local. Ainsi $\varphi'^{-1}(U) \cap \varphi''^{-1}(U)$ est un voisinage ouvert de a contenu dans A.

Si $b \in B$, $\varphi'(b)$ [resp. $\varphi''(b)$] appartient à un ouvert U' [resp. U''] dont l'image par f est un voisinage ouvert de $\varphi(b) = f[\varphi'(b)] = f[\varphi''(b)]$. Ainsi $\varphi'^{-1}(U') \cap \varphi''^{-1}(U'')$ est un voisinage ouvert de b contenu dans B. □

Lemme. 2.3. — ***Soient $x \in E$, $y = f(x)$ et z un point arbitraire de F. Alors il existe un chemin $\tilde{\gamma} : [0, 1] \to E$, de classe C^1, d'origine $\tilde{\gamma}(0) = x$, qui relève le segment $\gamma : [0, 1] \to F$ défini par $\gamma(t) = (1 - t).y + t.z$.***

En particulier f est surjective.

PREUVE. — Soit A l'ensemble des $a \in [0, 1]$ tels que $\gamma|_{[0,a[}$ se relève selon un chemin $[0, a[\to E$, de classe C^1 et d'origine x.

Montrons que A n'est pas vide. D'après le théorème d'inversion locale, il existe un voisinage ouvert U de x dont l'image par f est un voisinage ouvert V de y. Ainsi $\gamma^{-1}(V)$ contient un intervalle $[0, s[$ et $(f|_U)^{-1} \circ \gamma$ est de classe C^1, vaut x pour $t = 0$ et relève $\gamma|_{[0,s[}$.

Si $a, a' \in A$ et $a < a'$, d'après le lemme d'unicité le relèvement de $\gamma|_{[0,a']}$ coïncide sur $[0, a[$ avec celui de $\gamma|_{[0,a[}$.

Soit α la borne supérieure de A. Montrons que $\alpha \in A$. Avec une notation évidente, $f \circ \tilde{\gamma}(t) = \gamma(t)$ pour $0 \leqslant t < \alpha$, donc $Df[\tilde{\gamma}(t)] \circ \tilde{\gamma}'(t) = \gamma'(t)$. L'hypothèse sur Df implique donc $\| \tilde{\gamma}(t) \| \leqslant A.M$, où M est une borne supérieure des $\| \gamma'(t) \|$, $0 \leqslant t \leqslant 1$. Si t_n est une suite croissante de nombres tendant vers α, le théorème des accroissements finis entraîne

$$\| \tilde{\gamma}(t_p) - \tilde{\gamma}(t_q) \| \leqslant A.M.|t_p - t_q|.$$

La suite $\tilde{\gamma}(t_n)$ est donc de Cauchy et elle converge vers un point b. La continuité de f entraîne $f(b) = \lim f \circ \tilde{\gamma}(t_n) = \lim \gamma(t_n) = b$, et $\tilde{\gamma}(\alpha) = b$ est bien défini.

Montrons, par l'absurde, que $\alpha = 1$; moyennant quoi le lemme sera établi. Il existe un voisinage ouvert U' de $\tilde{\gamma}(\alpha)$ dont l'image par f est un voisinage ouvert V' de $\gamma(\alpha)$. Puisque $\alpha < 1$, $\gamma^{-1}(V')$ contient un intervalle ouvert $I =]\alpha - \varepsilon, \alpha + \varepsilon[$. Définissons $\sigma : I \to E$ par $(f|_{U'})^{-1} \cdot \gamma$, puis posons $\tilde{\gamma}_1(t) = \tilde{\gamma}(t)$ si $0 \leqslant t \leqslant \alpha$, $\tilde{\gamma}_1(t) = \sigma(t)$ si $\alpha \leqslant t < \alpha + \varepsilon$. Evidemment $\tilde{\gamma}_1$ est de classe C^1, $\tilde{\gamma}_1(0) = x$ et $\tilde{\gamma}_1$ relève $\gamma|_{[0,\alpha+\varepsilon[}$. Ceci contredit la définition de α. □

Lacet. 2.4. Un chemin sur F (resp. E), c'est-à-dire une application continue $\gamma : [0, 1] \to F$ (resp. E), est appelé un lacet si son origine $\gamma(0)$ coïncide avec son extrémité $\gamma(1)$.

Lemme. 2.5. — ***Un lacet γ de F d'origine $y = f(x)$ se relève suivant un lacet $\tilde{\gamma}$ de E d'origine x.***

En particulier f est injective.

PREUVE. — Quitte à effectuer une translation, on peut supposer $y = 0$.

Soit $Z = \{ 0 \leqslant t \leqslant 1 \} \times \{ 0 \leqslant s \leqslant 1 \}$. Nous allons montrer que $\varphi : Z \to F$, définie par $\varphi(t, s) = s.\gamma(t)$, admet un relèvement $\tilde{\varphi} : Z \to E$ tel que $\tilde{\varphi}(t, 0) = x$ pour

$$0 \leqslant t \leqslant 1, \tilde{\varphi}(0, 1) = \tilde{\varphi}(1, 1) = x.$$

Le lemme en résultera en prenant $\tilde{\gamma}(t) = \tilde{\varphi}(t, 1)$.

D'après 2.2. et 2.3. chaque segment $s \in [0, 1] \rightsquigarrow s.\gamma(t)$ possède un relèvement unique $\tilde{\varphi}_t(s)$ tel que $\tilde{\varphi}_t(0) = x$. Posons $\tilde{\varphi}(t, s) = \tilde{\varphi}_t(s)$. Il suffit de montrer que pour chaque $s \in [0, 1]$, $t \rightsquigarrow \tilde{\varphi}(t, s)$ est un lacet d'origine et d'extrémité x.

Pour cela, considérons l'ensemble A des $a \in [0, 1]$ tels que, pour $0 \leqslant s \leqslant a$, $t \rightsquigarrow \tilde{\varphi}(t, s)$ possède les propriétés voulues. *Mutatis mutandis* le raisonnement du lemme 2.3. montre que $A = [0, 1]$. Bornons-nous à prouver, par l'absurde, que $\alpha = \sup A = 1$.

Chaque point $\tilde{\varphi}(t, \alpha)$ est contenu dans un voisinage ouvert U_t dont l'image par f est un voisinage ouvert V_t de $f[\tilde{\varphi}(t, \alpha)] = \varphi(t, \alpha) = \alpha . \gamma(t)$. Puisque $[0, 1]$ est compact, un nombre fini de ces V_t recouvre l'image de $t \rightsquigarrow \alpha . \gamma(t)$; soient V' leur réunion et U' la réunion des U_t correspondants. Puisque $\alpha < 1$, V' contient l'image d'un lacet $t \rightsquigarrow (\alpha + \varepsilon) . \gamma(t)$ pour un $\varepsilon > 0$. Alors $t \rightsquigarrow (f|_{U'})^{-1} \circ \varphi(t, \alpha + \varepsilon)$ relève ce lacet, c'est lui-même un lacet d'origine x, et l'unicité du relèvement montre qu'il n'est autre que $\tilde{\varphi}(t, \alpha + \varepsilon)$. D'où la contradiction souhaitée.

Si x et x' sont deux points de E de même image $y = f(x) = f(x')$, l'image par f du segment qui les joint est un lacet $\tilde{\gamma}$ d'origine y. Ce lacet se relève, d'après ce qui précède, en un lacet unique d'origine x. Comme ce relèvement doit être aussi le segment xx', c'est que $x = x'$, f est donc injective. $\square$

PREUVE DU THÉORÈME. 2.6. — f est une bijection de E sur F de classe C^1. Puisque $Df(x)$ est inversible en chaque point $Df^{-1} = [Df \circ f^{-1}]^{-1}$ est continue et f^{-1} est de classe C^1.

Remarque. 2.7. — A posteriori f est un revêtement, c'est-à-dire une surjection telle que chaque $y \in F$ possède un voisinage ouvert V dont l'image inverse est l'union disjointe d'ouverts U_i (un seul, dans le cas présent), f réalisant une homéomorphie de chaque U_i sur V. A priori ce n'était pas évident : c'est l'hypothèse $\| [D\varphi]^{-1} \| \leqslant A$ qui entraîne la surjectivité.

Appendice E

RÉDUCTION DES ENDOMORPHISMES LINÉAIRES

Soient E un espace vectoriel complexe de dimension n et $\mathrm{End}(E) = \mathscr{L}(E; E)$. Si $A, B \in \mathrm{End}(E)$, nous noterons simplement $A.B$ l'application composée $A \circ B$ et 1 l'application identique Id_E.

1. Théorème de Hamilton-Cayley

Spectre et polynôme caractéristique. 1.1. — On appelle spectre de $A \in \mathrm{End}(E)$ l'ensemble des scalaires a tels que $A - a.1$ ne soit pas inversible. Les propriétés suivantes sont équivalentes :

a) a appartient au spectre de A ;

b) le noyau de $A - a.1$ n'est pas réduit à zéro ;

c) le déterminant $\mathrm{dét}(A - a.1)$ est nul.

Une valeur propre de A est un élément du spectre de A ; c'est une racine du polynôme de degré n : $k \rightsquigarrow \mathrm{dét}(A - k.1)$, qu'on appelle polynôme caractéristique de A.

Puisque la dimension de E est finie, à chaque valeur propre a correspond un vecteur propre $e \neq 0 : A(e) = a.e$.

Lemme. 1.2. — ***Si $A \in$ End (E) possède n valeurs propres distinctes $a_1, \dots, a_n$, de vecteurs propres correspondants $e_1, \dots, e_n$, alors ces vecteurs forment une base de E. De plus A annule son polynôme caractéristique p.***

PREUVE. — Supposons qu'il existe des scalaires c_i tels que $\sum c_i.e_i = 0$. Appliquons successivement aux deux membres de cette égalité $A, A^2, \dots, A^{n-1}$. On obtient n relations, ce qui s'écrit, sous forme matricielle

$$\begin{bmatrix} 1 & 1 \\ a_1 & a_n \\ \vdots & \vdots \\ (a_1)^{n-1} & (a_n)^{n-1} \end{bmatrix} \cdot \begin{bmatrix} c_1.e_1 \\ \\ \vdots \\ c_n.e_n \end{bmatrix} = 0 .$$

La matrice $n \times n$ est une matrice de Vandermonde de déterminant $\prod_{i<j} .(a_i - a_j) \neq 0$. Il en résulte $c_1 e_1 = \cdots = c_n e_n$, donc les c_i sont nuls et les e_i forment une base de E.

Puisque $A(e_i) = a_i.e_i$, la matrice de A dans cette base est la matrice diagonale $\mathrm{diag}(a_i)$. La matrice de $p(A)$ est donc la matrice diagonale $\mathrm{diag}\,[p(a_i)]$ qui est nulle puisque les a_i sont racines de p. On a donc bien $p(A) = 0$.

Lemme. 1.3. — ***L'ensemble des $A \in$ End (E) dont les n valeurs propres sont distinctes est un ouvert dense dans End (E).***

PREUVE. — Choisissons une base de E. Cela permet d'identifier A à sa matrice et, par conséquent, End (E) à $\mathbf{C}^{n^2}$.

Puisque nous sommes sur le corps **C**, le polynôme caractéristique p de A possède n racines $a_1, \ldots, a_n$, éventuellement confondues. Pour qu'une racine soit multiple il faut et il suffit qu'elle soit aussi racine du polynôme dérivé p'. Soient $b_1, \ldots, b_{n-1}$ les racines de p' ; p et p' ont une racine commune si, et seulement si, $0 = \prod_{s=1}^{n-1} \prod_{r=1}^{n} (a_r - b_s) = p(b_1) \ldots p(b_{n-1})$. Cette dernière expression est un polynôme symétrique en $b_1, \ldots, b_{n-1}$. C'est donc un polynôme en les coefficients de p et p', c'est-à-dire, en fin de compte, un polynôme $P(a_{11}, \ldots, a_{nn})$ en les coefficients de la matrice de A. On a donc défini une application polynomiale $f : A \rightsquigarrow P(a_{11}, \ldots)$ de End $(E) = \mathbf{C}^{n^2}$ dans **C**. Elle ne peut donc s'annuler sur un ouvert de End (E) sans être identiquement nulle. Ce n'est pas le cas : si les valeurs propres de A sont toutes distinctes, $f(A) \neq 0$. Il en résulte que $f(A) \neq 0$ sur le complémentaire U de $f^{-1}(0)$; qui est un ouvert dense dans End (E). □

Remarque. — Tout isomorphisme peut donc être approché d'aussi près qu'on le veut par un endomorphisme diagonalisable d'après 1.1.

Théorème de Hamilton-Cayley. 1.4. — ***Soit p_A le polynôme caractéristique de $A \in$ End (E). Alors $p_A(A) = 0$.***

PREUVE. — Si A appartient à l'ouvert U ci-dessus, toutes ses valeurs propres sont distinctes. D'après le lemme 1.1. on a donc $p_A(A) = 0$. Puisque l'application $A \in$ End $(E) \rightsquigarrow p_A(A) \in$ End (E) est continue et s'annule sur l'ouvert dense U, elle s'annule partout. □

2. Réduction

Lemme. 2.1. — ***Supposons que le polynôme caractéristique p de $A \in$ End (E) soit le produit de deux polynômes p_1 et p_2 premiers entre eux. Alors***

a) ***E est la somme directe des noyaux E_1 et E_2 de $p_1(A)$ et $p_2(A)$. De plus E_1 et E_2 sont invariants par A ;***

b) ***la restriction de A à E_i, $i = 1, 2$, admet p_i comme polynôme caractéristique, à une constante multiplicative près.***

PREUVE.

a) E_i, $i = 1, 2$, est invariant par A. Si $x \in E_i$, alors $p_i(A)\, x = 0$. Mais $p_i(A)$ et A commutent, donc $p_i(A).(Ax) = A.p_i(A)\, x = 0$, et l'on a bien $A.x \in E_i$.

b) Montrons que $p_1(A).E \subset E_2$.

Un vecteur de $p_1(A).E$ s'écrit $p_1(A).x$. Or, d'après le théorème de Hamilton-Cayley, $p_2(A).p_1(A) = p(A) = 0$, donc $p_2(A).p_1(A)\, x = 0$ et l'on a bien $p_1(A).x \in E_2$.

De même $p_2(A).x \in E_1$.

c) Montrons que $E = E_1 + E_2$.

Puisque p_1 et p_2 sont premiers entre eux, d'après le théorème de Bezout il existe des polynômes q_1 et q_2 tels que $p_1.q_1 + p_2.q_2 = 1$. On a donc $p_1(A).q_1(A) + p_2(A).q_2(A) = 1$. En sorte que, si $x \in E$, $p_1(A).q_1(A).x + p_2(A).q_2(A).x = x$. Posons $x_2 = p_1(A).q_1(A).x$, $x_1 = p_2(A).q_2(A).x$; on a $x = x_1 + x_2$ et, d'après *b*), $x_i \in E_i$.

d) Montrons que $E = E_1 \oplus E_2$.

Il reste à prouver que $E_1 \cap E_2 = 0$. Si $x \in E_1 \cap E_2$, $p_1(A).x = p_2(A).x = 0$. Puisqu'on a vu au *c*) que $x = p_1(A).q_1(A).x + p_2(A).q_2(A).x$ et que deux polynômes en A commutent, on a $x = q_1(A).[p_1(A).x] + q_2(A).[p_2(A).x] = q_1(A).0 + q_2(A).0 = 0$.

e) Finalement, montrons que les restrictions A_1 et A_2 de A à E_1 et E_2 admettent pour polynômes caractéristiques respectifs p_1 et p_2, à une constante multiplicative près.

Choisissons une base B de E qui soit réunion d'une base B_1 de E_1 et d'une base B_2 de E_2. Si, par abus de notations, on note encore A, A_1 et A_2 les matrices respectives de A, A_1 et A_2 dans les bases B, B_1 et B_2, on a

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}.$$

Cela montre que le polynôme caractéristique de A est le produit des polynômes caractéristiques p'_1 et p'_2 de A_1 et A_2. Ainsi $p = p'_1 . p'_2 = p_1 . p_2$.

Soit k une racine de p'_1 ; nous allons voir qu'elle est aussi racine de p_1. A k correspond un vecteur propre $x \neq 0$ de la restriction A_1 de A à E_1. Ainsi $x \neq 0$, $x \in E_1$ et $A(x) = k.x$.

Divisons le polynôme $p_1(t)$ par $(t - k)$: $p_1(t) = (t - k)\,q(t) + p_1(k)$. Remplaçons t par A : $p_1(A) = (A - k.1).q(A) + p_1(k).1$, et appliquons les deux membres à x. On a $p_1(A).x = 0$, car $x \in E_1$; $(A - k.1).x = 0$, car $A(x) = k.x$; il reste donc $p_1(k).x = 0$. Comme $x \neq 0$, cela montre que k est racine de p_1.

Ainsi chaque racine de p'_1 (resp. p'_2) est racine de p_1 (resp. p_2). Puisque p_1 et p_2 sont premiers entre eux, ils n'ont pas de racine commune, et comme $p'_1 . p'_2 = p_1 . p_2$, on voit que les racines de p'_i, $i = 1, 2$, sont toutes les racines de p_i, et que leurs ordres de multiplicité sont les mêmes. Par conséquent $p'_1 = c . p_1$, $p'_2 = c^{-1} . p_2$, où c est une constante.

Le lemme s'étend par récurrence au cas où p se décompose en un produit de m polynômes premiers entre eux. On a donc prouvé le résultat invoqué en (5. 3. chap. 5) :

Théorème. — *Soient $k_1, \dots, k_m$ les racines distinctes du polynôme caractéristique p de $A \in \mathrm{End}\,(E)$: $p(k) = (k_1 - k)^{r_1} \dots (k_m - k)^{r_m}$.*

Alors E est somme directe des noyaux $E_1, \dots, E_m$ de $(A - k_1.1)^{r_1}, \dots, (A - k_m.1)^{r_m}$. Les E_i sont invariants par A et la restriction de A à E_i admet pour polynôme caractéristique $(k_i - k)^{r_i}$. En particulier* dim *$E_i = r_i$.

3. Surjectivité de l'exponentielle

On a vu (2.3. chap. 5) que pour tout $X \in \mathrm{End}\,(E)$ l'exponentielle $\exp X$ est inversible. Réciproquement, nous allons voir que pour tout opérateur inversible A d'un espace vectoriel *complexe* E de dimension finie, il existe au moins un $X \in \mathrm{End}\,(E)$ tel que $A = \exp X$ (et par conséquent une infinité puisque, pour tout entier n, $\exp(X + 2\pi i n.1) = \exp(X)$).

1er *cas.* — Supposons que A ait pour polynôme caractéristique $P(k) = (a - k)^r$. Puisque A est inversible, $a \neq 0$ et il existe $x \in \mathbf{C}$ tel que $a = e^x$. Posons

$$N = a^{-1}.A - 1 \quad \text{et} \quad B = \sum_{i=1}^{r-1} (-1)^{i-1} N^i/i .$$

D'après le théorème de Hamilton-Cayley $N^r = 0$; il en résulte $\exp B = 1 + N$ par le calcul classique de $\exp[\mathrm{Log}\,(1 + u)]$. Par conséquent $\exp[x.1 + B] = e^x . \exp B = A$.

Cas général. — Adoptons les notations du théorème précédent. La restriction A_i de A à E_i admettant comme polynôme caractéristique $(k_i - k)^{r_i}$, d'après le premier cas il existe $X_i \in \mathrm{End}\,(E_i)$ tel que $\exp(X_i) = A_i$. Prenons une base de E formée par la réunion de bases de $E_1, \dots, E_n$. Si

$$X = \begin{pmatrix} X_1 & & 0 \\ & \ddots & \\ 0 & & X_m \end{pmatrix},$$

alors

$$e^X = \begin{pmatrix} e^{X_1} & & 0 \\ & \ddots & \\ 0 & & e^{X_m} \end{pmatrix} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix} = A$$

et X répond à la question.

Appendice F

ÉQUATIONS DIFFÉRENTIELLES LINÉAIRES A COEFFICIENTS PÉRIODIQUES

Soient E un espace de Banach, $A : \mathbf{R} \to \mathrm{End}\,(E)$ une fonction continue. Nous supposerons dans toute la suite que A est de période $T > 0$: $A(t + T) = A(t)$ pour tout $t \in \mathbf{R}$.

Voyons les conséquences de cette hypothèse sur les solutions de l'équation différentielle linéaire $x'(t) = A(t).x(t)$, $x(t) \in E$. La question n'est pas triviale, car bien que A soit de période T il n'en est pas nécessairement de même des solutions. C'est déjà faux si A est constant (7. chap. 5).

1. Les opérateurs de monodromie

Théorème. 1.1. — *La résolvante $R(a, t)$ de l'équation différentielle $x' = A.x$ vérifie $R(a + T, t + T) = R(a, t)$ pour tout t.*

Preuve. — D'après (1.4. chap. 6) on sait que $\frac{\mathrm{d}}{\mathrm{d}t} R(a, t) = A(t).R(a, t)$ et $R(a, a) = 1$ (l'application identique de E). Puisque $A(t) = A(t + T)$, $t \rightsquigarrow R(a + T, t + T)$ vérifie la même équation différentielle et la même condition initiale $R(a + t, a + T) = 1$. D'après l'unicité $R(a + T, t + T) = R(a, t)$. □

Conséquence. — La solution générale de $x' = A.x$ s'écrit $f(t) = R(a, t).f(a)$. En particulier $f(a + T) = R(a, a + T).f(a)$. Pour que les solutions de $x' = A.x$ soient toutes de période T il faut donc que $C(a) = R(a, a + T)$ se réduise à 1. On appelle $C(a)$ un *opérateur de monodromie* de $x' = A.x$. Notons que si A est constant, $C(a) = \mathrm{e}^{TA}$ ne dépend pas de a. Ce n'est pas le cas général.

Théorème. 1.2. — *Les opérateurs de monodromie sont tous conjugués et, par conséquent, ils ont mêmes valeurs propres.*

Preuve. — Le théorème précédent et la relation de Chasles pour le produit intégral entraînent $C(b) = R(b, b + T) = R(a + T, b + T).R(a, a + T).R(b, a) = R(a, b).C(a).R(a, b)^{-1}$.

Les valeurs propres de $C(.)$ sont appelées les *nombres caractéristiques* de $x' = A.x$. Si la dimension de E est finie et si A est constant, les nombres caractéristiques sont les $\mathrm{e}^{r.T}$, où les r sont les valeurs propres de A. □

Théorème. 1.3. — *Pour que toutes les solutions de $x' = A.x$ soient de période T, il faut et il suffit que $a \rightsquigarrow C(a)$ prenne la valeur 1.*

Preuve. — On a vu que c'est nécessaire. Montrons que c'est suffisant. Si $C(a) = 1$, la proposition précédente montre que $C(t) = 1$ pour tout t. Comme $f(t + T) = C(t).f(t)$, on a bien $f(t + T) = f(t)$ pour tout t. □

Cherchons s'il existe au moins une solution périodique.

Théorème. 1.4. [G. Floquet. (1883)]. — ***A chaque valeur propre k de $C(a)$ correspond une solution f de $x' = A.x$ telle que $f(t + T) = k.f(t)$ pour tout t.***

En particulier, pour qu'il existe une solution non nulle de période T, il faut et il suffit que 1 soit valeur propre de C.

PREUVE. — Soit $u \neq 0$ un vecteur propre de $C(a)$ de valeur propre k. Soit f la solution de $x' = A.x$ vérifiant $f(a) = u$. Alors $f(a + T) = C(a).f(a) = k.f(a)$. En sorte que, si l'on pose $g(t) = f(t + T) - k.f(t)$, g est la solution de $x' = A.x$ vérifiant $g(a) = 0$. D'après l'unicité $g(t) = 0$ pour tout t.

Si $k = 1$ la solution f ci-dessus est donc de période T. Réciproquement, si f est une solution non nulle de période T, $f(a) = f(a + T) = C(a).f(a)$ montre que $f(a)$ est un vecteur propre de $C(a)$ de valeur propre 1. □

Nous allons voir que les nombres caractéristiques de $x' = A.x$ sont invariants par changement de variable linéaire.

Théorème. 1.5. [A. M. Liapounov. (1892)]. — ***Soit S : R → GL (E) une fonction de classe C^1 et de même période T que A. Si dans l'équation $x' = A.x$ on fait le changement de variable $y(t) = S(t).x(t)$, on obtient une équation différentielle $y' = A_1.y$ dont l'opérateur de monodromie $C_1(a)$ est semblable à celui, $C(a)$, de l'équation initiale.***

PREUVE. — Un calcul immédiat montre que $y' = A_1.y$, où

$$A_1 = [S' + S.A].S^{-1}\,. \tag{1.6}$$

Ainsi S', S et A étant de période T, A_1 l'est aussi et l'on peut parler de l'opérateur de monodromie $C_1(a) = R_1(a, a + T)$, où R_1 est la résolvante de $y' = A_1.y$.

La solution générale de $y' = A_1.y$ est $y(t) = R_1(a, t).y(a) = R_1(a, t).S(a).x(a)$. Comme $x(t) = R(a, t).x(a)$, on a encore $y(t) = S(t).x(t) = S(t).R(a, t).x(a)$. En combinant ces résultats, et parce que $x(a)$ est arbitraire, on trouve $R_1(a, t) = S(t)\,R(a, t).S(a)^{-1}$. Prenons $t = a + T$; puisque $S(a + T) = S(a)$ on en déduit $C_1(a) = S(a).C(a).S(a)^{-1}$. □

Cela montre qu'étant données deux équations différentielles linéaires périodiques et de même période T, il n'est pas toujours possible de transformer l'une dans l'autre par un changement de variable linéaire de période T.

Cherchons s'il est possible de se ramener à une équation $y' = A_1.y$, où A_1 est constant.

2. Le théorème de Liapounov

Théorème. 2.1. [Liapounov. (1907)]. — ***Soit E un* espace vectoriel complexe de dimension finie. *Soit A : R → End (E) une fonction continue de période $T > 0$. Alors il existe au moins une fonction S : R → GL (E) de classe C^1 et de période T, telle que le changement de variable $y(t) = S(t).x(t)$ ramène l'équation $x' = A.x$ à une équation $y' = A_1.y$, où A_1 est constant.***

PREUVE. — Soit $C(a) = R(a, a + T)$ un opérateur de monodromie de $x' = A.x$. Il existe au moins un $B \in \mathrm{End}(E)$ tel que $C(a) = e^{TB}$ (Appendice E). Prenons $S(t) = e^{t.B}.R(a, t)^{-1}$, autrement dit, posons $x(t) = R(a, t).e^{-tB}.y(t)$. La relation (1.6) montre que y vérifie l'équation $y' = B.y$. Evidemment S est de classe C^1. Il nous reste à montrer qu'elle est de période T. Puisque $e^{TB} = R(a, a + T)$, le théorème 1.1. et la relation de Chasles pour le produit intégral entraînent

$$\begin{aligned} S(t + T) &= e^{(t+T).B}.R(t + T, a) = e^{tB}.R(a, a + T).R(t + T, a) = \\ &= e^{tB}.R(t + T, a + T) = e^{tB}.R(t, a) = S(t)\,. \end{aligned}$$
□

Bien que la détermination explicite de S nécessite celle de la résolvante, et en fin de compte la résolution de l'équation $x' = A.x$, on déduit d'importantes conséquences de ce résultat.

Forme des solutions. 2.2. — On connaît (5.3. chap. 5) la forme de la solution générale de l'équation à coefficients constants $y' = B.y$. Si $\lambda_1, \ldots, \lambda_m$ sont les valeurs propres distinctes de B, cette solution est une combinaison linéaire de vecteurs $e^{t.\lambda_i}.t^k.a$, où $a \in E$ et $k <$ ordre de multiplicité de λ_i. On en déduit que la solution générale de $x' = A.x$ est une combinaison linéaire de vecteurs $e^{t.\lambda}.t^k.S(t)^{-1}.a$, où $S(t)$ est de période $T > 0$ (donc bornée).

Puisque $C(a) = e^{TB}$, les valeurs propres de $C(a)$, c'est-à-dire les nombres caractéristiques, sont les $e^{T.\lambda_i}$. On appelle les λ_i les *exposants caractéristiques* de $x' = A.x$.

Voici l'une des nombreuses conséquences de ce résultat (Liapounov).

Pour que toutes les solutions de $x' = A.x$ tendent vers zéro lorsque $t \to +\infty$ (resp. $-\infty$) il faut et il suffit que les parties réelles des exposants caractéristiques soient négatives (resp. positives).

Bibliographie. — Le lecteur trouvera des exposés plus détaillés dans les ouvrages de F. R. Gantmacher et de N. Rouche et J. Mawhin.

Appendice G

LE THÉORÈME D'EXISTENCE ET DE DÉPENDANCE PAR RAPPORT AUX CONDITIONS INITIALES DES SOLUTIONS DES ÉQUATIONS DIFFÉRENTIELLES

On se propose de déduire les théorèmes (2.2.) et (3.4.) du chapitre 7 du théorème des fonctions implicites, selon une méthode due à J. Robbin [On the existence theorem for differential equations. Proc. Amer. Math. Soc. 19, pages 1005-1006, 1968].

Théorème. — ***Soient U un ouvert d'un espace de Banach E, X un champ de vecteurs de classe C^1 sur U. Alors, pour chaque $t_0 \in \mathbf{R}$ et chaque $x_0 \in U$, il existe un voisinage ouvert V de x_0 dans U, un intervalle $]t_0 - a, t_0 + a[$ et une application $\varphi :]t_0 - a, t_0 + a[\times V \to U$ tels que***

a) ***φ soit de classe C^1,***

b) ***$\varphi(t_0, x) = x$ pour $x \in V$;***

c) ***$\dfrac{d}{dt}\varphi(t, x) = X[\varphi(t, x)]$ pour $(t, x) \in]t_0 - a, t_0 + a[\times V$.***

PREUVE. — Sans nuire à la généralité, et afin d'alléger les notations, nous pouvons supposer que $t_0 = 0$, que x_0 est l'origine de 0 de E et que U est la boule $B_{2r}(0)$ de centre 0 et de rayon $2r$. Nous posons $U_0 = B_r(0)$ et $I = [-1, 1]$.

Rappelons que $C^0(I; E)$ [resp. $C^1(I; E)$] désigne l'espace de Banach des applications de classe C^0 [resp. C^1] de I dans E muni de la norme

$$\| f \|_{C^0} = \sup_{t \in I} \| f(t) \| \ [\text{resp. } \| f \|_{C^1} = \| f \|_{C^0} + \| f' \|_{C^0}].$$

L'ensemble $C_0^p(I; E)$ des $f \in C^p(I; E)$, $p = 0$ ou 1, tels que $f(0) = 0$ est un sous-espace fermé de $C^p(I; E)$; c'est donc un espace de Banach. Soit $C_0^p(I; U_0)$ le sous-ensemble des $f \in C_0^p(I; E)$ à valeurs dans U_0. C'est un ouvert de $C_0^p(I; E)$ car l'image $f(I)$ de chaque courbe $f \in C_0^p(I; U)$ est compacte et, par conséquent, maintenue à distance finie du complémentaire de U dans E; il en est donc de même de chaque courbe voisine de f. □

L'application F.

Si $x \in U_0$ et $f \in C_0^1(I; U_0)$ on a $\| x + f(t) \| < 2r$, donc $x + f(t) \in U$ pour $t \in I$ et $X[x + f(t)]$ est défini. Nous pouvons donc définir une application F de l'ouvert $\mathbf{R} \times U_0 \times C_0^1(I; U_0)$ de l'espace de Banach $\mathbf{R} \oplus E \oplus C_0^1(I; E)$ dans l'espace de Banach $C^0(I; E)$ par

$F(k, x, f) : t \rightsquigarrow \dfrac{df(t)}{dt} - k.X[x + f(t)]$.

L'application F est de classe C^1.

D'après (3. chap. 2) il suffit de montrer que les différentielles partielles existent et sont continues.

Evidemment F est différentiable par rapport à k et $D_1 F(k, x, f)$ est l'élément de $\mathscr{L}[\mathbf{R}; C^0(I; E)]$ qui s'identifie canoniquement (d'après 4.1. Appendice A) à l'élément $t \in I \rightsquigarrow -X[x + f(t)]$ de $C^0(I; E)$. Puisque f est continue, $\{ x + f(t) : t \in I \}$ est un compact sur lequel la fonction continue X est uniformément continue. Donnons-nous $\varepsilon > 0$, il lui correspond donc un $\delta > 0$ tel que $\| X[x + f(t)] - X(y) \| < \varepsilon$ si $t \in I$ et $\| x + f(t) - y \| < \delta$. Prenons $x_1, x_2 \in U_0$ tels que $\| x_1 - x_2 \| < \delta/2$ et $f_1, f_2 \in C_0^1(I; U_0)$ telles que $\| f_1 - f_2 \|_{C^1} < \delta/2$. Il en résulte $\| x_1 + f_1(t) - x_2 - f_2(t) \| < \delta$ pour $t \in I$, donc

$$\| D_1 F(k_1, x_1, f_1) - D_1 F(k_2, x_2, f_2) \| = \sup_{t \in I} \| X[x_1 + f_1(t)] - X[x_2 + f_2(t)] \| < \varepsilon .$$

Cela montre que $D_1 F$ est continue.

Puisque X est différentiable, F est différentiable par rapport à x et $D_2 F(k, x, f) = - k.DX[x + f]$. C'est l'élément de $\mathscr{L}[E; C^0(I; E)]$ qui fait correspondre à $y \in E$ l'élément $t \in I \rightsquigarrow - k.DX[x + f(t)].y$ de $C^0(I; E)$. Puisque DX est continue, on montre, comme pour $D_1 F$, que $D_2 F$ est continue.

Enfin, puisque X est différentiable, F est différentiable par rapport à f : si $h \in C_0^1(I; E)$, $D_3 F(k, x, f).h$ est l'élément de $\mathscr{L}[C_0^1(I; E); C^0(I; E)]$ qui associe à h la courbe

$$t \in I \rightsquigarrow \frac{\mathrm{d}h(t)}{\mathrm{d}t} - k.DX[x + f(t)].h(t) \quad \text{de} \quad C^0(I; E) .$$

On voit, comme pour $D_2 F$, que $D_3 F$ est continue.

L'application H.

Le calcul précédent montre que $D_3 F(0, 0, 0) = \dfrac{\mathrm{d}}{\mathrm{d}t}$; autrement dit $D_3 F(0, 0, 0)$ fait correspondre à $h \in C_0^1(I; E)$ la courbe $t \in I \rightsquigarrow \dfrac{\mathrm{d}h}{\mathrm{d}t}(t) \in E$ de $C^0(I; E)$. L'application linéaire continue $D_3 F(0, 0, 0)$ est donc surjective : $g \in C^0(I; E)$ est l'image de $t \in I \rightsquigarrow \displaystyle\int_0^t g(s).\mathrm{d}s$.

Elle est aussi injective : si $h'(t) = 0$ pour $t \in I$, alors $h(t) = \text{constante} = h(0) = 0$. En résumé $D_3 F(0, 0, 0)$ est un isomorphisme de $C_0^1(I; E)$ sur $C^0(I; E)$.

Comme d'autre part, $F(0, 0, 0) = 0$, le corollaire (5.2. chap. 3) du théorème des fonctions implicites dit qu'on peut trouver un voisinage $]-a, a[\times V$ de (o, o) dans $\mathbf{R} \times U_0$ et une application H de classe C^1 de ce voisinage dans $C_0^1(I; U_0)$ tels que $F[k, x, H(k, x)] = 0$ pour $(k, x) \in]-a, a[\times V$, c'est-à-dire

$$\text{(1)} \qquad \frac{\mathrm{d}}{\mathrm{d}t}[H(k, x)(t)] = k.X[x + H(k, x)(t)] \quad \text{pour } t \in I .$$

L'application φ.

Puisque $x \in V \subset U_0$ et que $H(k, x)(t) \in U_0$ alors $\varphi(t, x) = x + H(k, x)(t/k)$, où $-a \leqslant k \leqslant a$, vérifie $\| \varphi(t, x) \| < 2r$ et définit donc une application de $]-k, k[\times V$ dans U.

D'après la définition de H, φ est de classe C^1 en t et

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t, x) = \frac{\mathrm{d}}{\mathrm{d}t}[H(k, x)(t/k)] = k^{-1}.k.X[x + H(k, x)(t/k)] = X[\varphi(t, x)] .$$

D'autre part, puisque $H(k, x) \in C_0^1(I; U_0)$, on a $H(k, x)(0) = 0$ et par suite $\varphi(0, x) = x$.

On a donc exhibé, pour chaque $k \in [-a, a]$, une courbe intégrale de X d'origine x et définie sur $[-k, k]$. Deux quelconques de ces courbes coïncident donc sur l'intersection de leurs domaines de définition d'après l'unicité. Donc $H(k, x)(t/k)$ ne dépend pas de k. Si $t \neq 0$ on en déduit

$$\text{(2)} \qquad \varphi(t, x) = x + H(t, x)(1) .$$

Si $t = 0$, (1) montre que $\dfrac{\mathrm{d}}{\mathrm{d}t}[H(0, x)(t)] = 0$ pour $t \in I$; donc $H(0, x)(t) = \text{constante} = H(0, x)(0) = 0$ et par suite $H(0, x)(1) = 0$. La relation (2) vaut donc encore pour $t = 0$ puisque $\varphi(0, x) = x$. Comme H est de classe C^1 en (t, x), cette relation (2) montre qu'il en est de même pour φ.

Appendice H

SIMPLICITÉ DE SO (3)

Rappelons que SO (n) est le sous-groupe de O (n) formé des éléments a dont le déterminant dét $(a) = +1$.

Nous avons vu (1.9. chap. 9) que O (3) est une sous-variété C^∞ de End$(\mathbf{R}^3) = \mathbf{R}^9$. Nous allons utiliser le théorème d'inversion locale (3.5. chap. 9) pour démontrer que $G = \mathrm{SO}(3)$ est un groupe simple.

1. Préliminaires

Proposition. 1.1. — $G = \mathrm{SO}(3)$ ***est la composante connexe de*** $I = \mathrm{id}_{\mathbf{R}^3}$ ***de la sous-variété*** **O (3)** ***munie de la topologie induite par*** $\mathbf{R}^9$.

PREUVE. — L'application continue dét : O(3) $\to$ **R** ne prend que deux valeurs $+1$ et -1. Donc O (3) possède au moins deux composantes connexes, contenues respectivement dans $\text{dét}^{-1}(1) = G$ et $\text{dét}^{-1}(-1)$.

Il suffit donc de prouver que G est connexe. Un élément $a \in G$ est une rotation d'angle $\hat{a}$ autour d'un axe. Si $r(t)$ désigne la rotation d'angle $t.\hat{a}$ autour du même axe, alors $t \in [0, 1] \rightsquigarrow r(t) \in G$ est un arc d'origine I et d'extrémité a. □

Proposition. 1.2. — ***On définit un produit scalaire sur*** **End** $(\mathbf{R}^n)$ ***en posant*** $\langle c, d \rangle =$ **Trace** $(c \circ {}^t d)$, ${}^t d$ ***est le transposé de*** d.
De plus, si $s \in \mathrm{O}(\mathbf{R}^n)$, ***on a*** $\langle s \circ c, s \circ d \rangle = \langle c, d \rangle$ ***et*** $\langle s \circ c \circ s^{-1}, s \circ d \circ s^{-1} \rangle = \langle c, d \rangle$.

PREUVE. — La première partie est immédiate.

La seconde résulte de ${}^t s = s^{-1}$ et du fait que deux endomorphismes semblables possèdent la même trace. □

Corollaire. 1.3. — ***Soient*** $v \in$ **End** $(\mathbf{R}^n)$ ***et*** $a, b \in$ **SO** (n). ***Alors*** $2v = a \circ v \circ a^{-1} + b \circ v \circ b^{-1}$ ***implique*** $v \circ a = a \circ v$, $v \circ b = b \circ v$.

PREUVE. — On a $2.\| v \|^2 = 2 \langle v, v \rangle = \langle a \circ v \circ a^{-1}, v \rangle + \langle b \circ v \circ b^{-1}, v \rangle$.

Utilisons 1.2 et l'inégalité de Schwarz : $\langle a \circ v \circ a^{-1}, v \rangle \leqslant \| v \|^2$, l'égalité n'ayant lieu que si $a \cdot v \cdot a^{-1} = v$; et l'analogue en changeant a en b. L'inégalité initiale entraîne donc $a \circ v \circ a^{-1} = b \circ v \circ b^{-1} = v$. □

2. Simplicité de SO (3)

Théorème. — ***Un sous-groupe invariant*** $H \neq I$ ***de*** $G =$ **SO** (3) ***coïncide avec*** G.

PREUVE. — Prenons $a \in H - \{ I \}$; c'est une rotation d'axe D. Si X est un vecteur non nul porté par D, on a donc $a(X) = X$.

Soit r une rotation qui envoie X sur un vecteur orthogonal $Y = r(X)$. Alors

$$b = r \circ a \circ r^{-1} \in H \quad \text{et} \quad b(Y) = Y.$$

Considérons l'application $\varphi : G \to G$ définie par $\varphi(g) = g \circ a \circ g^{-1} \circ a^{-1} \circ g \circ b \circ g^{-1} \circ b^{-1}$. Evidemment φ est C^∞, $\varphi(I) = I$ et $\varphi(G) \subset H$. Montrons que φ est un difféomorphisme local de G au voisinage de I. D'après le théorème d'inversion locale (3.5. chap. IX), il suffit de démontrer que $T_I\,\varphi$ est injective.

Soit $v \in T_I\,G$. Prenons une courbe γ tracée sur G, d'origine I et vérifiant $\gamma'(0) = v$. Son image par φ est $\varphi \circ \gamma(t) = \gamma(t) \circ a \circ \gamma(t)^{-1} \circ a^{-1} \circ \gamma(t) \circ b \circ \gamma(t)^{-1} \circ b^{-1}$. D'après (2.8. chap. 4) et la règle de Leibniz, $(\varphi \circ \gamma)'(0) = 2\,v - a \circ v \circ a^{-1} - b \circ v \circ b^{-1}$. Supposons que $v \in \operatorname{Ker} T_I\,\varphi$. Alors $2\,v = a \circ v \circ a^{-1} + b \circ v \circ b^{-1}$ et le corollaire ci-dessus entraîne $v \circ a = a \circ v$, $v \circ b = b \circ v$. Or $a(X) = X$, $b(Y) = Y$; donc $a[v(X)] = v(X)$ et $b[v(Y)] = v(Y)$. Mais, puisque a et b ne se réduisent pas à I, leurs seuls vecteurs invariants sont portés par leur axe de rotation. Donc $v(X)$ est proportionnel à X, et $v(Y)$ à Y. Puisque v est antisymétrique (2.9. chap. 9), $v(X) = v(Y) = 0$. Comme X et Y sont orthogonaux, cela implique $v = 0$. Ainsi $T_I\,\varphi$ est injective, donc bijective.

D'après le théorème d'inversion locale, il existe donc un voisinage ouvert V de I tel que $\varphi(V)$ soit un voisinage ouvert de I. Et l'on a vu que $\varphi(V) \subset \varphi(G) \subset H$.

Montrons que H est un ouvert-fermé de G. Puisque G est connexe (proposition 1.1.), on aura prouvé que $G = H$. En effet, H étant un groupe et la translation à gauche L_h par $h \in H$ étant un difféomorphisme de G (voir 3.4. chap. 9), $L_h\,\varphi(V)$ est un voisinage ouvert de h contenu dans H. Cette translation à gauche par un élément $h \in SO(3)$ conservant également la distance euclidienne définie sur $\operatorname{End}(\mathbf{R}^3)$ en 1.2., on voit que H est fermé.

Note. — Une méthode similaire permet de démontrer la simplicité des groupes classiques, tels $SO(2\,k + 1)$, $k \neq 0$.

BIBLIOGRAPHIE

ABRAHAM R. et MARSDEN J. — *Foundations of Mechanics*, 2e éd., Benjamin, New York, 1978.

ARNOLD V. — *Equations différentielles ordinaires*, Mir, Moscou, 1974.

CARTAN H. — *Cours de calcul différentiel*, 2e éd., Hermann, Paris, 1977.

CODDINGTON E. A. et LEVINSON N. — *Theory of ordinary differential equations*, McGraw Hill, New York, 1955.

DIEUDONNÉ J. — *Eléments d'analyse, tome I : Fondements de l'analyse moderne*, Gauthier-Villars, Paris, 1968.

GANTMACHER F. R. — *The theory of matrices*, Chelsea, New York, 1959.

HIRSCH M. W. et SMALE S. — *Differential equations, dynamical systems, and linear algebra*, Academic Press, New York, 1974.

MALLIAVIN P. — *Géométrie différentielle intrinsèque*, Hermann, Paris, 1972.

MILNOR J. — *Topology from the differentiable point of view*, The University Press of Virginia, Charlottesville, 1965.

NELSON E. — *Topics in dynamics I : Flows*, Princeton University Press, Princeton, 1969.

ROUCHE N. et MAWHIN J. — *Equations différentielles ordinaires*, Masson, Paris, 1973.

YOUNG L. C. — *Lectures on the calculus of variations and optimal control theory*, 2e éd., Chelsea, New York, 1980.

INDEX ALPHABÉTIQUE DES MATIÈRES

MASSON, Editeur
120, bd St-Germain
75280 Paris Cedex 06
Dépôt légal : mars 1983

JOUVE
18, rue Saint-Denis
75001 Paris
10696. Dépôt légal : février 1983

Collection Maîtrise de mathématiques pures
dirigée par J. DIEUDONNÉ et P. MALLIAVIN

ALGÈBRE

ALGÈBRE LINÉAIRE ET GÉOMÉTRIE CLASSIQUE

Cours,
par J.E. et M.J. BERTIN.
1981, 148 pages.

Exercices,
par M.P. MALLIAVIN
et A. WARUSFEL.
1981, 128 pages.

LES GROUPES FINIS
ET LEURS REPRÉSENTATIONS COMPLEXES

Cours,
par M.P. MALLIAVIN.
1981, 96 pages.

Exercices,
par J.P. BÉZIVIN
et A. LÉVY-BRUHL.
1982, 108 pages.

ANALYSE

INTÉGRATION ET PROBABILITÉS. ANALYSE DE FOURIER
ET ANALYSE SPECTRALE

Cours,
par P. MALLIAVIN.
1982, 200 pages.

Exercices,
par G. LETAC.
1982, 152 pages.

CALCUL DIFFÉRENTIEL

Cours,
par A. AVEZ.
1983, 152 pages.

Exercices.
par B. EL MABSOUT
A paraître.

■ Le calcul différentiel est sans doute, avec l'algèbre linéaire, la partie des Mathématiques la plus sollicitée par les utilisateurs (Physiciens. Mécaniciens ou autres) : recherche et étude des solutions des équations différentielles, calcul des variations...

■ Il sert également de base à de nombreuses théories mathématiques : géométrie et topologie différentielles, équations aux dérivées partielles, ...

■ Issu d'un cours enseigné à l'Université PARIS VI et à l'Université de PALERME, ce livre expose les fondements du calcul différentiel dans une perspective moderne, c'est-à-dire libérée du recours systématique aux coordonnées. Le texte est illustré d'exemples concrets, empruntés le plus souvent à la Mécanique.

■ La matière de cet ouvrage peut être enseignée dans un cours semestriel de second cycle et être utile pour la préparation au concours de l'Agrégation.

ISBN : 2-225-7907