Calcul différentiel dans les espaces de Banach

I

1. RAPPEL DE NOTIONS RELATIVES AUX ESPACES DE BANACH ET AUX APPLICATIONS LINÉAIRES CONTINUES

Dans toute la suite, le corps de base K est le *corps réel* **R** ou le corps complexe **C**. On suppose connues les définitions relatives aux espaces vectoriels et leurs propriétés élémentaires. On rappelle que si E est un espace vectoriel *complexe* (c'est-à-dire sur le corps **C**), E possède une structure d'espace vectoriel *réel* sous-jacent : on se borne à considérer le produit λx d'un vecteur $x \in E$ par un scalaire λ lorsque $\lambda \in \mathbf{R}$.

Norme sur un espace vectoriel E 1.1

C'est une fonction $\rho : E \to \mathbf{R}^+$ (où $\mathbf{R}^+$ désigne l'ensemble des nombres réels $\geqslant 0$) ayant les propriétés suivantes :

(i) $\rho(0) = 0$;

(i') $(\rho(x) = 0) \Rightarrow (x = 0)$;

(ii) $\rho(x + y) \leqslant \rho(x) + \rho(y), \forall x, y \in E$;

(iii) $\rho(\lambda x) = |\lambda| \cdot \rho(x), \forall x \in E, \lambda \in K$.

Un espace vectoriel (e.v. en notation abrégée) muni de la donnée d'une norme s'appelle un *espace vectoriel normé* (e.v. normé). Quand la norme ρ est ainsi donnée, on note souvent $\|x\|$ la valeur $\rho(x)$ de la norme pour un vecteur x. Avec cette notation, les conditions (i) à (iii) s'écrivent :

(i) $\|0\| = 0$;

(i') $(\|x\| = 0) \Rightarrow (x = 0)$;

(ii) $\|x + y\| \leqslant \|x\| + \|y\|$;

(iii) $\|\lambda x\| \leqslant |\lambda| \cdot \|x\|$.

Soit E un e.v. normé ; on définit la *distance* de deux points x, y de E par la formule :

$$d(x, y) = \|x - y\|.$$

Puisque $\|x - y\| = \|y - x\|$ à cause de (iii) [remplacer x par $(x - y)$, et λ par -1], on a $d(x, y) = d(y, x)$. De plus, (ii) entraîne immédiatement

$$d(x, z) \leqslant d(x, y) + d(y, z)$$

(« inégalité du triangle »). Enfin, $d(x, y) = 0$ si et seulement si $x = y$. Donc E est un *espace métrique*; et, comme pour tout espace métrique, E a une structure *topologique*.

Pour cette topologie, la norme $u \mapsto \|u\|$ est une application continue $E \to \mathbf{R}$, parce que $|\,\|u\| - \|v\|\,| \leqslant \|u - v\|$.

Soit, pour $a \in E$ et $r > 0$, $B'(a, r)$ la *boule* de centre a et de rayon r, formée des points $x \in E$ tels que

$$d(x, a) \leqslant r, \qquad \text{c'est-à-dire} \quad \|x - a\| \leqslant r.$$

Alors un sous-ensemble $U \subset E$ est dit *ouvert* si, pour tout $a \in U$, il existe $r > 0$ tel que la boule $B'(a, r)$ soit contenue dans U. Ces ensembles ouverts définissent bien une topologie.

On vérifie que la boule $B'(a, r)$ est *fermée* (son complémentaire est un ensemble ouvert). En revanche, la « boule ouverte » $B(a, r)$, formée des x tels que $\|x - a\| < r$, est un ensemble ouvert.

La topologie de E est *séparée*, car si $x \neq y$, posons $d(x, y) = r$; alors les boules ouvertes $B(x, r/2)$ et $B(y, r/2)$ sont disjointes.

Une suite $(x_n)_{n \geqslant 0}$ de points de E a pour *limite* $a \in E$, et on écrit $\lim_{n \to \infty} x_n = a$, si la suite des distances $\|x_n - a\|$ tend vers zéro. On montre facilement que si

$$\lim_{n \to \infty} x_n = a, \qquad \lim_{n \to \infty} y_n = b, \quad \text{alors} \quad \lim_{n \to \infty} (x_n + y_n) = a + b.$$

De même, si

$$\lim_{n \to \infty} x_n = a \qquad \lim_{n \to \infty} \lambda_n = \mu,$$

alors

$$\lim_{n \to \infty} (\lambda_n x_n) = \mu a.$$

Une suite (x_n) s'appelle une *suite de Cauchy* si on a $\lim_{\substack{m \to \infty \\ n \to \infty}} \|x_m - x_n\| = 0$; ceci signifie que $\forall \varepsilon > 0$, $\exists N$ tel que

$$(m \geqslant N \text{ et } n \geqslant N) \Rightarrow \|x_m - x_n\| \leqslant \varepsilon.$$

On sait que toute suite *convergente* (c'est-à-dire qui a une limite) est une suite de Cauchy. Si la réciproque est vraie (c'est-à-dire si toute suite de Cauchy est convergente), on dit que l'espace métrique E est *complet*.

DÉFINITION. On appelle *espace de Banach* un espace vectoriel normé qui est *complet* pour la distance déduite de la norme. Si le corps de base est $\mathbf{R}$, on parle d'espace de Banach réel ; si c'est $\mathbf{C}$, on parle d'espace de Banach complexe.

Exemples d'espaces de Banach 1.2

Exemple 1. Considérons l'espace numérique réel $\mathbf{R}^n$, resp. l'espace numérique complexe $\mathbf{C}^n$. C'est un espace vectoriel réel, resp. un e.v. complexe. Considérons sur cet espace l'une des trois normes usuelles :

$$\rho_1(x) = \sum_{i=1}^{n} |x_i|,$$

$$\rho_2(x) = \sup_{1 \leqslant i \leqslant n} |x_i|,$$

$$\rho_3(x) = \sqrt{\sum_{i=1}^{n} |x_i|^2} \quad \text{(norme euclidienne)} ;$$

[on a noté $x_1, \ldots, x_n$ les coordonnées du vecteur x].

La topologie définie sur $\mathbf{R}^n$ (resp. $\mathbf{C}^n$) par l'une quelconque de ces normes est la *topologie-produit* $\mathbf{R} \times \ldots \times \mathbf{R}$ (n fois), resp. $\mathbf{C} \times \ldots \times \mathbf{C}$ (n fois). Pour qu'une suite de points ait pour limite $a = (a_1, \ldots, a_n)$ il faut et il suffit que, pour tout entier i tel que $1 \leqslant i \leqslant n$, la i-ième coordonnée des points de la suite ait pour limite a_i. Comme $\mathbf{R}$ (resp. $\mathbf{C}$) est complet, $\mathbf{R}^n$ (resp. $\mathbf{C}^n$) est aussi complet : c'est un *espace de Banach* pour l'une quelconque des normes ρ_1, ρ_2, ρ_3.

Exemple 2. Soit X un espace topologique. Soit $\mathscr{C}_b(X)$ l'ensemble de toutes les fonctions numériques $X \to \mathbf{R}$ qui sont *continues* et *bornées.* Dire que f est bornée, c'est dire que

$$\sup_{x \in X} |f(x)| \text{ est } \textit{fini}.$$

Il est clair que $\mathscr{C}_b(X)$ est un *espace vectoriel* (l'addition étant l'addition des fonctions, et le produit λf de f par un scalaire λ étant le produit de la fonction f et de la fonction constante λ). Posons

$$\|f\| = \sup_{x \in X} |f(x)|.$$

On vérifie (le faire à titre d'exercice !) que $\|f\|$ est une *norme* sur l'espace vectoriel $\mathscr{C}_b(X)$. On l'appelle la norme de la *convergence uniforme* des fonctions. De plus, cet espace est *complet* (exercice à détailler : cela tient à ce que la limite d'une suite uniformément convergente de fonctions continues bornées est continue. Ainsi $\mathscr{C}_b(X)$ est un *espace de Banach* (réel).

On pourrait faire de même avec les fonctions continues bornées à valeurs complexes : on obtiendrait un espace de Banach complexe.

Exemple 2 bis. On va généraliser l'exemple 2. Au lieu de considérer les fonctions continues bornées $X \to \mathbf{R}$, considérer les applications continues et bornées $X \to F$, où F désigne un espace de Banach donné ; par définition, $f : X \to F$ est *bornée* si

$$\|f\| = \sup_{x \in X} \|f(x)\|$$

est fini (dans le membre de droite, $\|f(x)\|$ désigne la norme de f(x) dans l'espace de Banach F). L'ensemble $\mathscr{C}_b(X ; F)$ de ces fonctions est encore un espace vectoriel (e.v. sur $\mathbf{R}$, si F est un e.v. réel ; e.v. sur $\mathbf{C}$, si F est un e.v. complexe). De plus $\|f\|$ définie ci-dessus est une norme sur cet espace vectoriel ; et il est *complet* parce que F est complet (le démontrer en exercice). Ainsi $\mathscr{C}_b(X, F)$ est un espace de Banach.

Exemple 3. Soit $\mathscr{L}_1[0, 1]$ l'espace vectoriel des fonctions numériques sur l'intervalle $[0, 1] \subset \mathbf{R}$, intégrables au sens de Lebesgue. Posons

$$\|f\| = \int_0^1 |f(t)|\, dt.$$

Ceci a toutes les propriétés d'une norme, excepté (i') : la relation $\|f\| = 0$ n'implique pas que f est identiquement nulle, mais seulement que f est nulle « presque partout » (c'est-à-dire sauf sur un ensemble de mesure nulle). Pour obtenir une vraie norme, on procède ainsi : considérons la relation d'équivalence $\mathscr{R}(f_1, f_2)$:

« f_1 et f_2 sont égales presque partout » ;

l'ensemble $L_1([0, 1])$ des classes d'équivalence a une structure d'espace vectoriel (c'est l'espace vectoriel quotient par le sous-espace vectoriel des f telles que $\|f\| = 0$). Si φ est une classe d'équivalence, on définit $\|\varphi\|$: c'est la valeur commune de $\|f\|$ pour les f de la classe φ. Alors $\|\varphi\|$ est une *norme* sur l'espace vectoriel $L_1([0, 1])$. De plus il résulte de la théorie de l'intégrale de Lebesgue que l'espace $L_1([0, 1])$ est *complet* (cela ne serait pas vrai si on utilisait l'intégrale de Riemann). Ainsi $L_1([0, 1])$ est un espace de Banach.

Exemple 3 bis. Il est analogue au précédent : on considère cette fois l'espace vectoriel $\mathscr{L}_2([0, 1])$ des « fonctions de carré intégrable » avec

$$\|f\| = \sqrt{\int_0^1 |f(t)|^2 \, dt}.$$

On passe au quotient par la relation d'équivalence $\mathscr{R}(f_1, f_2)$ comme ci-dessus. L'espace quotient $L_2([0, 1])$ est un espace de Banach.

Séries normalement convergentes dans un espace de Banach 1.3

DÉFINITION. Soit $(u_n)_{n \geqslant 0}$ une suite d'éléments $u_n \in E$, où E désigne un espace de Banach. On dit que la série de terme général u_n est normalement convergente si la série des normes

$$\sum_{n \geqslant 0} \|u_n\|,$$

qui est une série à termes $\geqslant 0$, est convergente.

THÉORÈME. *S'il en est ainsi, la série de terme général* u_n *est convergente* (c'est-à-dire : $\sum_{0 \leqslant n \leqslant p} u_n$ a une limite quand $p \to \infty$, limite notée $\sum_{n \geqslant 0} u_n$), *et on a*

$$\|\sum_{n \geqslant 0} u_n\| \leqslant \sum_{n \geqslant 0} \|u_n\|.$$

Nous ne démontrons pas ce théorème (cf. par ex. Choquet, loc. cit, p. 215–216 ; Choquet dit « absolument sommable » là où nous disons « normalement convergente »). Il est essentiel d'avoir supposé que E est un espace de Banach, car la démonstration utilise le critère de Cauchy.

Exemple. Reprenons l'espace de Banach $\mathscr{C}_b(X)$ (exemple 2 du n° 1.2). Dire que u_n est le terme général d'une série normalement convergente, où u_n est une fonction numérique continue et bornée sur l'espace topologique X, c'est dire qu'il existe une série convergente à termes $\varepsilon_n \geqslant 0$ telle que l'on ait, pour tout n,

$$|u_n(x)| \leqslant \varepsilon_n \quad \text{quel que soit } x \in X\,;$$

il suffit en effet de prendre $\varepsilon_n = \|u_n\| = \sup_{x \in X} |u_n(x)|$. On retrouve ainsi la notion usuelle de *convergence normale* d'une série de fonctions.

Applications linéaires continues 1.4

Soient E et F deux e.v. normés (tous deux sur le corps **R**, ou tous deux sur le corps **C**). On se propose de chercher un critère permettant de reconnaître si une application linéaire $f : E \to F$ est *continue* lorsqu'on munit E et F des topologies définies par leurs normes.

THÉORÈME 1.4.1. *Pour une application linéaire* $f : E \to F$, *les conditions suivantes sont équivalentes:*

(a) f *est continue en tout point de* E*;*
(b) f *est continue à l'origine* 0*;*
(c) $\|f(x)\|$ *est bornée sur la boule-unité* $\|x\| \leqslant 1$.

DÉMONSTRATION. Il est clair que (a) ⇒ (b). Montrons que (b) ⇒ (c) : supposons f continue au point 0 ; l'image réciproque f^{-1} de la boule-unité de F est un voisinage de 0 dans E, donc contient une boule $\|x\| \leqslant r$ pour un $r > 0$ convenable. Ainsi il existe $r > 0$ tel que

$$\|x\| \leqslant r \quad \text{entraîne} \quad \|f(x)\| \leqslant 1;$$

alors

$$\|x\| \leqslant 1 \quad \text{entraîne} \quad \|f(x)\| \leqslant 1/r,$$

car si on pose $y = rx$, on a

$$\|f(y)\| \leqslant 1\;; \quad \text{or} \quad \|f(y)\| = r \cdot \|f(x)\|.$$

Donc on vient de prouver que $\|f(x)\|$ est borné sur la boule-unité $\|x\| \leqslant 1$, ce qui démontre que (b) entraîne (c).

Montrons enfin que (c) entraîne (a). Si (c) est vrai, il existe $M > 0$ tel que $\|f(x)\| \leqslant M$ pour tout x tel que $\|x\| \leqslant 1$, d'où, pour tout x sans exception,

$$\|f(x)\| \leqslant M\|x\|.$$

(en effet, c'est évident si $\|x\| = 0$; et si $\|x\| = r > 0$, le vecteur $y = (1/r)x$ satisfait à $\|y\| = 1$, d'où $\|f(y)\| \leqslant M$, et $\|f(x)\| = r\|f(y)\| \leqslant rM = M\|x\|$). Montrons que, dans ces conditions, f est continue en un point quelconque $a \in E$; on a $f(x) - f(a) = f(x - a)$ puisque f est linéaire, donc il suffit que $\|x - a\| \leqslant \varepsilon/M$ pour que

$$\|f(x) - f(a)\| \leqslant M \cdot \frac{\varepsilon}{M} = \varepsilon,$$

et ceci prouve bien la continuité.

Notation. On notera $\mathscr{L}(E\,;F)$ l'ensemble de toutes les applications linéaires *continues* de E dans F. C'est évidemment un espace vectoriel (sous-espace vectoriel de l'espace de toutes les applications linéaires $E \to F$). Sur $\mathscr{L}(E\,;F)$, on pose

$$\|f\| = \sup_{\|x\| \leqslant 1} \|f(x)\|,$$

qui est fini (d'après le critère (c) du théorème 1.4.1). On a vu que, pour tout $x \in E$:

(1.4.1) $$\boxed{\|f(x)\| \leqslant \|f\| \cdot \|x\|} \qquad \text{(relation fondamentale)}$$

De plus, soit $M > 0$ tel que

(1.4.2) $$\|f(x)\| \leqslant M\,\|x\| \text{ pour tout } x \in E\,;$$

alors, pour $\|x\| \leqslant 1$, ceci donne $\|f(x)\| \leqslant M$; d'où

$$\sup_{\|x\| \leqslant 1} \|f(x)\| \leqslant M,$$

c'est-à-dire $\|f\| \leqslant M$. Ainsi $\|f\|$ *est le plus petit des nombres* $M \geqslant 0$ *tels que la relation* (1.4.2) *ait lieu.*

$\|f\|$ *est une norme sur l'espace vectoriel* $\mathscr{L}(E\,;F)$: la vérification est immédiate (la faire à titre d'exercice). Ainsi $\mathscr{L}(E\,;F)$ est un *espace vectoriel normé* ; il a donc une topologie parfaitement définie par la donnée des deux espaces normés E et F.

THÉORÈME 1.4.2. *Si* F *est un espace de Banach,* $\mathscr{L}(E\,;F)$ *est un espace de Banach.*

DÉMONSTRATION. Soit (f_n) une suite de Cauchy dans l'espace $\mathscr{L}(E\,;F)$. Pour chaque $r > 0$, considérons les restrictions des f_n à la boule $\|x\| \leqslant r$; ce sont des fonctions $f_n^{(r)}$ qui forment une suite de Cauchy dans l'espace vectoriel $\mathscr{C}_b(B'(o, r)\,;F)$ (cf. exemple 2 bis du n° 1.2). Or cet espace est complet puisque F est un espace de Banach. Donc la suite $f_n^{(r)}$ converge uniformément, dans la boule $\|x\| \leqslant r$, vers une fonction $f^{(r)}$ *continue et bornée.* Il est clair que, pour $r' < r$, la restriction de $f^{(r)}$ à la boule $\|x\| \leqslant r'$ est égale à $f^{(r')}$. Donc la collection des fonctions $f^{(r)}$ (fonctions qui se prolongent nutuellement) définit une fonction f dans tout l'espace E, telle que la restriction de f à la boule $\|x\| \leqslant r$ soit précisément $f^{(r)}$. Pour chaque $x \in E$, on a

$$f(x) = \lim_{n \to \infty} f_n(x),$$

puisque la convergence est uniforme sur chaque boule de centre 0. De là on déduit, si $x \in E$ et $y \in E$:

$$\begin{aligned} f(x + y) = \lim_{n \to \infty} f_n(x + y) &= \lim_{n \to \infty}(f_n(x) + f_n(y)) \\ &= \lim_{n \to \infty} f_n(x) + \lim_{n \to \infty} f_n(y) \\ &= f(x) + f(y), \end{aligned}$$

et on montre de la même manière que

$$f(\lambda x) = \lambda f(x).$$

Ainsi f est *linéaire*. On a vu que, sur chaque boule $\|x\| \leqslant r$, $\|f(x)\|$ est bornée ; donc f est *linéaire continue*.

Enfin $\|f - f_n\|$ tend vers 0, puisque

$$\|f - f_n\| = \sup_{\|x\| \leqslant 1} \|f(x) - f_n(x)\|$$

et que, sur la boule $\|x\| \leqslant 1$, la suite (f_n) converge uniformément vers f. On a ainsi montré que la suite de Cauchy (f_n) a une limite f, et il est donc prouvé que $\mathscr{L}(E\,;F)$ est un espace de Banach.

Composition des applications linéaires continues 1.5

Soient E, F, G trois e.v. normés, et soient $f : E \to F$, $g : F \to G$ deux applications linéaires continues. Alors $g \circ f : E \to G$ est une application linéaire continue (on sait en effet que la composée de deux applications linéaires est linéaire, et que la composée de deux applications continues est continue). Pour tout $x \in E$, on a

$$\|(g \circ f)(x)\| = \|g(f(x))\| \leqslant \|g\| \cdot \|f(x)\|\,;$$

or

$$\|f(x)\| \leqslant \|f\| \cdot \|x\|,$$

d'où finalement

$$\|(g \circ f)(x)\| \leqslant \|g\| \cdot \|f\| \cdot \|x\|.$$

D'après la propriété caractéristique de la norme d'une application linéaire (cf. 1.4), ceci entraîne

$$\boxed{\|g \circ f\| \leqslant \|g\| \cdot \|f\|}. \tag{1.5.1}$$

Isomorphismes d'espaces vectoriels normés ; normes équivalentes sur un e.v. normé 1.6

DÉFINITION. Une application $f : E \to F$ (où E et F sont des e.v. normés) est un *isomorphisme* si :

1° f est linéaire continue ;

2° il existe une g linéaire continue $F \to E$ telle que $g \circ f = \mathrm{id}_E$ (application identique de E) et $f \circ g = \mathrm{id}_F$.

Ces conditions impliquent que f est une bijection de E sur F, et que g est la bijection réciproque. D'autre part, il est clair que si f est une bijection linéaire, la bijection réciproque est linéaire. En revanche, si f est une bijection linéaire continue, il n'est pas certain que la bijection réciproque soit continue. De ces remarques, il suit une autre caractérisation des isomorphismes :

Pour que $f : E \to F$ soit un isomorphisme, il faut et il suffit que f soit un homéomorphisme (d'espaces topologiques) et soit *linéaire*.

Signalons sans démonstration un théorème, très important en Analyse, mais difficile à démontrer* :

THÉORÈME DE BANACH. *Si* E *et* F *sont des espaces de Banach, toute application linéaire continue bijective* $f : E \to F$ *est un isomorphisme.*

(Ce théorème dit que l'application réciproque $f^{-1} : F \to E$ est automatiquement continue.)

Il ne faut pas confondre *isomorphisme* et *isométrie* :

DÉFINITION. Une application $f : E \to F$ (où E et F sont des e.v. normés) est une *isométrie* si f est une bijection linéaire qui conserve la norme, c'est-à-dire

$$\|f(x)\| = \|x\| \quad \text{pour } x \in E.$$

Cette condition entraîne que $\|f(x)\|$ est bornée sur la boule-unité ; donc f est une application linéaire continue ; et le même raisonnement montre que l'application réciproque g est linéaire continue. Ainsi *toute isométrie est un isomorphisme* ; mais la réciproque n'est pas vraie : par exemple, une homothétie $x \mapsto \lambda x$ (où $\lambda \neq 0$) est un isomorphisme $E \to E$, mais n'est pas une isométrie si $|\lambda| \neq 1$.

DÉFINITION. On dit que deux normes ρ_1 et ρ_2, sur un même espace vectoriel E, sont *équivalentes* si elles définissent la même topologie.

On peut encore formuler cette définition comme suit : soit E_{ρ_1} l'e.v. normé qu'on obtient en munissant E de la norme ρ_1 ; et E_{ρ_2} celui qu'on obtient avec la norme ρ_2. L'application identique de E définit deux bijections

$$f_1 : E_{\rho_1} \to E_{\rho_2}, \qquad f_2 : E_{\rho_2} \to E_{\rho_1},$$

réciproques l'une de l'autre. Dire que ρ_1 et ρ_2 définissent la même topologie, c'est dire que f_1 et f_2 sont des *isomorphismes* d'e.v. normés. Pour cela, il faut et il suffit que f_1 et f_2 soient des applications *continues*.

Appliquons le critère de continuité d'une application linéaire (théorème 1.4.1) : la continuité de f_1 s'exprime par l'existence d'un $M > 0$ tel que

$$\rho_2(x) \leqslant M\rho_1(x) \quad \text{pour tout } x \in E\,;$$

de même, la continuité de f_2 s'exprime par l'existence d'un $M' > 0$ tel que

$$\rho_1(x) \leqslant M'\rho_2(x)$$

d'où :

PROPOSITION 1.6.1. *Pour que les normes* ρ_1 *et* ρ_2 *soient équivalentes, il faut et il suffit que le rapport* $\rho_2(x)/\rho_1(x)$ (qui est défini pour $x \neq 0$) *soit majoré et minoré par des nombres* >0.

* Voir par exemple N. Bourbaki, Espaces vectoriels topologiques, Chap. II.

THÉORÈME 1.6.2. *Sur l'espace vectoriel* $\mathbf{R}^n$, *toutes les normes sont équivalentes.*

DÉMONSTRATION. Notons

$$\|x\| = \sqrt{\sum_{i=1}^{n} |\xi_i|^2}$$

la norme euclidienne ($\xi_1, \ldots, \xi_n$ désignent les coordonnées de x). Soit ρ n'importe quelle autre norme ; montrons d'abord que $\rho : \mathbf{R}^n \to \mathbf{R}^+$ est continue (lorsqu'on munit $\mathbf{R}^n$ de la topologie-produit, qui est aussi celle définie par la norme euclidienne.) On a

$$|\rho(x) - \rho(y)| \leqslant \rho(x - y) \leqslant \sum_{i=1}^{n} |\xi_i - \eta_i| \rho(e_i),$$

en notant $(e_i, \ldots, e_n)$ la base canonique de $\mathbf{R}^n$. Cette inégalité montre que $\rho(y)$ tend vers $\rho(x)$ quand y tend vers x ; ρ est bien continue.

Sur la sphère-unité $\|x\| = 1$, qui est compacte, ρ est une fonction continue partout $\neq 0$; elle a donc une borne supérieure $M > 0$ et une borne inférieure $m > 0$. D'où aussitôt :

$$\rho(x) \leqslant M\|x\|, \qquad \rho(x) \geqslant m\|x\|,$$

ce qui prouve que ρ est équivalente à la norme euclidienne.

COROLLAIRE. Si E est un e.v. normé, toute application linéaire bijective $f : \mathbf{R}^n \to E$ est un isomorphisme. (En effet, si ρ désigne la norme sur E, $\rho \circ f$ est une norme sur $\mathbf{R}^n$; cette norme définit donc la même topologie que la norme euclidienne, d'où le résultat.)

THÉORÈME 1.6.3. *Soit* E *un e.v. normé de dimension finie. Alors* E *est un espace de Banach, et toute application linéaire de* E *dans un e.v. normé* F *est continue.*

DÉMONSTRATION. Soit n la dimension de E ; on a une application linéaire bijective $f : \mathbf{R}^n \to E$. D'après le corollaire précédent, f est un isomorphisme. Puisque $\mathbf{R}^n$ est complet, E est complet (c'est un espace de Banach). Soit maintenant $g : E \to F$ une application linéaire (F étant un e.v. normé) ; si on montre que

$$h = g \circ f : \mathbf{R}^n \to F$$

est continue, il s'ensuivra que $g = h \circ f^{-1}$ est continue.

Il reste donc seulement à montrer que toute application linéaire $h : \mathbf{R}^n \to F$ est continue. On a

$$h(\xi_1, \ldots, \xi_n) = \sum_{i=1}^{n} \xi_i h(e_i).$$

D'où

$$\|h(\xi_1, \ldots, \xi_n)\| \leqslant \sum_{i=1}^{n} |\xi_i| \cdot \|h(e_i)\|,$$

donc $h(\xi_1, \ldots, \xi_n)$ tend vers 0 quand le point $(\xi_1, \ldots, \xi_n)$ tend vers 0.

Remarque. On a des résultats analogues aux théorèmes 1.6.2 et 1.6.3 pour les espaces vectoriels *complexes* ; le rôle de $\mathbf{R}^n$ est alors joué par $\mathbf{C}^n$.

Soient E et F deux espaces vectoriels de dimension finie, $\dim E = m$, $\dim F = n$. Le choix d'une base de E et d'une base de F identifie l'espace vectoriel $\mathscr{L}(E ; F)$ à l'espace vectoriel des matrices à n lignes et m colonnes (les éléments de ces matrices étant dans le corps de base). La dimension de $\mathscr{L}(E ; F)$ est égale au produit mn.

Exemples d'espaces $\mathscr{L}(E ; F)$ 1.7

Exemple 1. Supposons que $E = \mathbf{R}$ dans le cas des e.v. réels, resp. $E = \mathbf{C}$ dans le cas des e.v. complexes. Raisonnons par exemple dans le cas réel. On va définir une *isométrie naturelle*

$$\mathscr{L}(\mathbf{R} ; F) \approx F.$$

Pour cela, nous associons à chaque $y \in F$ l'application linéaire $\lambda \mapsto \lambda y$ de $\mathbf{R}$ dans F ; elle est continue, puisque

$$\|\lambda y\| = \|y\| \cdot |\lambda|.$$

Ceci définit une application $\varphi : F \to \mathscr{L}(\mathbf{R} ; F)$, qui est évidemment linéaire. De plus la relation des normes montre que l'application linéaire $\varphi(y) : \mathbf{R} \to F$ a pour norme $\|y\|$. En sens inverse, partons d'une application linéaire continue $f : \mathbf{R} \to F$; associons-lui l'élément $f(1) \in F$; on définit ainsi une application ψ de $\mathscr{L}(\mathbf{R}, F)$ dans F, qui est évidemment linéaire. Il est immédiat que les applications φ et ψ sont réciproques l'une de l'autre ; chacune d'elles est donc une bijection. De plus, cette bijection est une *isométrie*, puisqu'on a vu que $\|\varphi(y)\| = \|y\|$. Par définition, ψ est l'isométrie naturelle de $\mathscr{L}(\mathbf{R} ; F)$ sur F.

Exemple 2. Soit E un espace de Banach réel ; alors $\mathscr{L}(E ; \mathbf{R})$ est un espace de Banach réel, qu'on appelle le *dual topologique* de E. Ses éléments sont les *formes linéaires continues* sur E.

Ne pas confondre avec le *dual algébrique*, qui comprend toute les formes linéaires continues ou non. Lorsque E est de dimension finie n, le dual topologique se confond avec le dual algébrique, il est aussi de dimension n. En général, nous noterons E^* le dual topologique de E ; E^* est muni de sa structure d'espace de Banach.

On définirait de même le dual topologique $\mathscr{L}(E ; \mathbf{C})$ d'un espace de Banach complexe.

Exemple 3. L'*algèbre* $\mathscr{L}(E ; E)$, E *étant un espace de Banach.*

On sait déjà que $\mathscr{L}(E ; E)$ est un espace de Banach. Mais on a en outre la loi de composition

$$(g, f) \mapsto g \circ f.$$

Appelons-la une *multiplication* (cette multiplication n'est pas commutative, en général). Cette multiplication jouit des propriétés suivantes :

$$\text{(1.7.1)} \qquad \begin{cases} (g_1 + g_2) \circ f = (g_1 \circ f) + (g_2 \circ f), \\ (\lambda g) \circ f = \lambda(g \circ f). \end{cases}$$

(conséquence de la définition de l'addition $g_1 + g_2$ des applications linéaires et de la multiplication d'une application linéaire g par un scalaire λ). De plus

$$\text{(1.7.2)} \qquad \begin{cases} (g \circ (f_1 + f_2) = (g \circ f_1) + (g \circ f_2) \\ g \circ (\lambda f) = \lambda(g \circ f), \end{cases}$$

parce que g est linéaire. Les relations (1.7.1) et (1.7.2) expriment que si l'on fixe f, l'application $g \mapsto g \circ f$ est linéaire, et que si l'on fixe g, l'application $f \mapsto g \circ f$ est linéaire. Une telle application est dite *bilinéaire* (cf. ci-dessous, n° 1.8).

Chaque fois que, dans un espace vectoriel A, on a défini une loi de composition interne (appelée multiplication) qui est *bilinéaire*, on dit qu'on a défini sur A une structure *d'algèbre* sur le corps de base. Cette algèbre est dite *associative* si la multiplication est associative. On voit donc ici que $\mathscr{L}(E\,;E)$ est une algèbre associative sur le corps $\mathbf{R}$ (resp. $\mathbf{C}$), suivant que E est un espace vectoriel réel ou complexe. Nous omettrons souvent le signe $\circ$ de la multiplication, écrivant simplement gf pour l'application composée $g \circ f$.

Sur l'algèbre $\mathscr{L}(E, E)$, nous avons une norme qui satisfait aux conditions habituelles d'une norme d'espace vectoriel (cf. 1.1), et en outre, d'après (1.5.1), satisfait à

$$\text{(1.7.3)} \qquad \| gf \| \leqslant \| g \| \cdot \| f \|$$

(propriété de la norme vis-à-vis de la multiplication).

Enfin, si E est un *espace de Banach* (ce qu'on va supposer jusqu'à la fin de ce numéro), $\mathscr{L}(E\,;E)$ est *complet* pour la norme (à cause du théorème 1.4.2). Nous dirons alors que $\mathscr{L}(E\,;E)$ est une *algèbre de Banach* : d'une façon précise, une algèbre de Banach A est une algèbre munie d'une norme satisfaisant à (1.7.3), et complète pour cette norme.

Remarque. Il ne faut pas croire que $\|gf\| = \|g\| \cdot \|f\|$; *par exemple* : prenons $E = R^2$; soit f (resp. g) l'application de projection sur le premier (resp. le second) axe de coordonnées. On a

$$gf = fg = 0,$$

et cependant

$$\|f\| = 1, \qquad \|g\| = 1.$$

Dans l'algèbre de Banach $\mathscr{L}(E\,;E)$, nous allons faire deux fois usage de la théorie des séries normalement convergentes.

THÉORÈME 1.7.1 ET DÉFINITION. *Si E est un espace de Banach, et si $f \in \mathscr{L}(E\,;E)$, la série*

$$\sum_{n \geqslant 0} \frac{1}{n!} f^n$$

est normalement convergente. On note $\exp f$ *sa somme.*

DÉMONSTRATION. Tout d'abord on convient que $f^0 = 1$, élément unité de l'algèbre (application identique $E \to E$). On a, d'après (1.7.3),

$$\|f^n\| \leqslant \|f\|^n,$$

donc la série des normes est majorée par :

$$\sum_{n \geqslant 0} \frac{1}{n!} \|f\|^n = \exp\|f\|$$

(fonction exponentielle usuelle d'une variable réelle), série qui converge.

C.Q.F.D.

Exercice. Démontrer que si $gf = fg$, on a $(\exp f)\cdot(\exp g) = (\exp g)\cdot(\exp f) = \exp(f+g)$; en particulier, puisque $\exp(0) = 1$, on a

$$(\exp f)\cdot(\exp(-f)) = 1,$$

donc $\exp f$ est un élément *inversible* de $\mathscr{L}(E\,;E)$.

Remarque. Ce qui précède est valable pour toute algèbre de Banach.

THÉORÈME 1.7.2. *Soit* E *un espace de Banach, et soit* $u \in \mathscr{L}(E\,;E)$ *tel que*

$$\|u\| < 1.$$

Alors $1 - u$ *a un inverse dans l'algèbre* $\mathscr{L}(E\,;E)$.

DÉMONSTRATION. La série

$$\sum_{n \geqslant 0} u^n = 1 + u + \ldots + u^n + \ldots$$

est normalement convergente, puisque $\|u^n\| \leqslant \|u\|^n$, et que la série géométrique $\sum_{n \geqslant 0} \|u\|^n$ est convergente en vertu de l'hypothèse $\|u\| < 1$. Soit v la somme $\sum_{n \geqslant 0} u^n$. On voit tout de suite que

$$vu = uv$$

est la somme de la série $\sum_{n \geqslant 1} u^n$. Alors

$$v(1-u) = (1-u)v = 1,$$

donc v est l'inverse de $1 - u$.

Remarque. Ce théorème est aussi valable dans toute algèbre de Banach.
Voici une conséquence du théorème 1.7.2. :

THÉORÈME 1.7.3. *Soient* E *et* F *deux espaces de Banach. Notons* Isom $(E\,;F)$ *le sous-ensemble de* $\mathscr{L}(E\,;F)$ *formé des isomorphismes* $E \to F$ (cf. Déf. du §1.6).
Alors :

(a) Isom $(E\,;F)$ *est ouvert dans* $\mathscr{L}(E\,;F)$;

(b) *l'application* $u \mapsto u^{-1}$ *de* Isom $(E\,;F)$ *dans* $\mathscr{L}(F\,;E)$ *est continue.*

DÉMONSTRATION. Observons d'abord que l'ensemble Isom $(E\,;F)$ peut être vide (si E et F ne sont pas isomorphes !). Dans ce cas, le théorème est trivialement vrai. Si Isom $(E\,;F)$ n'est pas vide, prenons un $u_0 \in$ Isom $(E\,;F)$. Pour prouver (a), nous devons montrer que tout $u \in \mathscr{L}(E\,;F)$ *assez voisin* de u_0 est encore un isomorphisme. Or, pour que $u : E \to F$ soit un isomorphisme, il faut et il suffit que

$$(u_0)^{-1}u : E \to E$$

soit un isomorphisme ; cherchons donc une condition suffisante pour que $(u_0)^{-1}u$ soit un isomorphisme, c'est-à-dire soit un élément inversible de $\mathscr{L}(E\,;E)$. Posons :

$$(u_0)^{-1}u = 1 - v.$$

Il *suffit* que $\|v\| < 1$, d'après le théorème 1.7.2. Or $v = 1 - u_0^{-1}u = u_0^{-1}(u_0 - u)$, d'où

$$\|v\| \leqslant \|u_0^{-1}\|\,\|u - u_0\|. \tag{1.7.4}$$

Donc, si

$$\|u - u_0\| < \frac{1}{\|u_0^{-1}\|},$$

on est sûr que $\|v\| < 1$, donc que u est un isomorphisme. Ceci prouve bien que tout u assez voisin de u_0 est un isomorphisme.
(*Attention* ! Ne pas croire que $\|u_0^{-1}\| = 1/\|u_0\|$).

Il reste à démontrer (b). On a

$$u^{-1} = (u_0(1 - v))^{-1} = (1 - v)^{-1}(u_0)^{-1},$$

d'où

$$u^{-1} - (u_0)^{-1} = [(1 - v)^{-1} - 1](u_0)^{-1}\,; \tag{1.7.5}$$

or

$$(1 - v)^{-1} = \sum_{n \geqslant 0} v^n, \qquad \text{d'où} \qquad (1 - v)^{-1} - 1 = \sum_{n \geqslant 1} v^n,$$

$$\|(1 - v)^{-1} - 1\| \leqslant \sum_{n \geqslant 1} \|v\|^n = \frac{\|v\|}{1 - \|v\|}.$$

Ainsi (1.7.5) entraîne

$$\|u^{-1} - (u_0)^{-1}\| \leqslant \|u_0^{-1}\| \cdot \frac{\|v\|}{1 - \|v\|}. \tag{1.7.6}$$

Lorsque u tend vers u_0, $\|v\|$ tend 0 d'après (1.7.4), donc u^{-1} tend vers $(u_0)^{-1}$ d'après (1.7.6). Ceci prouve bien que u^{-1} est une fonction continue de u lorsque u parcourt Isom(E ; F). Le théorème est ainsi démontré.

Remarque. Lorsque E et F sont de même dimension finie n, et qu'on identifie $\mathscr{L}(E\,;F)$ à l'espace des matrices à n lignes et n colonnes, on sait exprimer qu'une matrice f est inversible : la condition est que son déterminant $\det f$ soit $\neq 0$. L'application $f \mapsto \det f$ de $\mathscr{L}(E\,;F)$ dans $\mathbf{R}$ (resp. $\mathbf{C}$) étant continue, l'image réciproque du complémentaire de 0, qui est Isom (E ; F), est *ouverte.* Ceci donne, dans ce cas particulier, une nouvelle démonstration de la partie (a) du théorème. Le calcul de la matrice inverse permettrait de vérifier (b) dans ce cas particulier.

Applications multilinéaires continues 1.8

D'abord, un rappel d'algèbre : soient $E_1, \ldots, E_n$ et F des espaces vectoriels ; une application

$$f : E_1 \times \ldots \times E_n \to F$$

est dite *multilinéaire* (bilinéaire si $n = 2$, trilinéaire si $n = 3$) si, pour chaque entier $k \in [1, n]$, et pour chaque système d'éléments $a_i \in E_i$ $(i \neq k)$, l'application « partielle »

$$x_k \mapsto f(a_1, \ldots, a_{k-1}, x_k, a_{k+1}, \ldots, a_n)$$

de E_k dans F est linéaire. Autrement dit, lorsqu'on fixe toutes les variables sauf une, f doit dépendre linéairement de la variable restante. S'il en est ainsi on a $f(x_1, \ldots, x_n) = 0$ dès que l'un au moins des x_i est nul ; en particulier f s'annule à l'origine $(0, \ldots, 0)$. Observons que si f est multilinéaire, on a

$$f(\lambda_1 x_1, \ldots, \lambda_n x_n) = (\lambda_1 \ldots \lambda_n) f(x_1, \ldots, x_n). \tag{1.8.1}$$

Exemple. Prenons pour $E_1, \ldots, E_n$ et F le corps des scalaires ; le *produit* de n éléments du corps

$$\lambda_1 \lambda_2 \ldots \lambda_n,$$

considéré comme fonction de $\lambda_1, \lambda_2, \ldots, \lambda_n$, est une fonction multilinéaire.

Supposons maintenant que $E_1, \ldots, E_n$, F soient des espaces vectoriels normés. Alors $E_1 \times \ldots \times E_n$ a une structure d'espace topologique (comme produit d'espaces topologiques) ; cela a donc un sens de se demander si une application $f : E_1 \times \ldots \times E_n \to F$ est continue. Le théorème 1.4.1 se généralise comme suit :

THÉORÈME 1.8.1. *Soient* $E_1, \ldots, E_n$, F *des e.v. normés, et soit* $f : E_1 \times \ldots \times E_n \to F$ *une application multilinéaire. Les conditions suivantes sont équivalentes :*

(a) f *est continue en tout point* de $E_1 \times \ldots \times E_n$;

(b) f *est continue à l'origine* $(0, \ldots, 0) \in E_1 \times \ldots \times E_n$;

(c) $\|f(x_1, \ldots, x_n)\|$ *est borné sur le produit des boules-unité*

$$\|x_1\| \leqslant 1, \ldots, \|x_n\| \leqslant 1.$$

La démonstration procède comme celle du théorème 1.4.1. Il est évident que (a) $\Rightarrow$ (b). Pour montrer que (b) $\Rightarrow$ (c), on observe que si f est continue à l'origine, l'image réciproque de la boule-unité de f est un voisinage de $(0, \ldots, 0)$ dans $E_1 \times \ldots \times E_n$, donc il existe $r > 0$ tel que

$$(\|x_i\| \leqslant r \text{ pour tout } i) \Rightarrow \|f(x_1, \ldots, x_n)\| \leqslant 1.$$

Compte tenu de (1.8.1), on en déduit

$$(\|x_i\| \leqslant 1 \text{ pour tout } i) \Rightarrow \|f(x_1, \ldots, x_n)\| \leqslant \frac{1}{r^n},$$

ce qui prouve (c).

Supposons que f vérifie (c) ; soit $M > 0$ tel que

$$(\|x_i\| \leqslant 1 \text{ pour tout } i) \Rightarrow \|f(x_1, \ldots, x_n)\| \leqslant M.$$

On a alors, quels que soient les x_i :

$$\|f(x_1, \ldots, x_n)\| \leqslant M\|x_1\| \ldots \|x_n\|. \tag{1.8.2}$$

Montrons que, dans ces conditions, f est continue en un point arbitraire $(a_1, \ldots, a_n)$, ce qui prouvera que (c) $\Rightarrow$ (a). Formons la différence

$$\begin{aligned} &f(x_1, \ldots, x_n) - f(a_1, \ldots, a_n) \\ &\quad = f(x_1 - a_1, x_2, \ldots, x_n) + f(a_1, x_2 - a_2, x_3, \ldots, x_n) \\ &\quad + \ldots + f(a_1, \ldots, a_{n-1}, x_n - a_n). \end{aligned}$$

[Vérification immédiate, puisque f est une fonction additive de chaque variable séparément.] La norme du premier membre est majorée par la somme des normes des termes du second membre ; donc, compte tenu de (1.8.2) :

$$\begin{aligned} &\|f(x_1, \ldots, x_n) - f(a_1, \ldots, a_n)\| \\ &\quad \leqslant M\|x_1 - a_1\| \cdot \|x_2\| \ldots \|x_n\| + M\|x_2 - a_2\| \cdot \|a_1\| \cdot \|x_3\| \ldots \|x_n\| \\ &\quad + \ldots + M\|x_n - a_n\| \cdot \|a_1\| \ldots \|a_{n-1}\|. \end{aligned} \tag{1.8.3}$$

Supposons $\|x_i - a_i\| \leqslant \varepsilon$ pour tout i ; alors $\|x_i\| \leqslant \|a_i\| + \varepsilon$, donc il existe un nombre $A > 0$ tel que

$$(\|x_i - a_i\| \leqslant \varepsilon \text{ pour tout } i) \Rightarrow \|x_i\| \leqslant A \text{ pour tout } i.$$

L'inégalité (1.8.3) entraîne donc

$$\|f(x_1, \ldots, x_n) - f(a_1, \ldots, a_n)\| \leqslant MA^{n-1}\left(\sum_{i=1}^{n} \|x_i - a_i\|\right) \leqslant nMA^{n-1}\varepsilon \tag{1.8.4}$$

dès que $\|x_i - a_i\| \leqslant \varepsilon$ pour tout i. Visiblement, A peut être choisi indépendamment de $\varepsilon > 0$ dès que ε est assez petit. Alors (1.8.4) montre que $f(x_1, \ldots, x_n)$ tend vers $f(a_1, \ldots, a_n)$ quand simultanément x_1 tend vers $a_1, \ldots, x_n$ tend vers a_n. Donc f est *continue* au point $(a_1, \ldots, a_n)$, et la démonstration est achevée.

Notation. On notera $\mathscr{L}(E_1, \dots, E_n; F)$ l'ensemble des applications *n*-linéaires continues $E_1 \times \dots \times E_n \to F$. C'est évidemment un sous-espace vectoriel de l'espace vectoriel de toutes les applications $E_1 \times \dots \times E_n \to F$. Pour $f \in \mathscr{L}(E_1 \times \dots \times E_n; F)$ on posera

$$\|f\| = \sup \|f(x_1, \dots, x_n)\|$$

lorsque $x_1, \dots, x_n$ parcourent les boules-unité :

$$\|x_1\| \leqslant 1, \dots, \|x_n\| \leqslant 1.$$

On a alors, d'après (1.8.2) :

(1.8.5)
$$\|f(x_1, \dots, x_n)\| \leqslant \|f\| \cdot \|x_1\| \dots \|x_n\|,$$

et $\|f\|$ est le plus petit des $M > 0$ tels que (1.8.2) ait lieu.

Exercice. Vérifier que $\|f\|$ est bien une *norme* sur l'espace vectoriel $\mathscr{L}(E_1, \dots, E_n; F)$.

Autre exercice. Montrer que si F est un espace de Banach, l'e.v. normé $\mathscr{L}(E_1, \dots, E_n; F)$ est un espace de Banach. (Procéder comme dans le cas $n = 1$, cf. ci-dessus n° 1.4.)

Exemple d'une application bilinéaire continue. Soient E, F, G trois espaces vectoriels normés. Considérons l'application de composition :

$$\varphi : \mathscr{L}(F; G) \times \mathscr{L}(E; F) \to \mathscr{L}(E; G),$$

définie par

$$\varphi(g, f) = g \circ f.$$

On a déjà vu qu'elle est *bilinéaire*. On sait de plus que (cf. (1.5.1))

$$\|g \circ f\| \leqslant \|g\| \cdot \|f\| ;$$

donc si $\|f\| \leqslant 1$ et $\|g\| \leqslant 1$, on a $\|g \circ f\| \leqslant 1$. Ceci montre que l'application bilinéaire φ est continue, et que sa norme $\|\varphi\|$ est $\leqslant 1$.

L'isométrie naturelle $\mathscr{L}(E, F; G) \approx \mathscr{L}(E; \mathscr{L}(F; G))$ 1.

Nous allons d'abord définir une application

$$\varphi : \mathscr{L}(E, F; G) \to \mathscr{L}(E; \mathscr{L}(F; G))$$

comme suit : Soit $f \in \mathscr{L}(E, F; G)$; $f(x, y)$ est une fonction de deux variables $x \in E$ et $y \in F$; si on fixe x, l'application $y \mapsto f(x, y)$ est une application linéaire de F dans G, qu'on va noter f_x (application partielle). On a

$$\|f_x(y)\| = \|f(x, y)\| \leqslant \|f\| \cdot \|x\| \cdot \|y\|,$$

donc

(1.9.1)
$$\|f_x\| \leqslant \|f\| \cdot \|x\|,$$

et ceci montre notamment que f_x est une application linéaire *continue* (puisque sa norme est finie). Alors $x \mapsto f_x$ est une application $g : E \to \mathscr{L}(F\,;G)$; on vérifie aussitôt qu'elle est linéaire. De plus (1.9.1) s'écrit

$$\|g(x)\| \leqslant \|f\| \cdot \|x\|,$$

donc g est continue, et $\|g\| \leqslant \|f\|$. On a ainsi associé à toute $f \in \mathscr{L}(E, F\,;G)$ une $g \in \mathscr{L}(E\,;\mathscr{L}(F\,;G))$, qui par définition sera $\varphi(f)$. Ceci définit l'application φ. Il est immédiat que φ est linéaire. De plus, puisque φ transforme f en g, et que $\|g\| \leqslant \|f\|$, l'application linéaire φ est de norme $\leqslant 1$: $\|\varphi\| \leqslant 1$.

Maintenant, nous allons définir en sens inverse une application

$$\psi : \mathscr{L}(E\,;\mathscr{L}(F\,;G)) \to \mathscr{L}(E, F\,;G).$$

Pour cela, partons d'une application linéaire continue

$$g : E \to \mathscr{L}(F\,;G).$$

Pour $x \in E$, $g(x)$ est une application linéaire continue $F \to G$; donc, pour $x \in E$, $y \in F$, $g(x) \cdot y$ est une application *bilinéaire*

$$f : E \times F \to G.$$

De plus

$$\|g(x)\| \leqslant \|g\| \cdot \|x\|,$$

donc

$$\|f(x, y)\| = \|g(x) \cdot y\| \leqslant \|g(x)\| \cdot \|y\| \leqslant \|g\| \cdot \|x\| \cdot \|y\|,$$

ce qui prouve que f est bilinéaire *continue*, et que

$$\|f\| \leqslant \|g\|.$$

Ainsi chaque $g \in \mathscr{L}(E : \mathscr{L}(F\,;G))$ définit une $f \in \mathscr{L}(E, F\,;G)$; par définition, f sera $\psi(g)$. Ceci définit l'application ψ. Il est immédiat que ψ est linéaire. De plus, puisque ψ transforme g en f, et que $\|f\| \leqslant \|g\|$, l'application linéaire ψ est de norme $\leqslant 1$.

Maintenant il est clair que les deux applications φ et ψ sont réciproques l'une de l'autre. Ainsi $\psi \circ \varphi$ est l'application identique de $\mathscr{L}(E \times F\,;G)$, sa norme est donc 1. D'où

$$1 = \|\psi \circ \varphi\| \leqslant \|\psi\| \cdot \|\varphi\|,$$

et comme

$$\|\varphi\| \leqslant 1, \qquad \|\psi\| \leqslant 1,$$

on en conclut que

$$\|\varphi\| = 1, \qquad \|\psi\| = 1.$$

Par suite φ conserve la norme : c'est une *isométrie*.

C.Q.F.D.

2. APPLICATIONS DIFFÉRENTIABLES

Définition d'une application différentiable

Dans ce qui suit, on se donne deux espaces de Banach E et F, et un ouvert non vide $U \subset E$. On considère des applications $f : U \to F$. Chaque point $a \in U$ définit une *relation d'équivalence* dans l'ensemble de ces fonctions, comme suit :

DÉFINITION. On dit que $f_1 : U \to F$ et $f_2 : U \to F$ sont *tangentes* en un point $a \in U$ si la quantité

$$m(r) = \sup_{\|x-a\| \leqslant r} \|f_1(x) - f_2(x)\|,$$

qui est définie pour $r > 0$ assez petit (puisque U est ouvert), satisfait à la condition

$$\lim_{\substack{r \to 0 \\ r > 0}} \frac{m(r)}{r} = 0, \tag{2.1.1}$$

condition qu'on écrit aussi

$$m(r) = o(r). \tag{2.1.2}$$

Le lecteur vérifiera que la relation : « f_1 et f_2 sont tangentes en a » est bien une relation d'équivalence. On a, en particulier, la notion d'une f *tangente à* 0 au point a.

La condition (2.1.2) implique que la fonction $f_1 - f_2$ est *continue* au point a, et prend la valeur 0 au point a. Ainsi : deux fonctions tangentes en a prennent la même valeur au point a, et si l'une d'elles est continue en a, l'autre est aussi continue en a.

Exemple. Soit g une application linéaire $E \to F$ (continue ou non). Posons

$$f(x) = g(x - a),$$

et cherchons si f est tangente à 0 au point a. On a

$$m(r) = \|g\| \cdot r,$$

donc si $m(r)/r$ tend vers 0 avec r, on a $\|g\| = 0$, donc g est *identiquement nulle*. De là il résulte qu'étant donnée une application $f : U \to F$, il existe *au plus une* application linéaire $g : E \to F$ telle que les applications

$$x \mapsto f(x) - f(a)$$

et

$$x \mapsto g(x - a)$$

soient tangentes en a. De plus, si une telle g existe, la continuité de f en a entraîne la continuité de g à l'origine (donc partout, puisque g est linéaire), et réciproquement.

DÉFINITION. On dit que $f : U \to F$ est *différentiable* au point $a \in U$ si les conditions suivantes sont vérifiées :

(i) f est continue au point a ;

(ii) il existe une g linéaire $E \to F$ telle que les applications $x \mapsto f(x) - f(a)$ et $x \mapsto g(x - a)$ soient tangentes au point a.

Cette condition s'exprime ainsi :

$$\|f(x) - f(a) - g(x - a)\| = o(\|x - a\|). \tag{2.1.3}$$

Si f est différentiable au point a, l'unique g linéaire qu'elle définit est *continue*, d'après la remarque ci-dessus. C'est un élément de $\mathscr{L}(E ; F)$, qu'on notera $f'(a)$, et qu'on appellera la *dérivée* de l'application f au point a.

Une définition équivalente est celle-ci : f *est différentiable* au point $a \in U$ *s'il existe une* $g \in \mathscr{L}(E ; F)$ *telle que* (2.1.3) *ait lieu.* En effet, la continuité de g entraîne alors la continuité de f au point a.

Nous récrivons (2.1.3) avec la notation $f'(a)$:

$$\|f(x) - f(a) - f'(a)(x - a)\| = o(\|x - a\|). \tag{2.1.4}$$

Exemple. Soit F un espace de Banach réel, et soit U un ouvert de $\mathbf{R}$; alors $f : U \to F$ est une fonction d'une variable réelle. Compte tenu de l'isométrie canonique $\mathscr{L}(\mathbf{R} ; F) \approx F$, la différentiabilité de f au point a équivaut à l'existance d'un élément $c \in F$ tel que

$$\|f(x) - f(a) - (x - a)c\| = o(|x - a|) ;$$

autrement dit, le quotient

$$\frac{f(x) - f(a)}{x - a} \qquad \text{(pour } x \neq a\text{)}$$

doit avoir une limite $c \in F$ lorsque x tend vers a. On retrouve la définition usuelle de la dérivée d'une fonction d'une variable réelle, à valeurs dans un espace de Banach F.

Si cette limite c est notée $f'(a)$, l'application linéaire

$$t \to tf'(a)$$

de $\mathbf{R}$ dans F est l'élément de $\mathscr{L}(\mathbf{R} ; F)$ qui lui correspond dans l'isométrie naturelle $F \approx \mathscr{L}(\mathbf{R} ; F)$; c'est l'élément de $\mathscr{L}(\mathbf{R} ; F)$ que nous avions noté $f'(a)$ dans le cas général. Il n'y a pas d'inconvénient majeur à utiliser la même notation $f'(a)$ pour désigner l'élément de F et l'élément correspondant de $\mathscr{L}(\mathbf{R} ; F)$.

On a un phénomène analogue pour $\mathscr{L}(\mathbf{C} ; F)$ lorsque F est un espace de Banach complexe.

Revenons au cas général : $E \supset U \xrightarrow{f} F$.

DÉFINITION. On dit que f est *différentiable dans* U si f est différentiable en tout point de U.

Alors l'élément $f'(a) \in \mathscr{L}(E ; F)$ dépend de $a \in U$. Nous avons donc une applicatio $a \mapsto f'(a)$, que nous noterons simplement f' :

$$f' : U \to \mathscr{L}(E ; F).$$

C'est, par définition, *l'application dérivée* de l'application différentiable $f : U \to F$. O notera bien que l'application dérivée f' ne prend pas ses valeurs dans le même espac F que l'application f. Rappelons toutefois que si $E = \mathbf{R}$ (resp. $\mathbf{C}$), F étant un espace d Banach réel (resp. complexe), alors on peut identifier $\mathscr{L}(\mathbf{R} ; F)$ à F (resp. $\mathscr{L}(\mathbf{C} ; F)$ à F) Donc, pour une fonction f d'une variable réelle (resp. complexe), on peut identifie l'application dérivée f' à une application $U \to F$.

DÉFINITION. On dit que $f : U \to F$ est *continûment différentiable*, ou *de classe* C^1, si
1) f est différentiable dans U, c'est-à-dire différentiable en tout point de U ;
2) l'application dérivée $f' : U \to \mathscr{L}(E ; F)$ est *continue*.

(Ne pas oublier que $\mathscr{L}(E ; F)$ est muni d'une norme qui en fait un espace de Banach ; donc U et $\mathscr{L}(E ; F)$ sont des espaces topologiques.)

Remarque sur la notion de différentiabilité : soit toujours f une application continue $U \to F$, où F est un espace de Banach, et U un ouvert d'un espace de Banach E. Remplaçons la norme de E par une norme équivalente (cf. 1.6) ; notons $\|x\|_1$ la nouvelle norme d'un $x \in E$, l'ancienne étant notée $\|x\|$; remplaçons de même la norme de F par une norme équivalente. La topologie de E n'est pas changée, celle de F non plus : U reste ouvert, et f reste une application continue.

PROPOSITION 2.1.1. *Si f est différentiable au point $a \in U$ pour les anciennes normes, f est aussi différentiable au point a pour les nouvelles normes, et sa dérivée est la même.*

En effet, le fait que f est différentiable au point a et a pour dérivée un élément $g \in \mathscr{L}(E ; F)$ s'exprime par les conditions

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{\|f(x) - f(a) - g(x-a)\|}{\|x - a\|} = 0.$$

Puisque les normes $\|\ \|$ et $\|\ \|_1$ sont équivalentes dans E, on a

$$\frac{1}{\|x - a\|_1} \leqslant M \frac{1}{\|x - a\|},$$

M étant un nombre fixe. Puisque les normes $\|\ \|$ et $\|\ \|_1$ dans F sont équivalentes, on a

$$\|f(x) - f(a) - g(x-a)\|_1 \leqslant M'\|f(x) - f(a) - g(x-a)\|,$$

M' étant un nombre fixe. D'où :

$$\frac{\|f(x) - f(a) - g(x-a)\|_1}{\|x - a\|_1} < MM' \frac{\|f(x) - f(a) - g(x-a)\|}{\|x - a\|}.$$

Puisque par hypothèse le second membre tend vers 0 quand x tend vers a (en restant $\neq a$), il en est de même du premier membre.

C.Q.F.D.

Dérivée d'une fonction composée 2.2

Soient E, F, G trois espaces de Banach, soit U un ouvert de E, et soit V un ouvert de F. On considère deux applications continues

$$f : U \to F, \qquad g : V \to G,$$

et un point $a \in U$. On suppose que $b = f(a) \in F$ est, en fait, dans V. Alors $f^{-1}(V) \subset U$ est un ouvert de E qui contient a ; dans cet ouvert U′, l'application composée

$$g \circ f : U' \to G$$

est définie.

THÉORÈME 2.2.1. *Sous les hypothèses précédentes, si* f *est différentiable au point* a, *et si* g *est différentiable au point* $b = f(a)$, *alors* $h = g \circ f$ *est différentiable au point* a, *et on a*

$$h'(a) = g'(b) \circ f'(a). \tag{2.2.1}$$

Autrement dit, l'application linéaire $h'(a) : E \to G$ est composée de l'application linéaire $f'(a) : E \to F$ et de l'application linéaire $g'(f(a)) : F \to G$.

DÉMONSTRATION. On a par hypothèse

$$f(x) = f(a) + f'(a) \cdot (x - a) + \varphi(x - a), \tag{2.2.2}$$

où φ est une application tangente à 0 à l'origine, c'est-à-dire

$$\|\varphi(x - a)\| = o(\|x - a\|).$$

De même, on a par hypothèse

$$g(y) = g(b) + g'(b) \cdot (y - b) + \psi(y - b), \tag{2.2.3}$$

avec $\|\psi(y - b)\| = o(\|y - b\|)$.
Evaluons alors $h(x) - h(a) = g(f(x)) - g(f(a))$; appliquons (2.2.3) en y remplacant y par f(x), et b par f(a). On obtient

$$h(x) - h(a) = g'(f(a)) \cdot (f(x) - f(a)) + \psi(f(x) - f(a)).$$

Dans cette relation, remplacons $f(x) - f(a)$ par sa valeur tirée de (2.2.2), en tenant compte du fait que $g'(f(a))$ est une fonction *linéaire* $F \to G$:

$$\begin{aligned} h(x) - h(a) = (g'(f(a)) \circ f'(a)) \cdot (x - a) \\ + g'(f(a)) \cdot \varphi(x - a) \\ + \psi(f(x) - f(a)). \end{aligned}$$

Pour prouver que h est différentiable au point a et a pour dérivée $g'(f(a)) \circ f'(a)$, il suffit de montrer que le second et le troisième terme du membre de droite sont tangents à 0, c'est-à-dire

(2.2.4) $$\|g'(f(a)) \cdot \varphi(x - a)\| = o(\|x - a\|),$$

(2.2.5) $$\|\psi(f(x) - f(a))\| = o(\|x - a\|).$$

Or (2.2.4) résulte de

$$\|g'(f(a)) \cdot \varphi(x - a)\| \leqslant \|g'(f(a))\| \cdot \|\varphi(x - a)\| \; ;$$

(2.2.5) résulte du fait que

$$\|\psi(f(x) - f(a))\| = o(\|f(x) - f(a)\|)$$

et du fait que, en vertu de (2.2.2), l'inégalité

$$\|f(x) - f(a)\| \leqslant M \cdot \|x - a\|$$

(où M est donné $> \|f'(a)\|$) a lieu pour $\|x - a\|$ assez petit, comme cela résulte de (2.2.2).
Le théorème 2.2.1 est ainsi démontré.

Linéarité de la dérivée

Considérons le cas général : U est un ouvert d'un espace de Banach E, F est un espace de Banach. Soient f et g deux applications $U \to F$; leur *somme* h est l'application $h : U \to F$ définie par

$$h(x) = f(x) + g(x) \text{ (addition dans } F).$$

De même, le produit λf de f par un scalaire λ est l'application $k : U \to F$ définie par $k(x) = \lambda \cdot f(x)$.

PROPOSITION 2.3.1. Avec les notations précédentes, si f et g sont différentiables au point a, $h = f + g$ est différentiable au point a, et on a

$$h'(a) = f'(a) + g'(a).$$

Si f est différentiable au point a, $k = \lambda f$ est différentiable au point a, et on a

$$k'(a) = \lambda f'(a).$$

Autrement dit, l'ensemble des $f : U \to F$ qui sont différentiables au point $a \in U$ est un *sous-espace vectoriel* V_a de l'espace vectoriel de toutes les applications $U \to F$, et l'application $f \mapsto f'(a)$ est une application *linéaire* de V_a dans $\mathscr{L}(E\,;F)$. On verrait de même que l'ensemble des applications $f : U \to F$ qui sont *de classe* C^1 (dans U) est un sous-espace vectoriel du précédent.

Dérivées de fonctions particulières

PROPOSITION 2.4.1. *Si* $f : U \to F$ *est une application constante, elle est différentiable et sa dérivée* $f'(x)$ *est nulle, quel que soit* $x \in U$.

C'est évident à partir des définitions.

On verra plus loin (§3) que réciproquement, si f est différentiable et si $f'(x) = 0$ pour tout $x \in U$, si en outre U est *connexe*, alors f est constante dans U.

PROPOSITION 2.4.2. *Si* $f : U \to F$ *est la restriction d'une application linéaire continue* $E \to F$ (qu'on notera encore f), *elle est différentiable, et on a*

$$f'(x) = f \quad \text{pour tout } x \in U$$

[la dérivée est donc constante ; ne pas oublier que cette constante est un élément de $\mathscr{L}(E\,;F)$]. C'est encore évident à partir des définitions.

On va maintenant étudier la dérivée d'une *application bilinéaire continue*

$$f : E_1 \times E_2 \to F,$$

E_1, E_2, F désignant trois espaces de Banach. Mais d'abord, pour nous trouver dans le cadre voulu, nous devons définir sur $E_1 \times E_2$ une structure d'espace de Banach. Pour cela, on considère d'abord sur $E_1 \times E_2$ la structure d'*espace vectoriel*, produit des e.v. E_1 et E_2 : les opérations d'espace vectoriel sont définies par les formules

3

$$\begin{cases} (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2) \\ \lambda(x_1, x_2) = (\lambda x_1, \lambda x_2)\,; \end{cases}$$

en particulier, $(x_1, x_2) = (x_1, 0) + (0, x_2)$. Il reste à préciser quelle *norme* on choisit sur l'espace vectoriel $E_1 \times E_2$; on posera

$$\|(x_1, x_2)\| = \|x_1\| + \|x_2\|\,; \tag{2.4.1}$$

on vérifie que ceci est bien une norme, qu'elle définit sur $E_1 \times E_2$ la *topologie-produit* de E_1 et de E_2, et que pour cette norme $E_1 \times E_2$ est complet (parce que E_1 et E_2 sont complets, par hypothèse). *Remarque :* au lieu de la norme définie par (2.4.1), on pourrait prendre n'importe quelle autre norme équivalente à celle-là, par exemple $\sup(\|x_1\|, \|x_2\|)$.

THÉORÈME 2.4.3. *Si* $f : E_1 \times E_2 \to F$ *est bilinéaire continue,* f *est différentiable, et sa dérivée au point* (a_1, a_2) [avec $a_1 \in E_1$, $a_2 \in E_2$] *est définie par*

$$f'(a_1, a_2)\cdot(h_1, h_2) = f(h_1, a_2) + f(a_1, h_2)\,; \tag{2.4.2}$$

dans cette formule, on a $h_1 \in E_1$, $h_2 \in E_2$; le premier membre désigne la valeur de $f'(a_1, a_2) \in \mathscr{L}(E_1 \times E_2\,; F)$ sur le vecteur $(h_1, h_2) \in E_1 \times E_2$.

DÉMONSTRATION

$$f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) = f(h_1, a_2) + f(a_1, h_2) + f(h_1, h_2),$$

et tout sera donc démontré si on prouve que

$$\|f(h_1, h_2)\| = o(\|(h_1, h_2)\|).$$

Or $\|(h_1, h_2)\| = \|h_1\| + \|h_2\|$, tandis que

$$\|f(h_1, h_2)\| \leqslant \|f\| \cdot \|h_1\| \cdot \|h_2\| \leqslant \|f\| \cdot (\|h_1\| + \|h_2\|)^2.$$

Et il est clair que $(\|h_1\| + \|h_2\|)^2 = o(\|h_1\| + \|h_2\|)$.

C.Q.F.D.

Généralisation. Au lieu d'un produit de *deux* espaces de Banach E_1 et E_2, considérons un produit

$$E_1 \times \ldots \times E_n,$$

n étant un entier quelconque. Sur ce produit, on prend la structure d'espace vectoriel produit, et la norme

$$\|(x_1, \ldots, x_n)\| = \sum_{i=1}^{n} \|x_i\|.$$

Elle définit la topologie-produit. Soit

$$f : E_1 \times \ldots \times E_n \to F$$

une application multilinéaire continue. Alors le théorème 2.4.3. se généralise : f est différentiable, et on a

$$(2.4.3) \qquad f'(a_1, \ldots, a_n) \cdot (h_1, \ldots, h_n) = f(h_1, a_2, \ldots, a_n) + f(a_1, h_2, a_3, \ldots, a_n) + \ldots + f(a_1, \ldots, a_{n-1}, h_n).$$

[Dans $f(a_1, \ldots, a_n)$, on remplace successivement chaque a_i par h_i, sans toucher aux autres a_j ; on fait la somme des termes obtenus, et cela donne le second membre de (2.4.3.)]. La démonstration se fait par récurrence sur n ; elle est laissée au lecteur à titre d'exercice.

Voici un dernier exemple : dans le théorème 1.7.3, on a défini une application *continue*

$$u \mapsto u^{-1}$$

de Isom$(E ; F)$ (ouvert dans l'espace de Banach $\mathscr{L}(E ; F)$) sur Isom$(F ; E)$ (ouvert dans l'espace de Banach $\mathscr{L}(F ; E)$). Soit φ cette application ; on a donc $\varphi(u) = u^{-1}$. On peut considérer φ comme prenant ses valeurs dans l'espace de Banach $\mathscr{L}(F ; E)$, et cela a donc un sens de se demander si φ est différentiable. Sa dérivée sera alors un élément de

$$\mathscr{L}(\mathscr{L}(E ; F) ; \mathscr{L}(F ; E)).$$

THÉORÈME 2.4.4. *Avec les notations précédentes, φ est de classe C^1 dans l'ouvert* Isom$(E ; F) \subset \mathscr{L}(E ; F)$, *et sa dérivée est donnée par*

$$(2.4.4) \qquad \varphi'(u) \cdot h = -u^{-1} \circ h \circ u^{-1} \quad \text{pour } h \in \mathscr{L}(E ; F).$$

DÉMONSTRATION. Regardons d'abord ce que signifie le second membre de (2.4.4.). Le signe $\circ$ désigne ici la composition des applications linéaires continues :

$$F \xrightarrow{u^{-1}} E \xrightarrow{h} F \xrightarrow{u^{-1}} E,$$

de sorte que le second membre est un élément de $\mathscr{L}(F ; E)$, comme il se doit.

Pour démontrer (2.4.4), donnons à u un « accroissement » h :

$$\begin{aligned}\varphi(u+h)-\varphi(u) &= (u+h)^{-1}-u^{-1}\\ &= (u+h)^{-1}\circ(u-(u+h))\circ u^{-1}\\ &= -(u+h)^{-1}\circ h\circ u^{-1}.\end{aligned}$$

Pour prouver le théorème, il suffit de montrer que, $u \in \mathrm{Isom}(E ; F)$ étant fixé, la différence entre $(u+h)^{-1}\circ h\circ u^{-1}$ et la fonction (linéaire en h) $u^{-1}\circ h\circ u^{-1}$ est $o(\|h\|)$. Or

$$(u+h)^{-1}\circ h\circ u^{-1}-u^{-1}\circ h\circ u^{-1}=((u+h)^{-1}-u^{-1})\circ h\circ u^{-1},$$

d'où

$$\|(u+h)^{-1}\circ h\circ u^{-1}-u^{-1}\circ h\circ u^{-1}\| \leqslant \|(u+h)^{-1}-u^{-1}\|\cdot\|u^{-1}\|\cdot\|h\|.$$

Il suffit donc de montrer que $\|(u+h)^{-1}-u^{-1}\|$ tend vers 0 quand h tend vers 0 ; or il en est bien ainsi, puisque l'application $u \mapsto u^{-1}$ est *continue* (théorème 1.7.3).

Ainsi on a prouvé que φ est différentiable en tout point $u \in \mathrm{Isom}(E ; F)$, et que sa dérivée $\varphi'(u)$ est donnée par la formule (2.4.4). Pour montrer que φ est de classe C^1, il reste à démontrer que l'application

$$\varphi' : \mathrm{Isom}(E ; F) \to \mathscr{L}(\mathscr{L}(E ; F) ; \mathscr{L}(F ; E))$$

est *continue*. Introduisons pour cela une notation : pour $v \in \mathscr{L}(F ; E)$, $w \in \mathscr{L}(F ; E)$, notons $\psi(v, w)$ l'application linéaire

$$h \mapsto -v\circ h\circ w \text{ de } \mathscr{L}(E ; F) \text{ dans } \mathscr{L}(F ; E).$$

Alors (2.4.4) montre que

$$\varphi'(u)=\psi(u^{-1}, u^{-1}).$$

Or l'application $(v, w) \mapsto \psi(v, w)$ de $\mathscr{L}(F ; E) \times \mathscr{L}(F ; E)$ dans $\mathscr{L}(\mathscr{L}(E ; F) ; \mathscr{L}(F ; E))$ est bilinéaire ; elle est continue, car

$$\|\psi(v, w)\cdot h\| = \|v\circ h\circ w\| \leqslant \|v\|\cdot\|h\|\cdot\|w\|,$$

ce qui entraîne (cf. relation (1.4.2) et les lignes qui suivent) :

$$\|\psi(v, w)\| \leqslant \|v\|\cdot\|w\|.$$

Donc ψ est une application bilinéaire continue. Cela dit, l'application

$$u \mapsto \varphi'(u)=\psi(u^{-1}, u^{-1})$$

est composée de l'application *continue* $u \mapsto (u^{-1}, u^{-1})$ de $\mathrm{Isom}(E ; F)$ dans

$$\mathscr{L}(F ; E) \times \mathscr{L}(F ; E),$$

et de l'application *continue* $(v, w) \mapsto \psi(v, w)$. C'est donc une application continue.

C.Q.F.D.

N.B.—On verra plus tard que cette application est même différentiable.

CAS PARTICULIER DU THÉORÈME 2.4.4. Supposons $E = F = \mathbf{R}$ (resp. $E = F = \mathbf{C}$ dans le cas complexe). Une application linéaire $E \to F$ est alors définie par un *scalaire*, que nous noterons u ; et pour que l'application définie par u soit un isomorphisme, il faut et il suffit que $u \neq 0$. Ainsi Isom($\mathbf{R}$; $\mathbf{R}$) s'identifie à l'ouvert de $\mathbf{R}$ formé des éléments $u \neq 0$. Le théorème 2.4.4 dit, dans ce cas, que l'application $u \to 1/u$ est différentiable, et que sa dérivée est $-1/u^2$. Résultat bien classique!

Fonctions à valeurs dans un produit d'espaces de Banach

Supposons que l'espace F soit le produit d'un nombre fini k d'espaces de Banach :

$$F = F_1 \times \ldots \times F_k.$$

Introduisons les notations suivantes : pour chaque entier i tel que $1 \leqslant i \leqslant k$, soit

$$p_i : F \to F_i$$

l'application de projection du produit sur son i-ième facteur, et soit

$$u_i : F_i \to F$$

l'injection définie par

$$u_i(x_i) = (0, \ldots, x_i, \ldots, 0)$$

[0 partout sauf à la i-ième place]. On vérifie que p_i et u_i sont des applications linéaires continues, et qu'elles satisfont aux relations

$$\text{(2.5.1)} \qquad \begin{cases} p_i \circ u_i = 1_{F_i} \text{ (application identique de } F_i), \\ \sum_{i=1}^{k} u_i \circ p_i = 1_F \text{ (application identique de } F). \end{cases}$$

PROPOSITION 2.5.1. Avec les notations précédentes, soit $f : U \to F$ une application continue, U désignant toujours un ouvert d'un espace de Banach E. *Pour que* f *soit différentiable au point* $a \in U$, *il faut et il suffit que, pour chaque* i $(1 \leqslant i \leqslant k)$, la fonction $f_i = p_i \circ f : U \to F_i$ *soit différentiable au point* a, *et alors*

$$\text{(2.5.2)} \qquad f'(a) = \sum_{i=1}^{k} u_i \circ f_i'(a).$$

Démonstration facile. Les applications linéaires p_i et u_i sont différentiables. Donc si f est différentiable, l'application composée $p_i \circ f$ est différentiable (théorème 2.2.1) et a pour dérivée

$$f_i'(a) = p_i \circ f'(a) \in \mathscr{L}(E ; F_i).$$

Réciproquement, supposons que f_i' soit différentiable au point a, quel que soit l'entier

$i\,(1 \leqslant i \leqslant k)$; la deuxième relation (2.5.1) donne

$$\sum_{i=1}^{k} u_i \circ p_i \circ f = f,$$

c'est-à-dire $f = \sum_{i=1}^{k} u_i \circ f_i$; donc (théorème 2.2.1 et proposition 2.3.1) f est différentiable au point a, et

2.5

$$f'(a) = \sum_{i=1}^{k} u_i \circ f_i'(a).$$

C.Q.F.D.

Remarque. Pour que l'application $f' : U \to \mathscr{L}(E\,;F)$ soit continue, il faut et il suffit que $f_i' : U \to \mathscr{L}(E\,;F_i)$ soit continue pour chaque i.

Exemple. La proposition précédente s'applique notamment lorsque $F = \mathbf{R}^k$ (resp. $\mathbf{C}^k$) ; on prend alors

$$F_1 = \ldots = F_k = \mathbf{R} \text{ (resp. } = \mathbf{C}).$$

La donnée de $f : U \to \mathbf{R}^k$ équivaut à la donnée de k fonctions numériques $f_i : U \to \mathbf{R}$ (à savoir $f_i = p_i \circ f$) ; pour que f soit différentiable, il faut et il suffit que chaque f_i soit différentiable, et alors $f'(a)$ est l'application linéaire $E \to \mathbf{R}^k$ dont les k composantes sont $f_1'(a), \ldots, f_k'(a)$.

Application. Considérons, comme au nº 2.4, une application bilinéaire continue $f : E_1 \times E_2 \to F$; soient d'autre part $u : U \to E_1$ et $v : U \to E_2$ deux applications continues. La donnée de f permet de « multiplier » entre elles les fonctions u et v ; d'une façon précise, elles définissent $w : U \to F$ par la formule

$$w(x) = f(u(x), v(x)). \tag{2.5.3}$$

PROPOSITION 2.5.2. *Avec ces notations, supposons que* u *et* v *soient différentiables au point* $a \in U$; *alors* w *est différentiable en ce point, et* $w'(a)$ *est déterminée par la formule* :

$$w'(a) \cdot h = f(u'(a) \cdot h, v(a)) + f(u(a), v'(a) \cdot h), \quad \text{pour } h \in E. \tag{2.5.4}$$

DÉMONSTRATION. D'après la proposition 2.5.1, l'application $x \to (u(x), v(x))$ de U dans $E_1 \times E_2$ est différentiable au point a, et sa dérivée est l'application linéaire

$$h \to (u'(a) \cdot h, v'(a) \cdot h).$$

D'autre part, l'application $f : E_1 \times E_2 \to F$ est différentiable en tout point de $E_1 \times E_2$, puisqu'elle est bilinéaire continue (théorème 2.4.3). L'application w définie par (2.5.3) est la composée

$$U \xrightarrow{(u,\,v)} E_1 \times E_2 \xrightarrow{f} F\,;$$

donc (théorème 2.2.1) elle est différentiable au point a, et sa dérivée est égale à la composée des applications dérivées.

Explicitons cette dérivée : on doit, dans la relation (2.4.2), remplacer a_1 par $u(a)$, a_2 par $v(a)$, h_1 par $u'(a)\cdot h$, h_2 par $v'(a)\cdot h$; on obtient alors précisément le second membre de la relation (2.5.4) à démontrer.

C.Q.F.D.

Cas particulier. Supposons que $E = \mathbf{R}$, c'est-à-dire que u et v soient des fonctions d'une variable *numérique* x. Alors on sait que $u'(a)\cdot h$ est simplement $h\cdot u'(a)$ (produit de $u'(a)\in E_1$ par le scalaire h), que $v'(a)\cdot h$ est $h\cdot v'(a)$, et $w'(a)\cdot h$ est $h\cdot w'(a)$. La relation (2.5.4) donne alors, en y faisant $h = 1$:

$$w'(a) = f(u'(a), v(a)) + f(u(a), v'(a)). \tag{2.5.5}$$

On reconnaît ici la formule qui donne la dérivée d'un « produit » de deux fonctions u et v d'une variable numérique : par exemple, produit vectoriel de deux fonctions à valeurs dans $\mathbf{R}^3$, produit scalaire de deux fonctions à valeurs dans $\mathbf{R}^n$. Cette formule s'applique si $E_1 = E_2 = A$ est une *algèbre de Banach* (cf. 1.7), $f : A \times A \to A$ étant la multiplication dans cette algèbre ; dans ce cas (2.5.5) s'écrit

$$(uv)'(a) = u'(a)v(a) + v(a)u'(a).$$

Le cas le plus simple est celui où l'algèbre A est le corps des scalaires, et on retrouve alors la formule usuelle donnant la dérivée du produit de deux fonctions numériques.

Cas où U est un ouvert d'un produit d'espaces de Banach 2.6

On suppose maintenant que $E = E_1 \times \ldots \times E_n$, et que U est un ouvert de E. Soit $f : U \to F$ une application continue. Pour chaque $a = (a_1, \ldots, a_n) \in U$, considérons l'injection $\lambda_i : E_i \to E$ définie par

$$\lambda_i(x_i) = (a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n).$$

L'application composée $f \circ \lambda_i$ est définie dans l'ouvert $(\lambda_i)^{-1}(U) \subset E_i$, qui contient $a_i \in E_i$; on l'appelle la i-ième application partielle au point a.

PROPOSITION 2.6.1 ET DÉFINITION. Avec les notations précédentes, si f est différentiable au point a, alors, pour chaque entier i ($1 \leqslant i \leqslant n$), l'application partielle $f \circ \lambda_i$ est différentiable au point a_i. On note $f'_{x_i}(a)$, ou $\partial f/\partial x_i(a)$, ou $f'_{x_i}(a_1, \ldots, a_n)$, ou $\partial f/\partial x_i(a_1, \ldots, a_n)$ la dérivée de cette application partielle au point a_i ; c'est un élément de $\mathscr{L}(E_i ; F)$, qu'on appelle aussi la *dérivée partielle* de f par rapport à x_i. En outre, on a

$$f'(a)\cdot(h_1, \ldots, h_n) = \sum_{i=1}^{n} f'_{x_i}(a)\cdot h_i, \quad \text{pour } h_1 \in E_1, \ldots, h_n \in E_n. \tag{2.6.1}$$

DÉMONSTRATION. Notons $u_i : E_i \to E$ l'injection canonique, définie par

$$u_i(x_i) = (0, \ldots, 0, x_i, 0, \ldots, 0).$$

u_i est linéaire continue. On a évidemment

(2.6.2) $$\lambda_i(x_i) = a + u_i(x_i - a_i), \quad \lambda_i(a_i) = a\ ;$$

d'où

(2.6.3) $$\lambda_i'(x_i) = u_i \quad \text{pour tout } x_i \in E_i.$$

Si f est différentiable au point a, $f \circ \lambda_i$ est donc différentiable au point a_i (théorème 2.2.1), et $(f \circ \lambda_i)' = f'(a) \circ u_i$. Ainsi $f'_{x_i}(a)$ existe, et n'est autre que $f'(a) \circ u_i$.

La relation (2.6.1) résulte alors de la relation

$$\sum_{i=1}^{n} u_i \circ p_i = 1_E \quad \text{(cf. (2.5.1))},$$

qui donne

(2.6.4) $$\sum_{i=1}^{n} (f'(a) \circ u_i) \circ p_i = f'(a),$$

qui n'est qu'une autre manière d'écrire (2.6.1).

Remarque. Contrairement à ce qui avait lieu pour la proposition 2.5.1, la proposition 2.6.1 n'affirme pas que si les dérivées partielles partielles $f'_{x_i}(a)$ existent, la dérivée $f'(a)$ existe. On reviendra sur cette question au § 3.

Supposons maintenant que f soit différentiable en tout point de U, et soit

$$f' : U \to \mathscr{L}(E\ ; F)$$

l'application dérivée. Alors l'application « dérivée partielle »

$$f'_{x_i} : U \to \mathscr{L}(E_i\ ; F)$$

est composée de f′ et de l'application linéaire

(2.6.5) $$\mathscr{L}(E\ ; F) \to \mathscr{L}(E_i\ ; F)$$

qui, à toute application linéaire continue $\varphi : E \to F$, associe $\varphi \circ u_i : E_i \to F$; en effet, cela résulte de la relation

(2.6.6) $$f'_{x_i}(a) = f'(a) \circ u_i.$$

L'application linéaire (2.6.5) est de norme $\leqslant 1$, donc continue. Par suite, si l'application dérivée f′ est continue, les applications f'_{x_i} sont continues. La réciproque est vraie, car la relation (2.6.4) montre que l'application f′ est égale à la somme des applications composées

$$U \xrightarrow{f'_{x_i}} \mathscr{L}(E_i\ ; F) \to \mathscr{L}(E\ ; F),$$

où $\mathscr{L}(E_i\ ; F) \to \mathscr{L}(E\ ; F)$ est l'application linéaire qui, à $\varphi_i \in \mathscr{L}(E_i\ ; F)$, associe $\varphi_i \circ p_i \in \mathscr{L}(E\ ; F)$.

En résumé :

PROPOSITION 2.6.2. *Si* f *est différentiable en tout point de* U, *une condition nécessaire et suffisante pour que* f *soit de classe* C^1 *est que les* $f'_{x_i} : U \to \mathscr{L}(E_i\,;F)$ *soient des applications continues.*

Combinaison des cas étudiés en 2.5 et 2.6 2

Supposons à la fois que $E = E_1 \times \ldots \times E_n$ et $F = F_1 \times \ldots \times F_m$. Soit U un ouvert de E, et soit $f : U \to F$ une application différentiable au point $a = (a_1, \ldots, a_n) \in U$. Alors les $p_i \circ f = f_i$ (où $p_i : F \to F_i$ est la projection) sont différentiables au point a, donc ont des dérivées partielles $\partial f_i / \partial x_j(a)$ $(1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n)$. On a

$$\frac{\partial f_i}{\partial x_j}(a) \in \mathscr{L}(E_j\,;F_i),$$

et

(2.7.1) $$f'(a) = \sum_{i,j} u_i \circ \frac{\partial f_i}{\partial x_j}(a) \circ q_j,$$

où

$$\begin{cases} q_j : E \to E_j \text{ est la projection canonique,} \\ u_i : F_i \to F \text{ est l'injection canonique.} \end{cases}$$

Ainsi l'application linéaire $f'(a)$ est déterminée par la *matrice* des $(\partial f_i/\partial x_j)(a)$; c'est une matrice à m lignes et n colonnes (i est l'indice des lignes, la i-ième ligne correspondant à l'espace F_i ; j est l'indice des colonnes, la j-ième colonne correspondant à l'espace E_j). On vérifiera (exercice !) que si on a en outre un espace de Banach $G = G_1 \times \ldots \times G_p$, une application continue g d'un ouvert $V \subset G$ dans $U \subset E$, différentiable en un point $b \in V$ tel que $g(b) = a$, et si on note h l'application composée $f \circ g : V \to F$, la matrice $[(\partial h_i/\partial y_k)(b)]$ (où y_k varie dans G_k, $1 \leqslant k \leqslant p$) est égale au produit de la matrice $(\partial g_j/\partial y_k)(b)$ par la matrice $(\partial f_i/\partial x_j)(a)$; autrement dit :

(2.7.2) $$\frac{\partial h_i}{\partial y_k}(b) = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(a) \circ \frac{\partial g_j}{\partial y_k}(b).$$

Ceci résulte tout simplement du théorème de dérivation d'une fonction composée :

$$h'(b) = f'(a) \circ g'(b).$$

Ce qui précède s'applique notamment lorsque $E = \mathbf{R}^n$, $F = \mathbf{R}^m$ (avec $E_1 = \ldots = E_n = \mathbf{R}$, $F_1 = \ldots = F_m = \mathbf{R}$). Alors $(\partial f_i/\partial x_j)(a) \in \mathbf{R}$. Si en outre $G = \mathbf{R}^p$ (avec $G_1 = \ldots = G_p = \mathbf{R}$), les $(\partial g_j/\partial y_k)(b)$ sont aussi des scalaires, et dans le second membre de (2.7.2) le signe $\circ$ signifie simplement la multiplication des scalaires, comme dans le produit classique de deux matrices dont les éléments sont des scalaires.

Remarque finale : comparaison entre R-différentiabilité et C-différentiabilité 2.8

Comme on l'a déjà dit, la théorie précédente s'applique aussi bien aux espaces de Banach *réels* qu'aux espaces de Banach *complexes*. On va maintenant comparer les deux théories.

Soient E et F deux espaces de Banach sur le corps $\mathbf{C}$; on peut les considérer aussi comme des espaces de Banach sur le corps $\mathbf{R}$: il suffit de ne considérer le produit d'un vecteur et d'un scalaire que lorsque ce scalaire est *réel*. Par exemple, $\mathbf{C}$ est un espace vectoriel sur $\mathbf{C}$, de dimension 1 ; pour sa structure réelle sous-jacente c'est un espace vectoriel sur $\mathbf{R}$, de dimension 2.

E et F étant donc des espaces de Banach sur $\mathbf{C}$, soit U un ouvert de E, et soit $f : U \to F$ une application continue. Soit enfin $a \in U$. On peut envisager deux propriétés de f :

(i) f est différentiable au point a, pour les structures d'e.v. sur $\mathbf{C}$;

(ii) f est différentiable au point a, pour les structures d'e.v. sur $\mathbf{R}$.

Dans le premier cas, la dérivée $f'(a)$ est une application linéaire continue $E \to F$, où « linéaire » signifie : $\mathbf{C}$-linéaire. Pour bien préciser, on notera $\mathscr{L}_{\mathbf{C}}(E\,;F)$ l'espace vectoriel (normé complet) des applications $\mathbf{C}$-linéaires continues de E dans F. Dans le second cas, la dérivée $f'(a)$ est une application $\mathbf{R}$-linéaire continue $E \to F$. Pour bien préciser, on notera $\mathscr{L}_{\mathbf{R}}(E\,;F)$ l'espace de Banach des applications $\mathbf{R}$-linéaires continues de E dans F.

Une application $\mathbf{C}$-linéaire est, a fortiori, $\mathbf{R}$-linéaire ; donc $\mathscr{L}_{\mathbf{C}}(E\,;F) \subset \mathscr{L}_{\mathbf{R}}(E\,;F)$; l'espace de Banach $\mathscr{L}_{\mathbf{C}}(E\,;F)$ est un sous-espace de $\mathscr{L}_{\mathbf{R}}(E\,;F)$, et c'est d'ailleurs un sous-espace *fermé*, puisqu'il est *complet*.

La condition (i) ci-dessus exprime qu'il existe une $g \in \mathscr{L}_{\mathbf{C}}(E\,;F)$, nécessairement unique, telle que $\|f(x) - f(a) - g(x - a)\| = o(\|x - a\|)$. La condition (ii) exprime qu'il existe une $g \in \mathscr{L}_{\mathbf{R}}(E\,;F)$, nécessairement unique, telle que

$$\|f(x) - f(a) - g(x - a)\| = o(\|x - a\|).$$

Il est donc clair que la condition (i) entraîne la condition (ii) : *si* f *est* $\mathbf{C}$*-différentiable au point* a, f *est a fortiori* $\mathbf{R}$*-différentiable au point* a, *et sa dérivée* $f'(a)$ *au sens réel est la même que sa dérivée au sens complexe.*

Inversement, supposons que f soit $\mathbf{R}$-différentiable au point a, et soit $f'(a) \in \mathscr{L}_{\mathbf{R}}(E\,;F)$ sa dérivée. *Pour que* f *soit* $\mathbf{C}$*-différentiable au point* a, *il faut et il suffit que* $f'(a)$ *appartienne au sous-espace vectoriel* $\mathscr{L}_{\mathbf{C}}(E\,;F)$ *de* $\mathscr{L}_{\mathbf{R}}(E\,;F)$.

La théorie qui s'occupe spécialement des fonctions $\mathbf{C}$-différentiables fait l'objet d'une autre partie du cours de Mathématiques II : ces fonctions sont aussi appelées fonctions *holomorphes*, et on étudiera surtout les fonctions holomorphes dans un ouvert de $\mathbf{C}$.

3. THÉORÈME DES ACCROISSEMENTS FINIS; APPLICATIONS

La terminologie des « accroissements finis » s'explique par des raisons historiques : la notion d'accroissements « finis » s'oppose à celle d'accroissements « infinitésimaux » dont il était question autrefois en calcul différentiel.

Enoncé du théorème principal 3.1

THÉORÈME 3.1.1. *Soient* a *et* b *deux points de* $\mathbf{R}$ *tels que* $a < b$. *Notons* $[a, b]$ *le segment fermé qu'ils déterminent. Soient donnés deux applications continues*

$$f : [a, b] \to F, \quad g : [a, b] \to \mathbf{R},$$

où F *est un espace de Banach. Supposons que* f *et* g *soient différentiables en tout point de l'intervalle ouvert*]a, b[*et que*

$$\|f'(x)\| \leqslant g'(x) \quad \text{pour } a < x < b. \tag{3.1.1}$$

Alors on a

$$\|f(b) - f(a)\| \leqslant g(b) - g(a). \tag{3.1.2}$$

On va démontrer un théorème un peu plus fort, et dont la démonstration n'est pas plus difficile. Avant de l'énoncer, il faut donner une définition.

DÉFINITION. On dit qu'une application $f : [a, b] \to F$ admet une *dérivée à droite* en un point $x \in [a, b[$ si

$$\lim_{\substack{h \to 0 \\ h > 0}} \frac{1}{h}(f(x + h) - f(x))$$

existe ; cette limite se note alors $f'_d(x)$ et s'appelle la dérivée à droite de f au point x. C'est un élément de F. On définit de même, si elle existe, *la dérivée à gauche* de f en un point $x \in]a, b]$:

$$f'_g(x) = \lim_{\substack{h \to 0 \\ h < 0}} \frac{1}{h}(f(x + h) - f(x)).$$

Pour que f admette une dérivée $f'(x)$ en un point $x \in]a, b[$, il faut et il suffit que $f'_d(x)$ et $f'_g(x)$ existent, et soient égales : c'est évident.

THÉORÈME 3.1.2. Même énoncé que pour le théorème 3.1.1, sauf que l'on suppose seulement l'existence de $f'_d(x)$ et de $g'_d(x)$ en tout point $x \in]a, b[$, et que l'inégalité (3.1.1) est remplacée par

$$\|f'_d(x)\| \leqslant g'_d(x) \quad \text{pour } a < x < b. \tag*{(3.1.1)$_d$}$$

La conclusion est la même : c'est l'inégalité (3.1.2).

Ainsi on a *affaibli les hypothèses* du théorème 3.1.1, et la conclusion subsiste. Le théorème 3.1.2 est donc plus fort que le théorème 3.1.1.

DÉMONSTRATION DU THÉORÈME 3.1.2. Donnons-nous un nombre $\varepsilon > 0$. On va montrer

$$\|f(x) - f(a)\| \leqslant g(x) - g(a) + \varepsilon(x - a) + \varepsilon \tag{3.1.3}$$

pour tout $x \in [a, b]$. Un fois cela prouvé, on appliquera cette inégalité pour $x = b$, puis on fera tendre ε vers 0, ce qui, à la limite, donnera l'inégalité (3.1.2) de l'énoncé.

Introduisons l'ensemble U des $x \in [a, b]$ pour lesquels (3.1.3) est *fausse*, c'est-à-dire pour lesquels

$$\|f(x) - f(a)\| > g(x) - g(a) + \varepsilon(x - a) + \varepsilon. \tag{3.1.4}$$

On veut montrer que U est *vide*. On sait déjà que U est *ouvert* : en effet, puisque les fonctions f et g ont été supposées *continues*, chacun des deux membres de l'inégalité (3.1.4) est une fonction continue de x. Or si on considère une inégalité $\varphi(x) > 0$, où φ est une fonction continue à valeurs numériques, l'ensemble des points x qui vérifient cette inégalité est *ouvert*. Ainsi U est ouvert. Raisonnons par l'absurde, en supposant U non vide. Alors U aurait une *borne inférieure* c. On peut dire trois choses :

(i) $c > a$; en effet, la relation (3.1.3) est vraie pour tout x assez voisin de a, à cause de la continuité des deux membres ;

(ii) $c \notin U$, parce que U est ouvert : si c appartenait à U, il y aurait des x tels que $a < x < c$ et $x \in U$, et c ne serait pas la borne inférieure de U ;

(iii) $c < b$, sinon U devrait se réduire au point b, et ne serait donc pas ouvert.

Puisque $a < c < b$, on peut appliquer à c l'hypothèse de l'énoncé :

$$\|f'_d(c)\| \leqslant g'_d(c). \tag{3.1.5}$$

D'après la définition de $f'_d(c)$ et de $g'_d(c)$, il existe un intervalle $c \leqslant x \leqslant c + \eta$ (où $\eta > 0$) dans lequel on a

$$\|f'_d(c)\| \geqslant \left\|\frac{f(x) - f(c)}{x - c}\right\| - \frac{\varepsilon}{2}$$

$$g'_d(c) \leqslant \frac{g(x) - g(c)}{x - c} + \frac{\varepsilon}{2}.$$

Ces inégalités et (3.1.5) entraînent

$$\|f(x) - f(c)\| \leqslant g(x) - g(c) + \varepsilon(x - c). \tag{3.1.6}$$

Or on a vu que $c \notin U$; autrement dit, on a

$$\|f(c) - f(a)\| \leqslant g(c) - g(a) + \varepsilon(c - a) + \varepsilon. \tag{3.1.7}$$

Les inégalités (3.1.6) et (3.1.7) donnent

$$\begin{aligned}\|f(x) - f(a)\| &\leqslant \|f(x) - f(c)\| + \|f(c) - f(a)\| \\ &\leqslant g(x) - g(a) + \varepsilon(x - a) + \varepsilon.\end{aligned}$$

Ceci est valable pour $c \leqslant x \leqslant c + \eta$. Ainsi (3.1.3) est vraie pour $c \leqslant x \leqslant c + \eta$. Mais alors tout $x \leqslant c + \eta$ satisfait à (3.1.3), et la borne inférieure de U est donc $\geqslant c + \eta$. On arrive à une contradiction.

C.Q.F.D.

Remarque. On a un théorème analogue au théorème 3.1.2, en remplacant les dérivées à droite par les dérivées à gauche. Il s'en déduit en changeant x en $-x$.

Complément. Signalons un théorème encore plus fort que le théorème 3.1.2.

THÉORÈME 3.1.3. Soient $f : [a, b] \to F$ et $g : [a, b] \to \mathbf{R}$ deux applications continues. Supposons que, pour tout $x \in [a, b]$ *sauf peut-être pour ceux d'un ensemble dénombrable* D, $f'_d(x)$ et $g'_d(x)$ existent et satisfassent à $(3.1.1)_d$. Alors on a

$$\|f(b) - f(a)\| \leqslant g(b) - g(a). \tag{3.1.2}$$

Indication sur la démonstration. On range les points de D en une suite $x_1, x_2, \ldots, x_n, \ldots$; pour chaque $x \in [a, b]$, on note N_x l'ensemble des entiers $n > 0$ tels que $x_n < x$. Alors, au lieu de prouver (3.1.3) comme dans le théorème 3.1.2, on démontre

$$\|f(x) - f(a)\| \leqslant g(x) - g(a) + \varepsilon\left(\sum_{n \in N_x} 2^{-n}\right) + \varepsilon(x - a) + \varepsilon .$$

Une fois ceci prouvé, on fait $x = b$, puis on fait tendre ε vers 0.

Cas particuliers du théorème principal 3.2

Dans le théorème 3.1.2, supposons que l'espace F soit réduit à $\{0\}$. Alors l'hypothèse se réduit à $g'_d(x) \geqslant 0$, et la conclusion devient $g(b) \geqslant g(a)$. Comme on peut appliquer le résultat à deux points quelconques x_1 et x_2 de $[a, b]$, on obtient : $g(x_2) \geqslant g(x_1)$ chaque fois que $x_1 < x_2$. Ainsi :

COROLLAIRE 3.2.1. *Si* $g : [a, b] \to \mathbf{R}$ *est continue et possède une dérivée à droite* $g'_d(x) \geqslant 0$ *pour tout* $x \in]a, b[$, g *est croissante (au sens large) dans l'intervalle* $[a, b]$.

La réciproque est évidente : si une fonction croissante possède une dérivée à droite, celle-ci est $\geqslant 0$.

Appliquons maintenant le théorème 3.1.2 en supposant que $g(x) = kx$ (k constante $\geqslant 0$). L'hypothèse $(3.1.1)_d$ devient donc $\|f'_d(x)\| \leqslant k$. D'où le :

COROLLAIRE 3.2.2. *Soit* $f : [a, b] \to F$ *une application continue* (où F est un espace de Banach). *Supposons que* f *admette une dérivée à droite* $f'_d(x)$ *pour tout* $x \in]a, b[$, *et que*

$$\|f'_d(x)\| \leqslant k \quad (k \text{ constante } \geqslant 0).$$

Alors on a

$$\|f(b) - f(a)\| \leqslant k(b - a),$$

et, plus généralement :

$$\|f(x_2) - f(x_1)\| \leqslant k|x_2 - x_1| \quad \text{quels que soient } x_1, x_2 \in [a, b]. \tag{3.2.1}$$

Théorème des accroissements finis lorsque la variable est dans un espace de Banach 3.3

Jusqu'à présent f était une fonction d'une variable réelle. Soit maintenant U un ouvert d'un espace de Banach E, et soit $f : U \to F$ une application continue, F étant un espace de Banach. Rappelons que si a et b sont deux points de E, on appelle *segment* d'extrémités a et b l'ensemble des points $x \in E$ de la forme

$$x = (1 - t)a + tb, \quad \text{avec} \quad 0 \leqslant t \leqslant 1.$$

PROPOSITION 3.3.1. *Si* f *est différentiable dans* U, *et si le segment d'extrémités* a *et* b *est contenu dans* U, *on a*

$$\|f(b) - f(a)\| \leqslant \|b - a\| \cdot \sup_{0 \leqslant t \leqslant 1} \|f'((1 - t)a + tb)\|. \tag{3.3.1}$$

DÉMONSTRATION. Soit $h(t) = f((1 - t)a + tb)$. C'est une fonction différentiable de t, et on a (cf. théorème sur la dérivée d'une fonction composée, n° 2.2) :

$$h'(t) = f'((1 - t)a + tb) \cdot (b - a),$$

d'où

$$\|h'(t)\| \leqslant \|f'((1 - t)a + tb)\| \cdot \|b - a\|.$$

Appliquons le corollaire 3.2.2 (où f serait remplacée par h) : on obtient (3.3.1).

C.Q.F.D.

Supposons maintenant que l'ouvert U soit *convexe*, c'est-à-dire que, pour tout couple (a, b) de points de U, le segment d'extrémités a et b soit contenu dans U. Alors la proposition 3.3.1 entraîne aussitôt :

THÉORÈME 3.3.2. *Soit* U *un ouvert convexe d'un espace de Banach* E, *et soit* $f : U \to F$ *une application différentiable à valeurs dans un espace de Banach* F. *Supposons que :*

$$\|f'(x)\| \leqslant k \quad \text{pour tout} \quad x \in U.$$

Alors, quels que soient $x_1 \in U$, $x_2 \in U$, *on a*

$$\|f(x_2) - f(x_1)\| \leqslant k\|x_2 - x_1\|. \tag{3.3.2}$$

Une fonction f qui satisfait à (3.3.2) est, par définition, lipschitzienne de constante k, où k-*lipschitzienne* (la définition peut se donner chaque fois qu'on a deux espaces métriques et une application de l'un dans l'autre).

COROLLAIRE 3.3.3. *Sous les hypothèses précédentes, supposons que* $k = 0$, *c'est-à-dire que* $f'(x) = 0$ *pour tout* $x \in U$. *Alors* f *est constante dans* U.

On va voir maintenant que ce corollaire est, en réalité, valable non seulement si U est convexe, mais plus généralement si U est *connexe*.

Rappelons qu'un espace topologique X est dit *connexe* si, chaque fois que X est une réunion de deux *ouverts disjoints*, l'un deux est vide (et l'autre est X).

THÉORÈME 3.3.4. Soit U un ouvert *connexe* d'un espace de Banach E, et soit $f : U \to F$ une application différentiable à valeurs dans un espace de Banach. Si la dérivée $f'(x)$ est nulle pour tout $x \in U$, alors f est constante.

DÉMONSTRATION. Soit a un point quelconque de U ; U contient une boule ouverte B de centre a ; cette boule est convexe, donc f est constante dans B, d'après le corollaire 3.3.3. Ainsi f est *localement constante dans* U [par définition, on dit qu'une fonction définie sur un espace topologique est localement constante si chaque point possède un voisinage dans lequel la fonction est constante]. On n'a pas encore utilisé l'hypothèse de l'énoncé, suivant laquelle U est connexe. On va maintenant la faire intervenir ; pour cela, démontrons un lemme :

Lemme. Soit $f : X \to Y$ une application *continue* d'un espace topologique X non vide dans un espace topologique *séparé* Y. Si f est *localement constante*, et si X est *connexe*, alors f est *constante* dans X.

Ce lemme entraîne évidemment le théorème 3.3.4. Reste à prouver le lemme : soit $b \in Y$; l'image réciproque $f^{-1}(b)$ est un *fermé* de X, puisque f est continue, et que l'ensemble $\{b\} \subset Y$ est fermé (la topologie de Y étant séparée). D'autre part, $f^{-1}(b)$ est *ouvert*, puisque f est localement constante. Ainsi $f^{-1}(b)$ est à la fois ouvert et fermé ; donc X est réunion de l'ouvert $f^{-1}(b)$ et de son complémentaire qui est aussi ouvert. Puisque X est supposé connexe, l'un de ces deux ensembles est X. Cela dit, prenons un point $a \in X$, et choisissons $b = f(a)$; alors $f^{-1}(b)$ n'est pas vide, donc $f^{-1}(b) = X$, ce qui prouve que $f(x) = b$ pour tout $x \in X$.

C.Q.F.D.

Le théorème 3.3.4 donne un résultat meilleur que le corollaire 3.3.3 ; en effet, tout ouvert U convexe est connexe. Cela résulte de la proposition suivante qui donne un critère permettant de reconnaître si U est connexe :

PROPOSITION 3.3.5. *Soit* U *un ouvert d'un espace vectoriel normé* (*sur le corps* $\mathbf{R}$). *Les conditions suivantes sont équivalentes* :

(a) U *est connexe* ;

(b) *deux points quelconques de* U *peuvent être joints par un chemin dans* U ;

(c) *deux points quelconques de* U *peuvent être joints par une ligne brisée dans* U.

Avant la démonstration, il faut définir avec précision les termes utilisés dans l'énoncé des conditions (b) et (c).

DÉFINITION. Dans un espace topologique X, on appelle *chemin* une application *continue* φ du segment $[0, 1] \subset \mathbf{R}$ dans l'espace X ; le point $\varphi(0)$ s'appelle l'*origine*, le point $\varphi(1)$ s'appelle l'*extrémité* du chemin. On dit que deux points a et $b \in X$ peuvent être joints par un chemin s'il existe un chemin φ tel que $\varphi(0) = a$ et $\varphi(1) = b$.

DÉFINITION. Dans une partie A d'un espace vectoriel normé E sur le corps $\mathbf{R}$, on appelle *ligne brisée* un chemin $\varphi : [0, 1] \to A$ tel qu'il existe un nombre fini de points du segment $[0, 1]$:

$$t_0 = 0 < t_1 < \ldots < t_{n-1} < t_n = 1$$

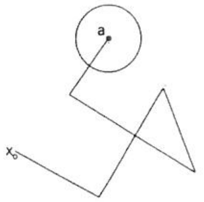
de façon que, dans chaque intervalle $[t_i, t_{i+1}]$ $(0 \leqslant i \leqslant n-1)$, la fonction φ soit somme d'une application *linéaire* et d'une constante (l'image de $[t_i, t_{i+1}]$ par φ est donc un segment de droite dans l'espace vectoriel E).

DÉMONSTRATION de la proposition 3.3.5. Il est clair que (c) $\Rightarrow$ (b). On va démontrer successivement (b) $\Rightarrow$ (a), puis (a) $\Rightarrow$ (c) ; cela prouvera l'équivalence de (a), (b), (c).

Démonstration de (*b*) $\Rightarrow$ (a) [cette démonstration vaut pour n'importe quel espace topologique, et pas seulement pour un ouvert d'un e.v. normé] : supposons (b) vrai, et raisonnons par l'absurde, en supposant qu'il existe deux ouverts non vides U_0 et U_1 contenus dans U, disjoints, et tels que U soit réunion de U_0 et U_1. Prenons un point $x_0 \in U_0$ et un point $x_1 \in U_1$; puisqu'on suppose (b), il existe un chemin $\varphi : [0, 1] \to U$ tel que $\varphi(0) = x_0$, $\varphi(1) = x_1$. Les ensembles $\varphi^{-1}(U_0)$ et $\varphi^{-1}(U_1)$ sont deux ouverts non vides, disjoints, du segment $[0, 1]$, qui en est la réunion. Ainsi le segment $[0, 1]$ ne serait pas connexe. C'est absurde, car on démontre en Topologie générale que tout segment de la droite numérique $\mathbf{R}$ est connexe.

Démonstration de (a) $\Rightarrow$ (c) : on peut supposer U non vide (sinon (a), (b) et (c) sont trivialement vrais). Choisissons donc $x_0 \in U$, et soit V l'ensemble des points de U que l'on peut joindre à x_0 par une ligne brisée dans U. On va montrer que V est à la fois ouvert et fermé dans U. Il en résulte que si U est connexe (hypothèse (a)), alors $V = U$, puisque V n'est pas vide ; donc que (a) $\Rightarrow$ (c).

V est *ouvert* dans U : soit $a \in V$ l'extrémité d'une ligne brisée l (contenue dans U) ayant pour origine x_0 (figure). Il existe une boule $B(a, r)$ de centre a, de rayon $r > 0$,



contenue dans U. Tout point $x \in B(a, r)$ peut être joint à a par un segment. En « mettant bout à bout » la ligne brisée l et ce segment, on obtient une ligne brisée d'origine x_0, d'extrémité $x \in B(a, r)$, et contenue dans U [Il faut naturellement retoucher le paramétrage initial de l, en s'arrangeant par exemple pour que l soit décrite lorsque t croît de 0 à $\frac{1}{2}$, et que le segment $[a, x]$ soit décrit quand t croît de $\frac{1}{2}$ à 1.] Ainsi a possède bien un voisinage dont tous les points x peuvent être joints à x_0 par une ligne brisée. Donc V est *ouvert*.

V est *fermé* dans U : soit $a \in U$ un point adhérent à V, et montrons que $a \in V$. Il existe une boule $B(a, r)$ contenue dans U ; puisque a est adhérent à V, il existe un point $b \in B(a, r) \cap V$. Ce point b peut être joint à x_0 par une ligne brisée contenue dans U, puisque $b \in V$; comme a peut être joint à b par un segment contenu dans U, il existe une ligne brisée d'origine x_0 et d'extrémité a, ce qui veut dire que $a \in V$.

C.Q.F.D.

Ainsi la proposition 3.3.5 est démontrée.

Remarque. Dans un espace topologique X, la *composante connexe* d'un point $x_0 \in X$ est définie comme le plus grand sous-ensemble connexe qui contient x_0 (on montre que parmi les parties connexes contenant x_0, il y en a une qui contient toutes les autres). Les composantes connexes de X forment une *partition* de X. Ici, pour un ouvert U d'un e.v. normé, la composante connexe de $x_0 \in U$ est l'ensemble V des points que l'on peut joindre à x_0 par une ligne brisée contenue dans U, ainsi qu'il résulte de la proposition 3.3.5. Dans la dernière partie de la démonstration de cette proposition, nous avons en fait prouvé que V est ouvert. Ainsi, si U est un ouvert d'un e.v. normé, les *composantes connexes de* U *sont des ensembles ouverts.*

Encore le théorème des accroissements finis 3.4

Soit E un e.v. normé. On sait ce qui est la *longueur* d'un segment d'origine a et d'extrémité b : c'est

$$d(a, b) = \|b - a\|.$$

Par définition, la *longueur d'une ligne brisée* est la somme des longueurs des segments qui la composent ; cette longueur est donc au moins égale à la distance $\|b - a\|$ de l'origine a à l'extrémité b.

DÉFINITION. Soit U un ouvert *connexe* d'un espace de Banach E. Pour a et $b \in U$, on note $d_U(a, b)$ la borne inférieure des longueurs des lignes brisées, *contenues dans* U, qui ont a pour origine et b pour extrémité. Cette définition est justifiée, car il existe de telles lignes brisées, en vertu de la proposition 3.3.5. On a

$$\begin{cases} d_U(a, b) = d_U(b, a) \\ d_U(a, c) \leqslant d_U(a, b) + d_U(b, c) \end{cases}$$

(le vérifier à titre d'exercice). Autrement dit, $d_U(a, b)$ est une distance dans l'espace topologique U.

Exercice. Montrer que cette distance définit sur U la même topologie que la distance $\|a - b\|$. Pour cela, on observera que si a est donné, on a $d_U(a, b) = \|a - b\|$ dès que b est assez voisin de a.

PROPOSITION 3.4.1. *Soit* U *un ouvert connexe d'un espace de Banach* E. *Soit* $f : U \to F$ *une application différentiable à valeurs dans un espace de Banach* F. *Supposons que*

$$\|f'(x)\| \leqslant k \quad \text{pour tout} \quad x \in U.$$

Alors, quels que soient x_1 et $x_2 \in U$, *on a* $\|f(x_2) - f(x_1)\| \leqslant k \cdot d_U(x_1, x_2)$. (Comparer cet énoncé celui du théorème 3.3.2.) La démonstration de la proposition 3.4.1 est laissée au lecteur à titre d'exercice.

Une liste d'exercices 3.5

(1) (Facile). Soit U un ouvert connexe d'un espace de Banach E ; soit $f : U \to F$ une application différentiable à valeurs dans un espace de Banach F. Montrer que si l'application $f' : U \to \mathscr{L}(E ; F)$ est constante, f est somme d'une constante et de la restriction à U d'une application linéaire (continue).

(2) Soit f une application continue d'un intervalle $[a, b]$ dans un espace de Banach F. Posons $g(x) = \|f(x)\|$. Montrer que si f est dérivable à droite en un point $x \in [a, b[$, alors

g est dérivable à droite en ce point et

$$|g'_d(x)| \leqslant \|f'_d(x)\|.$$

(Utiliser la convexité de la norme et l'exercice n° 6 à la fin de ce chapitre.) Montrer par un exemple simple que la dérivabilité de f n'entraîne pas nécessairement la dérivabilité de g.

(3) Soit f une application continue d'un segment $[a, b] \subset \mathbf{R}$ dans un espace de Banach F ayant une dérivée à droite en tout point $x \in]a, b[$. Soit C un sous-ensemble fermé convexe de F tel que $f'_d(x) \in C$ pour tout $x \in]a, b[$. Montrer que

$$\frac{f(b) - f(a)}{b - a} \in C.$$

[Calquer la marche de la démonstration sur celle du théorème 3.1.2. Montrer que pour $a < u < v < b$ et $\varepsilon > 0$ quelconque, l'ensemble

$$U_\varepsilon = \{x \in [u, v] ; \quad \frac{f(x) - f(u)}{x - u} \notin C_\varepsilon\}$$

est vide; C_ε désigne l'ensemble des éléments $y \in F$ tels que $d(y, C) \leqslant \varepsilon$ (on vérifie que C_ε est fermé convexe).]

Première application du théorème des accroissements finis : convergence d'une suite de fonctions différentiables 3.6

THÉORÈME 3.6.1. Soit U un ouvert *convexe* d'un espace de Banach E, et soit une suite d'applications différentiables

$$f_n : U \to F \quad (F : \text{espace de Banach}).$$

Faisons les hypothèses suivantes :

(i) il existe un point $a \in U$ tel que la suite des $f_n(a) \in F$ ait une limite ;

(ii) la suite des applications $f'_n : U \to \mathscr{L}(E ; F)$ *converge uniformément* dans U vers une $g : U \to \mathscr{L}(E ; F)$.

Alors, pour chaque $x \in U$, la suite des $f_n(x) \in F$ a une limite (qu'on note f(x)) ; la convergence de la suite $\{f_n\}$ vers f est *uniforme sur chaque partie bornée* de U ; enfin, *la fonction limite* f *est différentiable, et sa dérivée* $f'(x)$ *n'est autre que* $g(x)$.

DÉMONSTRATION. D'après le théorème 3.3.2 (qui est applicable puisqu'on a supposé U convexe), on a

$$(3.6.1) \qquad \|f_p(x) - f_p(a) - (f_q(x) - f_q(a))\| \leqslant \|x - a\| \cdot \sup_{y \in U} \|f'_p(y) - f'_q(y)\|.$$

A cause de l'hypothèse (ii), le second membre tend vers 0 quand p et q augmentent indéfiniment ; de plus la convergence est *uniforme* vis-à-vis de x, pourvu que $\|x - a\|$ reste borné, c'est-à-dire que x reste dans une partie bornée de U. Donc le premier membre de (3.6.1) tend vers 0 quand $p \to \infty$, $q \to \infty$, et ceci uniformément quand x reste dans une partie bornée de U. De plus, d'après (i), $f_p(a) - f_q(a)$ tend vers 0. Donc $\|f_p(x) - f_q(x)\|$ tend vers 0, uniformément en x dans toute partie bornée de U. Soit f la fonction limite ; chaque point de U a un voisinage borné, dans lequel f est limite uniforme de la suite des fonctions continues f_n, donc f est continue au voisinage de chaque point de U ; cela signifie simplement que f est *continue* dans U. Il reste à montrer que f est différentiable, et que $f'(x) = g(x)$. Fixons $x_0 \in U$. Il suffit de montrer que

(3.6.2) $$\|f(x) - f(x_0) - g(x_0)\cdot(x - x_0)\| = o(\|x - x_0\|).$$

Or on a évidemment

(3.6.3) $$\begin{aligned}\|f(x) - f(x_0) - g(x_0)\cdot(x - x_0)\| &\leqslant \|f(x) - f(x_0) - (f_n(x) - f_n(x_0))\| \\ &+ \|f_n(x) - f_n(x_0) - f'_n(x_0)\cdot(x - x_0)\| \\ &+ \|f'_n(x_0)\cdot(x - x_0) - g(x_0)\cdot(x - x_0)\|.\end{aligned}$$

Donnons-nous $\varepsilon > 0$. On peut majorer le premier des trois termes du second membre de (3.6.3), car (3.6.1) donne

$$\|f_p(x) - f_p(x_0) - (f_n(x) - f_n(x_0))\| \leqslant \varepsilon\|x - x_0\|$$

dès que p et n sont $\geqslant n_0$ (entier convenable dépendant de ε) ; donc, en passant à la limite quand $p \to \infty$:

(3.6.4) $$\|f(x) - f(x_0) - f_n(x) - f_n(x_0)\| \leqslant \varepsilon\|x - x_0\| \quad \text{pour } n \geqslant n_0.$$

D'autre part on a, à la limite,

$$\|f'_n(x_0) - g(x_0)\| \leqslant \varepsilon \quad \text{pour } n \geqslant n_0,$$

d'où, pour $n \geqslant n_0$:

(3.6.5) $$\|f'_n(x_0)\cdot(x - x_0) - g(x_0)\cdot(x - x_0)\| \leqslant \varepsilon\|x - x_0\|.$$

Ainsi, pour $n \geqslant n_0$, le premier et le troisième terme du second membre de (3.6.3) sont majorés chacun par $\varepsilon\|x - x_0\|$. *Fixons alors* n (par exemple $n = n_0$) ; dès que $\|x - x_0\| \leqslant h$ assez petit, on a

$$\|f_n(x) - f_n(x_0) - f'_n(x_0)\cdot(x - x_0)\| \leqslant \varepsilon\|x - x_0\|,$$

d'après la définition même de la dérivée $f'_n(x_0)$; ceci est une majoration du second terme du second membre de (3.6.3). Au total, (3.6.3) donne donc

$$\|f(x) - f(x_0) - g(x_0)\cdot(x - x_0)\| \leqslant 3\varepsilon\|x - x_0\| \quad \text{dès que} \quad \|x - x_0\| \leqslant h.$$

Pour tout $\varepsilon > 0$ existe un tel $h > 0$. Or ceci exprime précisément (3.6.2).

Remarque : si $E = \mathbf{R}$, le théorème 3.6.1 s'étend aux *dérivées à droite*.

On peut s'affranchir de l'hypothèse de *convexité* faite sur U dans le théorème 3.6.1 :

THÉORÈME 3.6.2. Soit U un ouvert connexe d'un espace de Banach E, et soit une suite d'applications différentiables

$$f_n : U \to F \quad (F : \text{espace de Banach}).$$

Faisons les hypothèses suivantes :

(i) il existe un point $a \in U$ tel que la suite des $f_n(a) \in F$ ait une limite ;

(ii) pour tout $x_0 \in U$, il existe une boule de centre x_0 dans laquelle la suite $\{f'_n\}$ converge uniformément.

Alors, pour chaque $x \in U$, la suite des $f_n(x) \in F$ a une limite (qu'on note $f(x)$) ; *tout point de* U *possède un voisinage dans lequel la convergence de la suite* $\{f_n\}$ *vers* f *est uniforme ; enfin,* f *est différentiable dans* U, *et* $f'(x) = g(x)$ *pour tout* $x \in U$.

Ce théorème est une conséquence facile du théorème 3.6.1. Le détail de la démonstration est laissé au lecteur, qui procédera ainsi : 1) l'ensemble des $x \in U$ tels que la suite $\{f_n(x)\}$ ait une limite est *ouvert et fermé* dans U (utiliser le théorème 3.6.1) ; 2) si $x_0 \in U$, et si $B(x_0, r)$ est une boule dans laquelle la suite $\{f'_n\}$ converge uniformément, la suite $\{f_n\}$ converge uniformément vers f dans $B(x_0, r)$ (appliquer encore le théorème 3.6.1) ; 3) que $f'(x) = g(x)$ résulte encore du théorème 3.6.1 appliqué à une boule convenable contenue dans U.

Deuxième application du théorème des accroissements finis : relation entre différentiabilité partielle et différentiabilité 3.7

Soient $E_1, \ldots, E_n$, F des espaces de Banach, et soit $E = E_1 \times \ldots \times E_n$. Soit U un ouvert de E, et soit $f : U \to F$ une application continue. Pour la notion de dérivée partielle f'_{x_i}, ou $\partial f/\partial x_i$, se reporter au n° 2.6.

THÉORÈME 3.7.1. *Avec les notations précédentes, pour que* f *soit de classe* C^1, *il faut et il suffit que* f *ait des dérivées partielles* $\partial f/\partial x_i$ *et que les applications*

$$\frac{\partial f}{\partial x_i} : U \to \mathscr{L}(E_i\,;F)$$

soient continues.

La condition est nécessaire, d'après les propositions 2.6.1 et 2.6.2. Il reste à montrer qu'elle est suffisante. Supposons donc que, pour tout $a \in U$, les dérivées partielles $(\partial f/\partial x_i)(a) \in \mathscr{L}(E_i\,;F)$ existent, et que les applications $\partial f/\partial x_i : U \to \mathscr{L}(E_i\,;F)$ soient continues. On veut montrer que f est alors de classe C^1. On va d'abord montrer que, pour tout a, $f'(a)$ existe (c'est-à-dire que f est différentiable au point a) ; la démonstration sera alors terminée, car la proposition 2.6.2 montre que l'application $f' : U \to \mathscr{L}(E\,;F)$ sera continue.

En résumé, il nous reste seulement à démontrer la proposition suivante :

PROPOSITION 3.7.2. Si les dérivées partielles $(\partial f/\partial x_i)(x)$ existent en tout point $x = (x_1, \ldots, x_n) \in U$, et si les applications $\partial f/\partial x_i : U \to \mathscr{L}(E_i\,;F)$ sont *continues au point* a, *alors* f *est différentiable au point* a.

La démonstration va utiliser de manière essentielle le théorème des accroissements finis. On veut montrer que

$$\|f(x_1, \ldots, x_n) - f(a_1, \ldots, a_n) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) \cdot (x_i - a_i)\|$$

est $o(\|x - a\|)$, c'est-à-dire $o(\|x_1 - a_1\| + \ldots + \|x_n - a_n\|)$, d'après la définition de la norme sur un produit d'espaces de Banach. Or on a l'identité évidente :

$$f(x_1, \ldots, x_n) - f(a_1, \ldots, a_n) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) \cdot (x_i - a_i)$$

$$= f(x_1, x_2, \ldots, x_n) - f(a_1, x_2, \ldots, x_n) - \frac{\partial f}{\partial x_1}(a) \cdot (x_1 - a_1)$$

$$+ f(a_1, x_2, \ldots, x_n) - f(a_1, a_2, \ldots, x_n) - \frac{\partial f}{\partial x_2}(a) \cdot (x_2 - a_2)$$

$$+ \ldots + f(a_1, \ldots, a_{n-1}, x_n) - f(a_1, \ldots, a_{n-1}, a_n) - \frac{\partial f}{\partial x_n}(a) \cdot (x_n - a_n).$$

Il suffit donc de montrer que, $\varepsilon > 0$ étant donné, il existe $\eta > 0$ tel que les inégalités

$$\|x_1 - a_1\| \leqslant \eta, \ldots, \|x_n - a_n\| \leqslant \eta \tag{3.7.1}$$

entraînent

$$\left\{\begin{array}{l} \|f(x_1, x_2, \ldots, x_n) - f(a_1, x_2, \ldots, x_n) - \frac{\partial f}{\partial x_1}(a) \cdot (x_1 - a_1)\| \leqslant \varepsilon \|x_1 - a_1\| \\ \|f(a_1, x_2, \ldots, x_n) - f(a_1, a_2, \ldots, x_n) - \frac{\partial f}{\partial x_2}(a) \cdot (x_2 - a_2)\| \leqslant \varepsilon \|x_2 - a_2\| \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \|f(a_1, \ldots, a_{n-1}, x_n) - f(a_1, \ldots, a_{n-1}, a_n) - \frac{\partial f}{\partial x_n}(a) \cdot (x_n - a_n)\| \leqslant \varepsilon \|x_n - a_n\|. \end{array}\right. \tag{3.7.2}$$

Montrons par exemple que, ε étant donné, on peut choisir η de facon que (3.7.1) entraîne la première inégalité ; la démonstration serait analogue pour les autres. On aura ainsi à faire un choix de η pour chacune des n inégalités (3.7.2), et alors on pourra trouver un η (le plus petit) qui marche pour les n inégalités (3.7.2). Regardons donc le premier membre de la première des inégalités (3.7.2). Soit ξ_1 une variable (élément de l'espace E_1, suffisamment voisin de a_1) ; posons

$$g(\xi_1) = f(\xi_1, x_2, \ldots, x_n) - \frac{\partial f}{\partial x_1}(a) \cdot (\xi_1 - a_1).$$

On veut majorer $\|g(x_1) - g(a_1)\|$. Or g a une dérivée, par hypothèse : c'est

$$g'(\xi_1) = \frac{\partial f}{\partial x_1}(\xi_1, x_2, \ldots, x_n) - \frac{\partial f}{\partial x_1}(a_1, \ldots, a_n).$$

Puisque $(\partial f/\partial x_1)(x)$ est une fonction de x qui est continue au point a (par hypothèse !), il existe un $\eta > 0$ tel que les inégalités (3.7.1) entraînent

$$\left\| \frac{\partial f}{\partial x_1}(x_1, x_2, \ldots, x_n) - \frac{\partial f}{\partial x_1}(a_1, a_2, \ldots, a_n) \right\| \leqslant \varepsilon.$$

S'il en est ainsi, et si $\xi_1 = (1 - t)a_1 + tx_1$ est un point du segment d'origine a_1 et d'extrémité x_1 (dans l'espace vectoriel E_1), on a aussi

$$\left\| \frac{\partial f}{\partial x_1}(\xi_1, x_2, \ldots, x_n) - \frac{\partial f}{\partial x_1}(a_1, a_2, \ldots, a_n) \right\| \leqslant \varepsilon,$$

car $\|\xi_1 - a_1\| \leqslant \|x_1 - a_1\| \leqslant \eta$. D'après la proposition 3.3.1 on conclut donc que

$$\|g(x_1) - g(a_1)\| \leqslant \varepsilon \|x_1 - a_1\|.$$

Et c'est justement ce que nous voulions démontrer. La proposition 3.7.2 est donc établie.

Remarque. La proposition 3.7.2 et le théorème 3.7.1 s'appliquent notamment lorsque $E_1 = \mathbf{R}, \ldots, E_n = \mathbf{R}$, donc $E = \mathbf{R}^n$. Alors $\partial f / \partial x_i$ sont des applications $U \to F$.

Troisième application du théorème des accroissements finis : notion de fonction strictement différentiable 3.8

Dans tout ce qui suit, U désigne un ouvert d'un espace de Banach E, et F un espace de Banach ; on considère des applications de U dans F.

DÉFINITION. $f : U \to F$ est *strictement tangente à zéro* au point $a \in U$, si les conditions suivantes sont vérifiées :

(i) $f(a) = 0$;

(ii) pour tout $\varepsilon > 0$ il existe $r > 0$ tel que, dans la boule $\|x - a\| \leqslant r$, f soit ε-lipschitzienne.

S'il en est ainsi, on a en particulier pour $\|x - a\| \leqslant r$,

$$\|f(x)\| = \|f(x) - f(a)\| \leqslant \varepsilon \|x - a\|,$$

donc f est *tangente à zéro* au point a. (Cf. la définition donnée en 2.1.) Ainsi « f strictement tangente à zéro » entraîne « f tangente à zéro », ce qui est heureux pour la terminologie choisie.

DÉFINITION. On dit que f_1 et f_2 sont strictement tangentes au point $a \in U$ si $f_1 - f_2$ est strictement tangente à 0. Vérifier (exercice !) qu'on obtient ainsi une *relation d'équivalence* entre fonctions $U \to F$.

DÉFINITION. $f : U \to F$ est *strictement différentiable* au point $a \in U$ s'il existe une *application linéaire continue* $g : E \to F$ telle que les applications

$$x \mapsto f(x) - f(a) \quad \text{et} \quad x \mapsto g(x - a)$$

soient strictement tangentes au point a.

S'il en est ainsi, ces deux applications sont a fortiori tangentes ; donc f est différentiable au point a, et g est égale à la dérivée $f'(a)$. Ainsi :

Pour que f soit strictement différentiable au point a, il faut et il suffit que f soit différentiable au point a, et que pour tout $\varepsilon > 0$ il existe $r > 0$ tel que l'application

$$x \mapsto f(x) - f(a) - f'(a) \cdot (x - a) = g(x)$$

soit ε- lipschitzienne dans la boule $\|x - a\| \leqslant r$. Explicitons : cela signifie que

$$(3.8.1) \qquad \begin{cases} f(x) - f(y) = f'(a) \cdot (x - y) + \|x - y\| \cdot \psi(x, y), \\ \text{avec } \lim\limits_{\substack{x \to a \\ y \to a}} \|\psi(x, y)\| = 0. \end{cases}$$

THÉORÈME 3.8.1. *Si* $f : U \to F$ *est différentiable* dans U, *et si l'application* $f' : U \to \mathscr{L}(E\,;F)$ *est continue au point* a, *alors* f *est strictement différentiable au point* a.

Ce critère de stricte différentiabilité se démontre grâce au théorème des accroissements finis. En effet, posons

$$g(x) = f(x) - f(a) - f'(a) \cdot (x - a).$$

g est différentiable, et on a

$$g'(x) = f'(x) - f'(a),$$

donc $\lim\limits_{x \to a} \|g'(x)\| = 0$ par hypothèse. Etant donné $\varepsilon > 0$, il existe $r > 0$ tel que

$$\|g'(x)\| \leqslant \varepsilon \quad \text{pour} \quad \|x - a\| \leqslant r.$$

D'après le théorème des accroissements finis (sous la forme du théorème 3.3.2), on conclut que g est ε-lipschitzienne dans la boule $\|x - a\| \leqslant r$.

C.Q.F.D.

4. INVERSION LOCALE D'UNE APPLICATION DE CLASSE C^1. THÉORÈME DES FONCTIONS IMPLICITES

Difféomorphismes de classe C^1 4.1

DÉFINITION. Soient E et F deux espaces de Banach, V un ouvert de E, W un ouvert de F. On dit que $f : V \to W$ est un difféomorphisme de classe C^1 (ou un C^1-difféomorphisme) si f est bijective, de classe C^1 (considérée comme application de V dans F), et si en outre l'application réciproque $g = f^{-1} : W \to V$ est de classe C^1 (considérée comme application de W dans E).

Erreur à ne pas commettre : une application de classe C^1, $f : V \to W$, peut être un *homéomorphisme* sans être un difféomorphisme de classe C^1 ; autrement dit, l'homéomorphisme réciproque $f^{-1} = W \to V$ n'est pas nécessairement de classe C^1. Par exemple, la fonction d'une variable réelle x :

$$y = x^3 = f(x)$$

définit un homéomorphisme de $\mathbf{R}$ sur $\mathbf{R}$; elle est de classe C^1, mais l'application réciproque

$$x = y^{\frac{1}{3}} = g(y)$$

n'est pas différentiable à l'origine ; en effet, la dérivée $f'(x)$ est $3x^2$ qui s'annule pour $x = 0$; si $g'(0)$ existait, on aurait $g'(0)f'(0) = 1$ (dérivée d'une application composée), ce qui est absurde. D'une manière générale :

PROPOSITION 4.1.1. *Soit* $f : V \to W$ *un homéomorphisme de classe* C^1 (V désigne un ouvert de Banach E, W un ouvert de l'espace de Banach F). *Pour que* f *soit un difféomorphisme de classe* C^1, *il faut et il suffit que, pour tout* $x \in V$, $f'(x)$ *appartienne à* $\mathrm{Isom}(E\,;F)$.

On va d'abord démontrer un lemme :

Lemme. Soit $f : V \to W$ un *homéomorphisme* ; supposons f différentiable en un point $a \in V$. Pour que $g = f^{-1}$ soit différentiable au point $b = f(a) \in W$, il faut et il suffit que $f'(a) \in \mathrm{Isom}(E\,;F)$, et alors

$$g'(b) = (f'(a))^{-1}.$$

La condition est nécessaire, car si g est différentiable au point b, le théorème sur la dérivée d'une application composée donne

$$g'(b) \circ f'(a) = 1_E, \qquad f'(a) \circ g'(b) = 1_F,$$

ce qui prouve que $f'(a)$ est un isomorphisme de E sur F, et que $g'(b)$ est l'isomorphisme réciproque. Montrons que la condition est suffisante : on suppose $f'(a) \in \mathrm{Isom}(E\,;F)$, et on veut montrer que g est différentiable au point b. Puisque f est différentiable au point a, on a, en posant $y = f(x)$ pour x voisin de a :

$$y - b = f'(a)\cdot(x - a) + \|x - a\|\cdot\varphi(x - a), \tag{4.1.1}$$

avec

$$\lim_{x \to a} \varphi(x - a) = 0.$$

Effectuons sur les deux membres la transformation linéaire $(f'(a))^{-1}$:

$$x - a = (f'(a))^{-1}\cdot(y - b) - \|x - a\|\,(f'(a))^{-1}\cdot\varphi(x - a)\,; \tag{4.1.2}$$

tout revient à prouver que

$$\|x - a\|\,(f'(a))^{-1}\cdot\varphi(x - a) = o(\|y - b\|).$$

Posons, pour abréger :

$$(f'(a))^{-1}\cdot\varphi(x - a) = \psi(x - a)\,;$$

ceci tend vers 0 quand x tend vers a, puisque $(f'(a))^{-1}$ est une application linéaire *continue* de F dans E. La relation (4.1.2) entraîne

$$\|(f'(a))^{-1} \cdot (y - b)\| \geqslant \|x - a\| (1 - \|\psi(x - a)\|),$$

d'où

$$\|x - a\| \leqslant \|y - b\| \cdot \frac{\|(f'(a))^{-1}\|}{1 - \|\psi(x - a)\|}$$

(dès que $\|x - a\|$ est assez petit pour que $\|\psi(x - a)\| < 1$). D'où

$$\|x - a\| \cdot \|\psi(x - a)\| \leqslant \|y - b\| \cdot \| (f'(a))^{-1} \| \cdot \frac{\|\psi(x - a)\|}{1 - \|\psi(x - a)\|}$$

$$= o(\|y - b\|).$$

C.Q.F.D.

Ayant établi le lemme, passons à la *démonstration de la proposition 4.1.1*. La condition de l'énoncé est évidemment nécessaire ; réciproquement, si $f'(x) \in \mathrm{Isom}(E ; F)$ pour tout $x \in V$, il résulte du lemme que g est différentiable en tout point $y \in W$, et que

(4.1.3) $$g'(y) = (f'(g(y)))^{-1}.$$

Il reste à démontrer que g est de classe C^1, c'est-à-dire que l'application

$$g' : W \to \mathscr{L}(F ; E)$$

est *continue*. Or (4.1.3) montre que cette application est composée de trois applications :

1) l'application $y \mapsto g(y)$ de W dans V, qui est continue puisque f est homéomorphisme ;
2) l'application $x \mapsto f'(x)$ de V dans Isom(E ; F), qui est continue puisque f a été supposée de classe C^1 ;
3) l'application $u \mapsto u^{-1}$ de Isom(E ; F) dans $\mathscr{L}(F ; E)$, qui est continue (théorème 1.7.3).

Ceci achève la démonstration.

Le théorème d'inversion locale 4.2

Jusqu'à présent, on avait supposé que $f : V \to W$ était un homéomorphisme. On veut maintenant s'affranchir de cette hypothèse. Le théorème fondamental est le suivant :

THÉOREME 4.2.1. *Soit* U *un ouvert d'un espace de Banach* E, *et soit* $f : U \to F$ *une application de classe* C^1 (F *étant un espace de Banach*). *Supposons que, en un point* $a \in U$, *on ait*

$$f'(a) \in \mathrm{Isom}(E ; F).$$

Alors il existe un voisinage ouvert V *de* a ($V \subset U$) *et un voisinage ouvert* W *de* $b = f(a)$, *tels que* f *soit un* C^1*-difféomorphisme de* V *sur* W.

La démonstration de ce théorème prendra un certain temps (voir ci-dessous, 4.3, 4.4 et 4.5). Avant de le prouver, on va tout de suite en déduire un :

COROLLAIRE 4.2.2. *Pour que* $f : U \to F$, *de classe* C^1, *soit un* C^1-*difféomorphisme de* U *sur un ouvert de* F, *il faut et il suffit que :*
(i) f *soit une injection* ;
(ii) $f'(x) \in \mathrm{Isom}(E\,;F)$ *pour tout* $x \in U$.

DÉMONSTRATION DU COROLLAIRE. Les deux conditions sont évidemment nécessaires. Réciproquement, supposons-les remplies ; la condition (ii) entraîne que $f : U \to F$ est une *application ouverte* (c'est-à-dire : pour tout ouvert $V \subset U$, l'image $f(V)$ est un ouvert de F) ; cela résulte en effet du théorème 4.2.1, qui montre que si $a \in U$, l'image par f de tout voisinage ouvert de a contient un voisinage ouvert de $f(a)$. En particulier, $f(U)$ est un *ouvert* de F. Si nous montrons que f est un homéomorphisme de U sur $f(U)$, alors nous saurons grâce à la proposition 4.1.1, que f est un C^1-difféomorphisme de U sur $f(U)$. Or, d'après (i), f est une *bijection* de U sur $f(U)$; cette bijection est une application qui est à la fois continue et ouverte ; puisque f est ouverte, $g = f^{-1} : f(U) \to U$ est continue, donc f est bien un homéomorphisme de U sur $f(U)$.

C.Q.F.D.

Démonstration du théorème d'inversion locale : première réduction 4.3

Plaçons-nous dans les hypothèses du théorème 4.2.1 (qu'il s'agit de démontrer). Puisque f est de classe C^1, f est *strictement différentiable* au point a (cf. théorème 3.8.1). Admettons pour un instant la proposition suivante :

PROPOSITION 4.3.1. *Soit* U *un ouvert d'un espace de Banach* E, *et soit* $f : U \to F$ *une application continue* (F étant un espace de Banach). *Supposons que* f *soit strictement différentiable au point* $a \in U$ *et que* $f'(a) \in \mathrm{Isom}(E\,;F)$. *Alors il existe un voisinage ouvert* V' *de* a ($V' \subset U$) *et un voisinage ouvert* W' *de* $b = f(a)$, *tels que* f *soit un homéomorphisme de* V' *sur* W'. *De plus l'homéomorphisme réciproque est strictement différentiable au point* $f(a)$.

Ceci étant admis, nous voyons que sous les hypothèses du théorème 4.2.1, $f'(x)$ existe pour tout $x \in V'$; de plus il existe un voisinage ouvert V de a ($V \subset V'$) tel que $f'(x) \in \mathrm{Isom}(E\,;F)$, car $\mathrm{Isom}(E\,;F)$ étant *ouvert* dans $\mathscr{L}(E\,;F)$ (cf. théorème 1.7.3), son image réciproque par l'application *continue* f' est un ouvert de V' qui contient a. Soit $W = f(V)$; W est ouvert dans W', puisque f est un homéomorphisme de V' sur W' (d'après la proposition 4.3.1, que nous admettons provisoirement) ; de plus f est un homéomorphisme de V sur W. On est alors dans les conditions d'application de la proposition 4.1.1, qui permet de conclure que f est un C^1-difféomorphisme de V sur W. Ainsi nous avons démontré le théorème 4.2.1, sous réserve que la proposition 4.3.1 soit établie.

Démonstration de la proposition 4.3.1 4.4

Plaçons-nous dans les hypothèses de la proposition 4.3.1. L'application linéaire continue $(f'(a))^{-1}$ applique F sur E ; considérons l'application composée

$$f_1 = (f'(a))^{-1} \circ f : U \to E$$

(rappelons que U est un ouvert de E). On vérifie facilement que f_1 est strictement différentiable au point $a \in U$, et que $f_1'(a) = 1_E$ [Exercice : faire cette vérification.] Puisque f_1 est strictement différentiable, on peut associer à chaque $k > 0$ un $r > 0$ tel que l'application $x \mapsto x - f_1(x) = \varphi(x)$ soit k-lipschitzienne dans la boule $\|x - a\| \leqslant r$. Choisissons une fois pour toutes un k tel que $0 < k < 1$, ce qui détermine un $r > 0$. Dans la boule $\|x - a\| \leqslant r$, l'application φ est donc *contractante*, et on peut alors appliquer la théorie des approximations successives traitée dans le cours de MATH. I. D'une façon précise, on va rappeler ci-dessous (et démontrer) le résultat dont on a besoin ici, et qui nous permet de conclure à l'existence d'un voisinage ouvert V de a (contenu dans la boule $\|x - a\| \leqslant r$) tel que f_1 soit un *homéomorphisme* de V sur un voisinage ouvert W_1 de $b_1 = f_1(a)$. Comme $f'(a)$ est un homéomorphisme de E sur F, on voit alors que

$$f = f'(a) \circ f_1$$

est un homéomorphisme de V sur W (transformé de W_1 par $f'(a)$), W ouvert de F contenant $b = f(a)$. Donc la proposition 4.3.1 est démontrée (à cela près que ce qui est noté V' et W' dans l'énoncé s'appelle V et W dans la démonstration).

Voici le résultat précis que nous avons admis, et qui nous a permis de prouver la proposition 4.3.1 :

THÉORÈME 4.4.1. Soit $B(a, r)$ la boule ouverte $\|x - a\| < r$ d'un espace de Banach E, et soit

$$f : B(a, r) \to E$$

une application continue, telle que l'application

$$x \mapsto x - f(x) = \varphi(x)$$

soit *contractante* (c'est-à-dire k-lipschitzienne, avec $k < 1$). Soit $f(a) = b$. Alors il existe un *ouvert* V contenant a, contenu dans $B(a, r)$, tel que f soit *homéomorphisme de* V *sur la boule ouverte* $B(b, (1 - k)r)$; de plus l'application réciproque

$$g = f^{-1} : B(b, (1 - k)r) \to B(a, r)$$

est $[1/(1 - k)]$-lipschitzienne.

Démonstration du théorème 4.4.1 4.5

Soient x et $x' \in B(a, r)$; on a

$$f(x) - f(x') = (x - x') - (\varphi(x) - \varphi(x'))$$

d'où

$$\|f(x) - f(x')\| \geqslant \|x - x'\| - \|\varphi(x) - \varphi(x')\|,$$

ce qui, compte tenu du fait que φ est k-lipschitzienne, donne

$$\|f(x) - f(x')\| \geqslant (1 - k) \cdot \|x - x'\|. \tag{4.5.1}$$

Lemme. Pour tout $y \in B(b, (1-k)r)$, *il existe un* x *et un seul dans* $B(a, r)$ *tel que* $f(x) = y$.

Démonstration de l'*unicité* : si on a $f(x) = f(x')$, l'inégalité (4.5.1) montre que $x = x'$.

Démonstration de l'*existence* : on va construire le x cherché par la méthode des approximations successives. On définit par récurrence sur n, une suite de points

$$(4.5.2) \qquad \begin{cases} x_0 = a, \quad x_1 = y + \varphi(x_0), \dots \\ x_{n+1} = y + \varphi(x_n), \dots . \end{cases}$$

Pour que cette définition récurrente soit licite, on doit s'assurer de proche en proche que $x_n \in B(a, r)$, ce qui permet alors de définir x_{n+1}, car φ est définie dans $B(a, r)$. D'une façon précise, on va montrer, par récurrence sur n, que

$$(4.5.3) \qquad \|x_n - a\| \leqslant \frac{1-k^n}{1-k}\|y - b\|,$$

et comme par hypothèse $\|y - b\| < (1-k)r$, on aura bien $\|x_n - a\| < r$. Pour $n = 1$, on a

$$x_1 - a = y + \varphi(a) - a = y - f(a) = y - b,$$

donc (4.5.3) est vérifié pour $n = 1$. Supposons (4.5.3) vrai pour n ($n \geqslant 1$), et montrons-le pour $n + 1$; on a, d'après (4.5.2),

$$x_{n+1} - x_n = \varphi(x_n) - \varphi(x_{n-1}),$$

d'où

$$\|x_{n+1} - x_n\| \leqslant k\|x_n - x_{n-1}\|$$

et par suite (par récurrence)

$$(4.5.4) \qquad \|x_{n+1} - x_n\| \leqslant k^n\|x_1 - a\| = k^n\|y - b\|.$$

Cette inégalité, jointe à (4.5.3), donne

$$\begin{aligned} \|x_{n+1} - a\| &\leqslant \|x_n - a\| + \|x_{n+1} - x_n\| \\ &\leqslant \left(\frac{1-k^n}{1-k} + k^n\right) \cdot \|y - b\| = \frac{1-k^{n+1}}{1-k}\|y - b\| \end{aligned}$$

ce qui prouve (4.5.3) où n serait remplacé par $n + 1$. Maintenant, (4.5.4) prouve que la série de terme général $x_{n+1} - x_n$ est normalement convergente, donc la suite (x_n) est une suite de Cauchy. Soit x sa limite ; par passage à la limite dans (4.5.3), on obtient

$$\|x - a\| \leqslant \frac{1}{1-k}\|y - b\| < r,$$

et par passage à la limite dans (4.5.2) on trouve

$$x = y + \varphi(x),$$

c'est-à-dire $y = f(x)$. *Ainsi le lemme est démontré.*

Introduisons maintenant une notation : pour $y \in B(b, (1-k)r)$, notons $g(y)$ l'unique $x \in B(a, r)$ tel que $f(x) = y$. On définit ainsi une application

$$g : B(b, (1-k)r) \to B(a, r).$$

L'inégalité (4.5.1) montre que si y et y' sont deux points de $B(b, (1-k)r)$, on a

$$\|g(y) - g(y')\| \leqslant \frac{1}{1-k}\|y - y'\|.$$

Donc la fonction g est $[1/(1-k)]$-lipschitzienne ; en particulier, g est *continue*. Soit $V \subset B(a, r)$ l'*image* de l'application g. On a

$$V = f^{-1}(B(b, (1-k)r)),$$

image réciproque d'un ouvert ; comme f est continue, V est *ouvert* dans $B(a, r)$, donc ouvert dans E. Il est clair que les applications $f : V \to B(b, (1-k)r)$ et

$$g : B(b, (1-k)r) \to V$$

sont bijectives, réciproques l'une de l'autre ; comme elles sont continues, ce sont des *homéomorphismes*.

Le théorème 4.4.1 est donc démontré. En même temps, la démonstration du *théorème d'inversion locale* (théorème 4.2.1) est achevée, puisque nous avions réduit sa démonstration à celle de la proposition 4.3.1, puis celle de la prop. 4.3.1 à la démonstration du théorème 4.4.1.

Théorème d'inversion locale dans le cas de la dimension finie 4.6

Dans le théorème 4.2.1, on suppose que $f'(a)$ est un isomorphisme linéaire $E \to F$. Donc les hypothèses du théorème impliquent que les espaces de Banach E et F sont isomorphes. Lorsque E et F sont de dimension finie, cela implique qu'ils ont la *même dimension*. Envisageons donc le cas où $E = \mathbf{R}^n$, $F = \mathbf{R}^n$. L'application $f : U \to F$ est alors définie par n fonctions numériques de n variables réelles dans l'ouvert U :

$$f_i(x_1, \ldots, x_n) \qquad (1 \leqslant i \leqslant n).$$

On les suppose de classe C^1. L'application linéaire $f'(a) \in \mathscr{L}(\mathbf{R}^n, \mathbf{R}^n)$ est définie par la matrice des dérivées partielles

$$\frac{\partial f_i}{\partial x_j}(a_1, \ldots, a_n)$$

(i est l'indice des lignes, j l'indice des colonnes). Dire que $f'(a) \in \mathrm{Isom}(\mathbf{R}^n, \mathbf{R}^n)$ équivaut à dire que le *déterminant* de cette matrice est $\neq 0$. On le note souvent

$$\frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)}(a_1, \ldots, a_n)$$

[il s'agit de la valeur, au point $a = (a_1, \ldots, a_n)$, de ce déterminant] et on l'appelle le *jacobien* de la transformation f au point a.

Le théorème d'inversion locale dit que si ce jacobien est $\neq 0$ au point a, il existe un ouvert V contenant a et contenu dans U, et un ouvert W contenant $b = f(a)$, tels que f soit un C^1-difféomorphisme de V sur W. L'application réciproque g est alors définie par n fonctions $g_i(y_1, \ldots, y_n)$ qui sont de classe C^1 dans W.

Théorème des fonctions implicites 4.7

On envisage la situation suivante : E, F, G désignent trois espaces de Banach, U un ouvert de $E \times F$, et $f : U \to G$ une application de classe C^1 ; f est donc une fonction de deux variables $f(x, y)$, où $x \in E$, $y \in F$, le couple (x, y) restant dans U.

Soit (a, b) un point de U, et supposons

$$f(a, b) = 0.$$

On se propose d'étudier les solutions (x, y) de l'équation

$$f(x, y) = 0$$

qui sont « suffisamment voisines » de (a, b). Pour cela, faisons l'hypothèse suivante :

(H) *la dérivée partielle* $f'_y(a, b) \in \mathscr{L}(F ; G)$ *est un isomorphisme de* F *sur* G.

THÉORÈME 4.7.1. (Théorème des fonctions implicites.) *Sous les hypothèses précédentes, il existe dans* $E \times F$ *un voisinage ouvert* V *de* (a, b) *contenu dans* U, *il existe dans* E *un voisinage ouvert* W *de* a, *et il existe une application de classe* C^1

$$g : W \to F$$

qui jouissent de la propriété suivante : la relation

$$(x, y) \in V \quad \text{et} \quad f(x, y) = 0 \tag{4.7.1}$$

est équivalente à la relation

$$x \in W \quad \text{et} \quad y = g(x). \tag{4.7.2}$$

Commentaire. Dans le voisinage V de (a, b), les solutions de l'équation $f(x, y) = 0$ sont données par (4.7.2) ; autrement dit, dans V l'équation $f(x, y) = 0$ se résout par $y = g(x)$, g de classe C^1 dans W.

Remarque. Puisque, par hypothèse, on a

$$(a, b) \in V \quad \text{et} \quad f(a, b) = 0,$$

et puisque $a \in W$, l'équivalence de (4.7.1) et (4.7.2) montre que l'on a $g(a) = b$.

DÉMONSTRATION DU THÉORÈME 4.7.1. On va se ramener au théorème d'inversion locale (théorème 4.2.1). Pour cela, considérons l'application

$$f_1 : U \to E \times G$$

définie par

$$f_1(x, y) = (x, f(x, y)), (x \in E, y \in F). \tag{4.7.3}$$

f_1 est de classe C^1 dans U, car ses deux composantes x et f(x, y) sont de classe C^1 dans U. Sa dérivée $f'_1(a, b)$ est définie par une matrice

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

où $\alpha \in \mathscr{L}(E\,;E)$, $\beta \in \mathscr{L}(F\,;E)$, $\gamma \in \mathscr{L}(E\,;G)$, $\delta \in \mathscr{L}(F\,;G)$. En fait, le calcul des dérivées partielles de f_1 montre que :

$$\begin{cases} \alpha = 1_E, & \beta = 0 \\ \gamma = f'_x(a, b), & \delta = f'_y(a, b). \end{cases}$$

Ainsi $f'_1(a)$ est l'application linéaire

$$(h, k) \mapsto (h, f'_x(a, b) \cdot h + f'_y(a, b) \cdot k) \tag{4.7.4}$$

de $E \times F$ dans $E \times G$. Puisque $f'_y(a, b) \in \mathrm{Isom}(E\,;F)$, il est clair que (4.7.4) est un *isomorphisme* $E \times F \to E \times G$, dont l'isomorphisme réciproque est

$$(h', k') \mapsto (h', (f'_y)^{-1} \cdot k' - (f'_y)^{-1} \circ f'_x \cdot h').$$

Nous pouvons donc appliquer à f_1, au voisinage du point $(a, b) \in U$, le théorème d'inversion locale. Il dit ceci :

Il existe dans $E \times F$ un voisinage ouvert V de (a, b), contenu dans U, et dans $E \times G$ un voisinage ouvert W_1 de $(a, 0) = f_1(a, b)$, tels que f_1 soit un C^1-difféomorphisme de V sur W_1.

Soit g_1 le difféomorphisme réciproque : il a la forme

$$g_1(x, z) = (x, g(x, z)), \quad \text{avec} \quad x \in E, z \in G$$

tels que $(x, z) \in W_1$. Ceci définit une fonction

$$g : W_1 \to F$$

de classe C^1. Puisque f_1 et g_1 sont deux homéomorphismes réciproques, on a équivalence entre les deux conditions suivantes :

(i) $(x, y) \in V$ et $f(x, y) = z$

(ii) $(x, z) \in W_1$ et $g(x, z) = y$.

Dans ces relations, faisons $z = 0$; la condition (i) devient (4.7.1) ; voyons ce que devient la condition (ii). Si on identifie E à un sous-espace vectoriel de $E \times F$, en identifiant $x \in E$ à $(x, 0) \in E \times F$, la relation $(x, 0) \in W_1$ exprime que x appartient à l'intersection de W_1 et de E ; cette intersection est un *ouvert* W de E, qui contient a [puisque W_1 contient le point (a, 0)]. Posons d'autre part

$$g(x, 0) = g(x)\,;$$

c'est une fonction de classe C^1 dans l'ouvert W. Alors la condition (ii), pour $z = 0$, s'écrit

$$x \in W \quad \text{et} \quad y = g(x).$$

C'est précisément (4.7.2), dont l'équivalence avec (4.7.1) est ainsi démontrée.

C.Q.F.D.

L'ouvert W qui intervient dans l'énoncé du théorème 4.7.1 n'est peut-être pas connexe. Mais il contient un ouvert connexe W' contenant a (par exemple une boule ouverte de centre a). Il est clair que *la relation*

$$x \in W' \quad \textit{et} \quad y = g(x)$$

entraîne

$$(x, y) \in U \quad \text{et} \quad f(x, y) = 0.$$

Je dis que la fonction g est *la seule fonction continue dans* W' *qui jouisse de ces propriétés.* D'une façon précise :

PROPOSITION 4.7.2. Soit W' un ouvert *connexe* de E, contenant a et contenu dans W, et soit $h : W' \to F$ une fonction *continue* jouissant des propriétés suivantes :

$$\begin{cases} h(a) = b, \qquad (x, h(x)) \in U \text{ pour tout } x \in W', \\ f(x, h(x)) = 0. \end{cases}$$

Alors h *est identique à* g *dans* W'.

Principe de la démonstration (laissée à titre d'exercice) : Soit A l'ensemble des $x \in W'$ tels que $h(x) = g(x)$; observer que $a \in A$, et que A est *fermé* dans W' ; montrer que A est *ouvert* dans W'. Conclure parce que W' est connexe.

Cas où E, F, G *sont de dimension finie.* Les hypothèses du théorème 4.7.1 impliquent alors que F et G ont même dimension. Supposons donc $E = \mathbf{R}^n$, $F = \mathbf{R}^p$, $G = \mathbf{R}^p$. On se donne un système d'équations

$$f_i(x_1, \ldots, x_n ; y_1, \ldots, y_p) = 0 \qquad (1 \leqslant i \leqslant p), \tag{4.7.5}$$

où les f_i sont des fonctions numériques de classe C^1 dans l'ouvert U ; on suppose que le jacobien

$$\frac{\partial(f_1, \ldots, f_p)}{\partial(y_1, \ldots, y_p)}$$

est $\neq 0$ au point $(a_1, \ldots, a_n ; b_1, \ldots, b_p)$. On conclut que le système (4.7.5) est équivalent au système

$$y_i = g_i(x_1, \ldots, x_n), \qquad 1 \leqslant i \leqslant p$$

(où les g_i sont de classe C^1), tout au moins quand $(x_1, \ldots, x_n)$ est assez voisin de $(a_1, \ldots, a_n)$ et $(y_1, \ldots, y_p)$ assez voisin de $(b_1, \ldots, b_p)$. Pour la formulation précise, reprendre le V et le W de l'énoncé du théorème 4.7.1.

5. DÉRIVÉES D'ORDRE SUPÉRIEUR

Dérivée seconde 5.1

Soient toujours E et F deux espaces de Banach, U un ouvert de E, et $f : U \to F$ une application que nous supposons différentiable dans U. On a alors une application dérivée

$$f' : U \to \mathscr{L}(E\,;F),$$

et on peut se demander si, à son tour, elle est différentiable.

DÉFINITION. On dit que f est *deux fois différentiable* au point $a \in U$ si l'application f' est différentiable au point a ; on note alors $f''(a)$ la dérivée (au point a) de f' ; on a

$$f''(a) \in \mathscr{L}(E\,;\mathscr{L}(E\,;F)).$$

Remarque. Sans supposer f différentiable dans tout U, on dit plus généralement que f est 2 fois différentiable au point $a \in U$ si : 1) f est différentiable dans un voisinage V de a ; 2) l'application $f' : V \to \mathscr{L}(E\,;F)$ est différentiable au point a.

DÉFINITION. On dit que f est deux fois différentiable dans U si elle est deux fois différentiable en tout point de U (autrement dit : f est différentiable dans U, et l'application $f' : U \to \mathscr{L}(E\,;F)$ est différentiable dans U). S'il en est ainsi, l'application $x \mapsto f''(x)$ est une application

$$f'' : U \to \mathscr{L}(E\,;\mathscr{L}(E\,;F)).$$

DÉFINITION. On dit que f est de classe C^2 (ou deux fois continûment différentiable) dans U si f est deux fois différentiable, et si l'application f'' est *continue*. Condition équivalente : f' est de classe C^1 dans U.

Rappelons (cf. 1.9) qu'on a défini une isométrie canonique

$$\mathscr{L}(E\,;\mathscr{L}(E\,;F)) \approx \mathscr{L}(E, E\,;F). \tag{5.1.1}$$

Par cette bijection, $f''(a)$ définit un élément de $\mathscr{L}(E, E\,;F)$, c'est-à-dire une *application bilinéaire continue* $E \times E \to F$. Par abus de langage, nous dirons souvent que $f''(a)$ est un élément de $\mathscr{L}(E, E\,;F)$. Si on explicite (5.1.1) en se reportant à 1.9, on trouve que l'application $E \times E \to F$ définie par $f''(a)$ est la suivante :

$$(h, k) \mapsto (f''(a) \cdot h) \cdot k. \tag{5.1.2}$$

Expliquons cela : h et k désignent deux vecteurs de E ; puisque $f''(a)$ est une application linéaire continue $E \to \mathscr{L}(E\,;F)$, la valeur de $f''(a)$ sur le vecteur $h \in E$ est un élément

$$f''(a)\cdot h \in \mathscr{L}(E\,;F).$$

Ainsi $f''(a)\cdot h$ est une application linéaire continue $E \to F$; sa valeur sur un vecteur $k \in E$ est notée

$$(f''(a)\cdot h)\cdot k.$$

On a ainsi précisé le sens de (5.1.2).

THÉORÈME 5.1.1. *Si* $f : U \to F$ *est deux fois différentiable au point* a, *la dérivée seconde* $f''(a) \in \mathscr{L}(E, F\,;F)$ *est une application bilinéaire symétrique ; autrement dit :*

(5.1.3) $$(f''(a)\cdot h)\cdot k = (f''(a)\cdot k)\cdot h, \qquad \forall h \in E \quad \text{et} \quad \forall k \in E.$$

DÉMONSTRATION. On va introduire la fonction

$$A(h, k) = f(a + h + k) - f(a + h) - f(a + k) + f(a),$$

qui est évidemment symétrique : $A(h, k) = A(k, h)$. Supposons qu'on ait démontré la relation :

(5.1.4) $$\|A(h, k) - (f''(a)\cdot k)\cdot h\| = o((\|h\| + \|k\|)^2).$$

Je dis que (5.1.3) en résultera facilement. En effet, si dans (5.1.4) on échange h et k, on obtient

$$\|A(h, k) - (f''(a)\cdot h)\cdot k\| = o((\|h\| + \|k\|)^2)\,;$$

cette relation et (5.1.4) entraînent

(5.1.5) $$\|(f''(a)\cdot k)\cdot h - (f''(a)\cdot h)\cdot k\| = o((\|h\| + \|k\|)^2),$$

puisque

$$\|(f''(a)\cdot k)\cdot h - (f''(a)\cdot h)\cdot k\| \leqslant \|(f''(a)\cdot k)\cdot h - A(h, k)\| + \|A(h, k) - (f''(a)\cdot h)\cdot k\|.$$

(5.1.5) dit que, $\varepsilon > 0$ étant donné, il existe $\eta > 0$ tel que l'on ait

(5.1.6) $$\|(f''(a)\cdot k)\cdot h - (f''(a)\cdot h)k\| \leqslant \varepsilon(\|h\| + \|k\|)^2$$

dès que $\|h\| + \|k\| \leqslant \eta$. Or, pour tout scalaire λ, on a

$$\|(f''(a)\cdot \lambda k)\cdot(\lambda h) - (f''(a)\cdot \lambda h)\cdot \lambda k\| = |\lambda|^2\cdot\|(f''(a)\cdot k)\cdot h - (f''(a)\cdot h)\cdot k\|.$$

Etant donnés h et k *quelconques* dans E, on peut toujours trouver un $\lambda \neq 0$ tel que $\|\lambda h\| + \|\lambda k\| \leqslant \eta$; donc d'après (5.1.6) [où h et k seraient remplacés par λh et λk], on a

$$|\lambda|^2\cdot\|(f''(a)\cdot k)\cdot h - (f''(a)\cdot h)\cdot k\,\| \leqslant \varepsilon|\lambda|^2(\|h\|^2 + \|k\|^2)\,;$$

en divisant par $|\lambda|^2 \neq 0$, on trouve donc que l'inégalité (5.1.6) est *vraie quels que soient* h *et* k ; comme $\varepsilon > 0$ a été choisi arbitrairement, on conclut que la relation (5.1.3) est vraie, ce qui démontre le théorème 5.1.1.

Ainsi, nous venons de voir que pour démontrer le théorème, il suffit de prouver la relation (5.1.4).

DÉMONSTRATION de (5.1.4). Partons de l'inégalité évidente :

$$\begin{aligned} &\|A(h, k) - (f''(a)\cdot k)\cdot h\| \\ &\quad\leqslant \|A(h, k) - f'(a + k)\cdot h + f'(a)\cdot h\| \\ &\qquad + \|f'(a + k)\cdot h - f'(a)\cdot h - (f''(a)\cdot k)\cdot h\|. \end{aligned} \tag{5.1.7}$$

On va majorer chacune des quantités du second membre, c'est-à-dire

$$\|A(h, k) - f'(a + k)\cdot h + f'(a)\cdot h\| \tag{5.1.8}$$

et

$$\|f'(a + k)\cdot h - f'(a)\cdot h - (f''(a)\cdot k)\cdot h\|. \tag{5.1.9}$$

Commençons par (5.1.9). On a

$$\|f'(a + k)\cdot h - f'(a)\cdot h - (f''(a)\cdot k)\cdot h\| \leqslant \|h\|\cdot\|f'(a + k) - f'(a) - f''(a)\cdot k\|.$$

D'après la définition de la dérivée de la fonction f' au point a, on a

$$\|f'(a + k) - f'(a) - f''(a)\cdot k\| = o(\|k\|).$$

Donc la quantité (5.1.9) est $\|h\|\cdot o(\|k\|)$, et est *a fortiori* $\|h\|\cdot o(\|h\| + \|k\|)$.

On va maintenant majorer (5.1.8). Considérons la fonction auxiliaire

$$B(h) = f(a + k + h) - f(a + h) - f'(a + k)\cdot h + f'(a)\cdot h.$$

Alors (5.1.8) n'est autre que $\|B(h) - B(0)\|$. D'après l'inégalité des accroissements finis (proposition 3.3.1), on a

$$\|B(h) - B(0)\| \leqslant \|h\|\cdot \sup_{0\leqslant t\leqslant 1} \|B'(th)\|.$$

On a

$$B'(h) = f'(a + k + h) - f'(a + h) - f'(a + k) + f'(a) ;$$

donc (5.1.8) est majoré par

$$\|h\|\cdot \sup_{0\leqslant t\leqslant 1} \|f'(a + k + th) - f'(a + th) - f'(a + k) + f'(a)\|. \tag{5.1.10}$$

On va maintenant majorer (5.1.10) ; d'après la définition de f''(a), on a

$$\begin{cases} f'(a + k + th) = f'(a) + f''(a)\cdot(k + th) + o(\|k + th\|) \\ f'(a + th) = f'(a) + f''(a)\cdot(th) + o(\|th\|) \\ f'(a + k) = f'(a) + f''(a)\cdot k + o(\|k\|). \end{cases}$$

On en déduit, par combinaison :

$$\|f'(a + k + th) - f'(a + th) - f'(a + k) + f'(a)\|$$
$$= o(\|k + th\|) + o(\|th\|) + o(\|k\|).$$

Puisque $\|k + th\| \leqslant \|k\| + \|h\|$ et $\|th\| \leqslant \|h\|$ quel que soit t ($0 \leqslant t \leqslant 1$), on voit que l'expression (5.1.10) est de la forme $o(\|h\| + \|k\|)$, et par conséquent (5.1.8) est majoré par

$$\|h\| \cdot o(\|h\| + \|k\|).$$

Finalement, chacune des quantités (5.1.8) et (5.1.9) est $\|h\| \cdot o(\|h\| + \|k\|)$; donc leur somme aussi. Alors (5.1.7) montre que

$$\|A(h, k) - (f''(a) \cdot k) \cdot h\| = \|h\| \cdot o(\|h\| + \|k\|).$$

Cela signifie que pour tout $\varepsilon > 0$ il existe $\eta > 0$ tel que l'on ait

$$\|A(h, k) - (f''(a) \cdot k) \cdot h\| \leqslant \varepsilon\|h\| \cdot (\|h\| + \|k\|)$$

dès que $\|h\| + \|k\| \leqslant \eta$. *A fortiori*, l'inégalité $\|h\| + \|k\| \leqslant \eta$ entraîne

$$\|A(h, k) - (f''(a) \cdot k) \cdot h\| \leqslant \varepsilon(\|h\| + \|k\|)^2,$$

ce qui démontre (5.1.4).

La démonstration du théorème 5.1.1 est ainsi achevée.

Cas où E est un produit $E_1 \times \ldots \times E_n$ 5.2

On suppose toujours que U est un ouvert de E, et que $f : U \to F$ est deux fois différentiable au point $a \in U$. Ceci implique (par définition) que f est différentiable en tout point x d'un voisinage de a. D'après (2.6.1), on a

$$f'(x) \cdot (h_1, \ldots, h_n) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) \cdot h_j, \quad \text{pour} \quad h_j \in E_j. \tag{5.2.1}$$

La même formule, appliquée à f' au lieu de f, donne

$$f''(a) \cdot (k_1, \ldots, k_n) = \sum_{i=1}^{n} \frac{\partial f'}{\partial x_i}(a) \cdot k_i, \quad \text{pour} \quad k_i \in E_i. \tag{5.2.2}$$

Par conséquent

$$(f''(a) \cdot (k_1, \ldots, k_n)) \cdot (h_1, \ldots, h_n) = \sum_{i=1}^{n} \left(\frac{\partial f'}{\partial x_i}(a) \cdot k_i \right) \cdot (h_1, \ldots, h_n). \tag{5.2.3}$$

Pour bien comprendre ce que signifie le second membre de cette relation, il faut penser que

$$\frac{\partial f'}{\partial x_i}(a) \in \mathscr{L}(E_i ; \mathscr{L}(E ; F)),$$

donc

$$\frac{\partial f'}{\partial x_i}(a) \cdot k_i \in \mathscr{L}(E\,;\,F),$$

et la valeur de ceci sur le vecteur $(h_1, \ldots, h_n) \in E$ est un élément de F.

Pour calculer $\partial f'/\partial x_i(a)$, on utilise la relation (5.2.1) qui dit ce qu'est f' ; en dérivant (5.2.1) par rapport à x_i on obtient

$$\left(\frac{\partial f'}{\partial x_i}(a) \cdot k_i\right) \cdot (h_1, \ldots, h_n) = \sum_{j=1}^{n} \left(\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(a) \cdot k_i\right) \cdot h_j. \tag{5.2.4}$$

Notons $(\partial^2 f/\partial x_i \partial x_j)(a)$ la valeur, au point a, de $\partial/\partial x_i(\partial f/\partial x_j)$; c'est un élément de $\mathscr{L}(E_i\,;\,\mathscr{L}(E_j\,;\,F)) \approx \mathscr{L}(E_i, E_j\,;\,F)$. Si dans (5.2.3) on remplace le second membre par sa valeur tirée de (5.2.4), on obtient enfin :

$$(f''(a) \cdot (k_1, \ldots, k_n)) \cdot (h_1, \ldots, h_n) = \sum_{i,j} \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \cdot k_i\right) \cdot h_j. \tag{5.2.5}$$

Telle est la *relation fondamentale* qui exprime $f''(a) \in \mathscr{L}(E, E\,;\,F)$ à l'aide des dérivées partielles

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \in \mathscr{L}(E_i, E_j\,;\,F).$$

Elle est, pour la dérivée seconde, l'analogue de ce qu'était (2.7.1) pour la dérivée première.

Exprimons maintenant que $f''(a) : E \times E \to F$ est une application bilinéaire *symétrique* (théorème 5.1.1). Par échange de k_i et h_i (pour chaque i), (5.2.5) donne

$$\sum_{i,j} \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \cdot k_i\right) \cdot h_j = \sum_{i,j} \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \cdot h_i\right) \cdot k_j,$$

ou encore, en échangeant les indices de sommation i et j dans le second membre :

$$\sum_{i,j} \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \cdot k_i\right) \cdot h_j = \sum_{i,j} \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a) \cdot h_j\right) \cdot k_i.$$

Ceci est une *identité* en $k_1, \ldots, k_n, h_1, \ldots, h_n$. On en déduit

$$\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \cdot k_i\right) \cdot h_j = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a) \cdot h_j\right) \cdot k_i \tag{5.2.6}$$

pour tout couple (i, j). Ceci exprime que l'application bilinéaire

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) : E_j \times E_i \to F$$

est composée de l'application de symétrie $E_j \times E_i \to E_i \times E_j$ [qui envoie (h_j, k_i) en (k_i, h_j)] et de l'application bilinéaire

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) : E_i \times E_j \to F.$$

On peut dire, brièvement, que *les deux applications bilinéaires* $(\partial^2 f/\partial x_i \partial x_j)(a)$ *et* $(\partial^2 f/\partial x_j \partial x_i)(a)$ *se déduisent l'une de l'autre par l'échange des variables* $k_i \in E_i$ *et* $h_j \in E_j$. En particulier, $(\partial^2 f/\partial x_i \partial x_j)(a)$, noté aussi $(\partial^2 f/\partial (x_i)^2)(a)$, est une application bilinéaire *symétrique* $E_i \times E_i \to F$.

Remarque. Dans tout ce qui précède, on a supposé a priori l'existence de $f''(a)$, — ce qui entraînait l'existence des dérivées partielles $(\partial^2 f/\partial x_i \partial x_j)(a)$. Or on a une condition *suffisante* pour que f soit deux fois dérivable au point a ; en effet, en appliquant *deux fois* la proposition 3.7.2, on trouve ceci :

PROPOSITION 5.2.1. Pour que $f''(a)$ existe, il suffit que les $\partial f/\partial x_j$ existent en tout point $x \in U$ et soient des fonctions *continues* dans U, et que les dérivées partielles $\partial/\partial x_i(\partial f/\partial x_j)$ existent en tout point $x \in U$ et soient des fonctions *continues au point* a [comme application $U \to \mathscr{L}(E_i, E_j ; F)$].

Cas particulier où $E = \mathbf{R}^n$. Dans ce cas, nous prenons $E_i = \mathbf{R}$ pour $i = 1, \ldots, n$. Alors $\mathscr{L}(E_i ; F)$ s'identifie à F, comme on l'a déjà dit souvent ; et

$$\mathscr{L}(E_i ; \mathscr{L}(E_j ; F)) = \mathscr{L}(\mathbf{R} ; \mathscr{L}(\mathbf{R} ; F))$$

s'identifie ainsi à F. Par cette identification, si on note pour un instant $c_{ij} \in F$ l'élément de F défini par $(\partial^2 f/\partial x_i \partial x_j)(a)$, l'application bilinéaire correspondante $\mathbf{R} \times \mathbf{R} \to F$ n'est autre que

$$(\lambda_i, \lambda_j) \mapsto \lambda_i \lambda_j c_{ij}.$$

D'après le résultat ci-dessous, on a

$$\lambda_i \lambda_j c_{ij} = \lambda_j \lambda_i c_{ji}, \qquad \text{quels que soient } \lambda_i \text{ et } \lambda_j.$$

On en déduit $c_{ij} = c_{ji}$ (faire par exemple $\lambda_i = 1$, $\lambda_j = 1$). Ainsi :

PROPOSITION 5.2.2. *Si* $f : U \to F$ *est une fonction deux fois différentiable de* n *variables réelles, on a*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a) \in F.$$

C'est le classique *théorème de Schwarz* ; mais il est souvent énoncé sous les hypothèses de la proposition 5.2.1, qui sont des conditions suffisantes, mais non nécessaires, pour que $f''(a)$ existe.

Le théorème de Schwarz ne fait qu'exprimer, dans le cas particulier d'une fonction de n variables réelles, le théorème 5.1.1 qui vaut en toute généralité.

Dérivées successives 5.3

Soit $f : U \to F$ une fonction deux fois différentiable. On a alors l'application « dérivée seconde » :

$$f'' : U \to \mathscr{L}_2(E ; F),$$

en notant pour abréger, $\mathscr{L}_2(E\,;F)$ l'espace de Banach $\mathscr{L}(E, E\,;F)$ formé des applications bilinéaires continues $E \times E \to F$. En général, on notera $\mathscr{L}_n(E\,;F)$ l'espace des applications multilinéaires continues

$$\underbrace{E \times \ldots \times E}_{\text{n facteurs}} \to F.$$

On peut se demander si l'application f'' est elle-même différentiable. Si elle l'est au point $a \in U$, on notera $f'''(a)$, ou encore $f^{(3)}(a)$, la dérivée de f'' au point a ; c'est un élément de $\mathscr{L}(E\,;\mathscr{L}_2(E\,;F)) \approx \mathscr{L}_3(E\,;F)$.

Par récurrence sur n, on définit : « f est n fois différentiable au point a », et on dit ce qu'on entend par la dérivée n-ième $f^{(n)}(a) \in \mathscr{L}_n(E\,;F)$. Supposons ces notions déjà définies pour $n - 1$; on dira que f est n fois différentiable en a s'il existe un voisinage ouvert V de a tel que f soit $n - 1$ fois différentiable en chaque point de V, et si l'application $x \mapsto f^{(n-1)}(x)$ de V dans $\mathscr{L}_{n-1}(E\,;F)$ est différentiable au point a ; alors la dérivée de $f^{(n-1)}$ au point a se note $f^{(n)}(a)$ et s'appelle la dérivée n-ième de f au point a. C'est un élément de $\mathscr{L}_n(E\,;F)$.

Pour $h_1, \ldots, h_n \in E$, on notera $f^{(n)}(a)\cdot(h_1, \ldots, h_n)$ la valeur de $f^{(n)}(a) : E \times \ldots \times E \to F$ pour l'élément $(h_1, \ldots, h_n) \in E \times \ldots \times E$.

DÉFINITION. On dit que f est *de classe* C^n dans U (ou encore, que f est n fois continûment différentiable dans U), si f est n fois différentiable en tout point de U, et si l'application

$$f^{(n)} : U \to \mathscr{L}_n(E\,;F)$$

est *continue.*

Ayant ainsi défini $f^{(n)}$ pour $n \geqslant 1$ (lorsque cette dérivée n-ième existe), on convient de poser

$$f^{(0)} = f \text{ (dérivée zéro-ième !)}.$$

On dira que f est *de classe* C^0 si f est continue.

DÉFINITION. On dit que $f : U \to F$ est de classe C^∞ si elle est de classe C^n pour tout n.

Exercice : Pour cela, il suffit que $f^{(n)}$ existe pour tout n ; on dit aussi, dans ce cas, que f est *indéfiniment différentiable.*

Remarque. Pour que f soit n fois différentiable au point a ($n \geqslant 1$), il faut et il suffit que $f'(x)$ existe en tout point x d'un voisinage ouvert V de a, et que l'application $f' : V \to F$ soit $n - 1$ fois différentiable au point a ; alors

$$f^{(n)}(a) = (f')^{(n-1)}(a).$$

De même, on a pour $n \geqslant 2$,

$$f^{(n)}(a) = (f'')^{(n-2)}(a), \quad \text{etc.}$$

La vérification est laissée à titre d'exercice.

Du théorème fondamental 5.1.1, on va déduire sans peine :

THÉORÈME 5.3.1. Si f est n fois différentiable au point a, la dérivée $f^{(n)}(a) \in \mathscr{L}_n(E\,;F)$ est une *application multilinéaire symétrique* $E \times \ldots \times E \to F$. Autrement dit, si $h_1, \ldots, h_n$ sont n vecteurs de E, et si σ désigne une permutation quelconque sur $[1, 2, \ldots, n]$, on a

$$f^{(n)}(a) \cdot (h_1, h_2, \ldots, h_n) = f^{(n)}(a) \cdot (h_{\sigma(1)}, h_{\sigma(2)}, \ldots, h_{\sigma(n)}). \tag{5.3.1}$$

DÉMONSTRATION. La question ne se pose que pour $n \geqslant 2$. Pour $n = 2$, c'est déjà démontré (théorème 5.1.1). On va procéder par récurrence : soit $n \geqslant 3$, et supposons le théorème démontré pour $n - 1$. Alors $f^{(n)}(a)$ est la dérivée de l'application

$$f^{(n-1)} : V \to \mathscr{L}_{n-1}(E\,;F),$$

qui par hypothèse existe dans un voisinage V de a. Par l'hypothèse de récurrence, $f^{(n-1)}$ prend ses valeurs dans le sous-espace de $\mathscr{L}_{n-1}(E\,;F)$ formé des applications $(n-1)$-linéaires *symétriques*. Donc, pour $h_1 \in E$, $f^{(n)}(a) \cdot h_1$ est un élément de cet espace ; autrement dit :

$$(f^{(n)}(a) \cdot h_1) \cdot (h_2, \ldots, h_n)$$

est une fonction *symétrique* de $h_2, \ldots, h_n$. Or c'est

$$f^{(n)}(a) \cdot (h_1, h_2, \ldots, h_n),$$

et on voit donc déjà que l'application multilinéaire $f^{(n)}(a) : E^n \to F$ est une fonction symétrique des $n - 1$ dernières variables. Il suffira donc de montrer que

$$f^{(n)}(a) \cdot (h_1, h_2, \ldots, h_n)$$

ne change pas de valeur lorsqu'on permute h_1 et h_2 ; en effet, on sait que toute permutation sur n éléments est composée d'un nombre fini de « transpositions », dont chacune consiste à permuter deux éléments consécutifs. On sait déjà que cela ne change rien si ces 2 éléments sont h_i et h_{i+1}, avec $2 \leqslant i \leqslant n - 1$; et si on prouve qu'il en est de même pour h_1 et h_2, tout sera démontré. Or $f^{(n)}(a)$ est la dérivée seconde de $f^{(n-2)}$, donc

$$(f^{(n)}(a) \cdot h_1) \cdot h_2 \in \mathscr{L}_{n-2}(E\,;F)$$

est symétrique en h_1 et h_2, d'après le théorème 5.1.1 appliqué à la fonction $f^{(n-2)}$.

C.Q.F.D.

5.4 Exemples de fonctions n fois différentiables

PROPOSITION 5.4.1. Toute application *bilinéaire continue*

$$\varphi : E_1 \times E_2 \to F$$

est de classe C^∞ ; plus précisément, φ'' est une application constante, et les dérivées $\varphi^{(n)}$ sont donc nulles pour $n > 2$.

DÉMONSTRATION. D'après le théorème 2.4.3, on sait que φ est différentiable, et que

$$\varphi'(x_1, x_2) \cdot (h_1, h_2) = \varphi(h_1, x_2) + \varphi(x_1, h_2).$$

Cette relation montre que l'application

$$\varphi' : E_1 \times E_2 \to \mathscr{L}(E_1, E_2 ; F)$$

est une fonction *linéaire continue* du point $(x_1, x_2) \in E_1 \times E_2$. Donc sa dérivée φ'' est *constante* ; la valeur de cette constante est l'élément de $\mathscr{L}_2(E_1 \times E_2 ; F)$ [où $E_1 \times E_2$ est pris avec sa structure d'espace de Banach] qui, à deux éléments (h_1, h_2) et (k_1, k_2) de l'espace vectoriel $E_1 \times E_2$, associe

$$\varphi(h_1, k_2) + \varphi(k_1, h_2).$$

La démonstration est achevée.

THÉORÈME 5.4.2 (*dérivées d'une fonction composée*). *Soient* $U \subset E$ *et* $V \subset F$ *deux ouverts d'espaces de Banach,* $f : U \to V$ *et* $g : V \to G$ *deux applications continues.*

(i) *Si* f *est* n *fois différentiable au point* $a \in U$, *et si* g *est* n *fois différentiable au point* $b = f(a) \in V$, *alors* $h = g \circ f : U \to G$ *est* n *fois différentiable au point* a.

(ii) *Si* f *et* g *sont de classe* C^n, *alors* $h = g \circ f$ *est de classe* C^n.

DÉMONSTRATION. Le théorème est vrai pour $n = 1$. En effet, cela résulte du théorème 2.2.1 (dérivée d'une fonction composée), qui donne

$$h'(x) = g'(f(x)) \circ f'(x), \tag{5.4.1}$$

formule qui montre que si f' et g' sont des fonctions continues, h' est une fonction continue (d'où l'assertion (ii) pour $n = 1$). On va prouver (i) et (ii) par récurrence sur n, en supposant que ce soit vrai pour $n - 1$ (avec $n \geqslant 2$).

Raisonnons par exemple dans le cas de la propriété (ii) ; le raisonnement serait tout semblable pour (i). On veut montrer que h est de classe C^n, ou, ce qui revient au même, que h' est de classe C^{n-1}. Or la relation (5.4.1) exprime que h' est composée de deux applications :

1°) l'application $x \mapsto (g'(f(x)), f'(x))$ de U dans $\mathscr{L}(F ; G) \times \mathscr{L}(E ; F)$;
2°) l'application $(v, u) \mapsto v \circ u$ de $\mathscr{L}(F ; G) \times \mathscr{L}(E ; F)$ dans $\mathscr{L}(E ; G)$.

La seconde application est bilinéaire continue (cf. la fin du n° 1.8), donc est de classe C^∞ (proposition 5.4.1). La première application prend ses valeurs dans un espace-produit ; ses deux composantes sont

$$x \mapsto g'(f(x)) \quad \text{et} \quad x \mapsto f'(x).$$

Par hypothèse, la seconde de ces applications est de classe C^{n-1}. Quant à la première, elle est composée de

$$U \xrightarrow{f} V \xrightarrow{g'} \mathscr{L}(F\,;G)\,;$$

f est de classe C^n, et *a fortiori* de classe C^{n-1} ; g' est aussi de classe C^{n-1}. *Par l'hypothèse de récurrence*, la composée $g' \circ f$ est de classe C^{n-1}. Ainsi l'application 1°) est de classe C^{n-1} (parce que chacune de ses deux composantes est de classe C^{n-1}) ; l'application 2°) est aussi de classe C^{n-1} (et même C^∞). *Par l'hypothèse de récurrence* (appliquée une deuxième fois) leur composée est de classe C^{n-1}. Cette composée est h', et la récurrence est démontrée.

THÉORÈME 5.4.3. Soient E et F deux espaces de Banach ; on note toujours Isom(E ; F) l'ouvert de $\mathscr{L}(E\,;F)$ formé des isomorphismes linéaires de E sur F. Alors *l'application* $\varphi : \mathrm{Isom}(E\,;F) \to \mathscr{L}(F\,;E)$, *telle que*

$$\varphi(u) = u^{-1} \in \mathrm{Isom}(F\,;E),$$

est de classe C^∞.

DÉMONSTRATION. D'après le théorème 2.4.4, on sait déjà que φ est de classe C^1, et que

$$\varphi'(u)\cdot h = -u^{-1} \circ h \circ u^{-1} \quad \text{pour} \quad h \in \mathscr{L}(E\,;F). \tag{5.4.2}$$

$\varphi'(u)$ est un élément de $\mathscr{L}(\mathscr{L}(E\,;F)\,;\mathscr{L}(F\,;E))$. Comme dans la démonstration du théorème 2.4.4, introduisons l'application bilinéaire continue

$$\psi : \mathscr{L}(F\,;E) \times \mathscr{L}(F\,;E) \to \mathscr{L}(\mathscr{L}(E\,;F)\,;\mathscr{L}(F\,;E))$$

définie par

$$\psi(v, w)\cdot h = -v \circ h \circ w.$$

La relation (5.4.2) s'écrit alors

$$\varphi'(u) = \psi(\varphi(u), \varphi(u)) \tag{5.4.3}$$

[puisque $u^{-1} = \varphi(u)$]. C'est là une « équation différentielle » à laquelle satisfait la fonction φ. On va en déduire, par récurrence sur n, que φ est de classe C^n.

On le sait pour $n = 1$. Soit $n \geqslant 2$, et supposons prouvé que φ est de classe C^{n-1}. On va montrer que φ' est de classe C^{n-1} (c'est-à-dire : φ est de classe C^n). Or (5.4.3) exprime que l'application φ' est composée de deux applications :

1°) l'application $u \mapsto (\varphi(u), \varphi(u))$ de Isom(E ; F) dans $\mathscr{L}(F\,;E) \times \mathscr{L}(F\,;E)$;

2°) l'application bilinéaire ψ.

La première est de classe C^{n-1} par l'hypothèse de récurrence, et la deuxième est de classe C^∞ (proposition 5.4.1). Donc leur composée est de classe C^{n-1} (théorème 5.4.2).

C.Q.F.D.

Exercice. Démontrer la formule explicite suivante qui donne la dérivée n-ième de φ :

$$\varphi^{(n)}(u) \cdot (h_1, \ldots, h_n) = (-1)^n \sum_{\sigma} u^{-1} \circ h_{\sigma(1)} \circ u^{-1} \circ \ldots \circ u^{-1} \circ h_{\sigma(n)} \circ u^{-1},$$

où la sommation est étendue aux $n!$ permutations σ de $[1, \ldots, n]$.

THÉORÈME 5.4.4. *Soient* E *et* F *deux espaces de Banach, et soient* $V \subset E$ *et* $W \subset F$ *des ouverts. Soit*

$$f : V \to W$$

un C^1*-difféomorphisme* (cf. 4.1). *Si l'application* f *est de classe* C^n, *alors l'homéomorphisme réciproque* $g = f^{-1}$ *est aussi de classe* C^n. [On dit alors que f est un C^n-*difféomorphisme*].

DÉMONSTRATION. Pour $n = 1$, l'assertion de l'énoncé est une tautologie. De plus, on sait que, pour $y \in W$:

$$g'(y) = (f'(g(y)))^{-1} ; \qquad (5.4.3)$$

ceci exprime que l'application g' est composée de trois applications :

— l'application $g : W \to V$;
— l'application $f' : V \to \mathrm{Isom}(E ; F)$;
— l'application $\mathrm{Isom}(E ; F) \to \mathscr{L}(F ; E)$ définie par $u \mapsto u^{-1}$.

On va alors démontrer le théorème par récurrence sur n. Supposons-le vrai pour $n - 1$ $(n \geqslant 2)$; alors, sous les hypothèses de l'énoncé, la deuxième et la troisième application ci-dessus sont de classe C^{n-1} (la troisième est même de classe C^∞, d'après le théorème 5.4.3). Quant à la première, c'est g ; *par l'hypothèse de récurrence,* elle est de classe C^{n-1}. Alors l'application g', composée de trois applications de classe C^{n-1}, est de classe C^{n-1} (théorème 5.4.2).

C.Q.F.D.

Remarque. Si un homéomorphisme $f : V \to W$ est de classe C^n $(n \geqslant 1)$ (resp. de classe C^∞) et si $f'(x) \in \mathrm{Isom}(E ; F)$ pour tout $x \in V$, alors f est un C^n-difféomorphisme (resp. un C^∞-difféomorphisme). [Pour $n = 1$, ceci n'est pas autre chose que la proposition 4.1.1 ; en la joignant au théorème 5.4.4 ci-dessus, on obtient le résultat énoncé].

COROLLAIRE 5.4.5. Dans le « théorème d'inversion locale » (théorème 4.2.1), si on suppose que f est non seulement de classe C^1, mais de classe C^n, on conclut que la restriction de f à V [notation du théorème 4.2.1] *est un* C^n*-difféomorphisme* de V sur W.

De même, dans le « théorème des fonctions implicites » (théorème 4.7.1), si on suppose que l'application $(x, y) \mapsto f(x, y)$ est non seulement de classe C^1, mais de classe C^n, on conclut [avec les notations du théorème 4.7.1] que *l'application* $g : W \to F$ *est de classe* C^n.

Commençons par une formule préliminaire. Soient E, F et G trois espaces de Banach, et $\varphi : E \times F \to G$ une application bilinéaire continue. Soient d'autre part

$$u : U \to E \quad \text{et} \quad v : U \to F$$

deux applications $n + 1$ fois différentiables, U désignant un intervalle ouvert de la *droite numérique* **R**. Les dérivées successives $u^{(i)}$, $v^{(i)}$ prennent leurs valeurs respectivement dans E et F.

Lemme. Sous les hypothèses précédentes, l'application

$$t \mapsto \sum_{p=0}^{n} (-1)^p \varphi(u^{(p)}(t), v^{(n-p)}(t))$$

de U *dans* G *a pour dérivée*

$$t \mapsto \varphi(u(t), v^{(n+1)}(t)) + (-1)^n \varphi(u^{(n+1)}(t), v(t)).$$

La vérification est laissée au lecteur : on applique la formule qui donne la dérivée d'une fonction bilinéaire de deux fonctions d'une variable numérique (cf. (2.5.5)).

On va appliquer ce lemme dans le cas particulier suivant : $E = \mathbf{R}$, $G = F$, l'application $\varphi : \mathbf{R} \times F \to F$ étant la multiplication d'un vecteur de F par un scalaire. De plus nous prenons

$$u(t) = \frac{1}{n!}(1 - t)^n ;$$

qui est de classe C^n, avec $u^{(n+1)}(t) = 0$. On obtient :

PROPOSITION 5.5.1. *Si* v *est une fonction* $(n + 1)$ *fois différentiable d'une variable réelle* $t \in U$, *à valeurs dans l'espace de Banach* F, *on a*

$$\frac{d}{dt}[v(t) + (1 - t)v'(t) + \ldots + \frac{1}{n!}(1 - t)^n v^{(n)}(t)] = \frac{1}{n!}(1 - t)^n v^{(n+1)}(t). \tag{5.5.1}$$

(la notation $(d/dt)f$ désigne la dérivée d'une fonction f de la variable réelle t).

COROLLAIRE 5.5.2. Supposons en outre que $U \supset [0, 1]$, et que $v^{(n+1)}$ soit *continue*. Alors

$$v(1) - v(0) - v'(0) - \frac{1}{2}v''(0) \ldots - \frac{1}{n!}v^{(n)}(0) = \int_0^1 \frac{(1 - t)^n}{n!} v^{(n+1)}(t)\, dt. \tag{5.5.2}$$

En effet, si $t \mapsto f(t)$ a une dérivée continue f' pour $t \in [0, 1]$, on sait que

$$f(1) - f(0) = \int_0^1 f'(t)\, dt$$

[cf. cours de Math I]. On applique ce résultat en prenant ici

$$f(t) = v(t) + (1-t)v'(t) + \ldots + \frac{1}{n!}(1-t)^n v^{(n)}(t). \tag{5.5.3}$$

COROLLAIRE 5.5.3. Sous les hypothèses de la proposition 5.5.1, supposons que

$$\|v^{(n+1)}(t)\| \leqslant M \quad \text{pour} \quad t \in [0, 1]. \tag{5.5.4}$$

Alors on a

$$\|v(1) - v(0) - v'(0) - \frac{1}{2}v''(0) \ldots - \frac{1}{n!}v^{(n)}(0)\| \leqslant \frac{M}{(n+1)!}. \tag{5.5.5}$$

DÉMONSTRATION. On va appliquer le théorème 3.1.1 (théorème des accroissements finis), en remplaçant dans ce théorème l'intervalle $[a, b]$ par $[0, 1]$, la fonction f par celle donnée par (5.5.3), et en prenant

$$g(t) = -M\frac{(1-t)^{n+1}}{(n+1)!}.$$

La relation (5.5.1) dit que

$$\|f'(t)\| \leqslant \frac{(1-t)^n}{n!}\|v^{(n+1)}(t)\| ;$$

donc, d'après l'hypothèse (5.5.4) :

$$\|f'(t)\| \leqslant M\frac{(1-t)^n}{n!} = g'(t).$$

Le théorème des accroissements finis 3.1.1 permet de conclure :

$$\|f(1) - f(0)\| \leqslant g(1) - g(0),$$

ce qui donne précisément l'inégalité (5.5.5) à démontrer.

Les corollaires 5.5.2 et 5.5.3 constituent deux cas particuliers de la « formule de Taylor », qu'on va maintenant considérer dans le cas général.

Formule de Taylor : cas général 5.6

Désormais, U désigne un ouvert d'un espace de Banach E, et F un espace de Banach ; on considère une application

$$f : U \to F.$$

Soient a et $a + h$ deux points de U, tels que le segment $[a, a+h]$ soit contenu dans U (par exemple, si U est convexe, il suffit que $a \in U$ et $a + h \in U$ pour que le segment soit contenu dans U ; si U est un ouvert quelconque, et a un point de U, alors $a + h \in U$ pour tout vecteur $h \in E$ de norme assez petite).

Considérons la fonction

$$v(t) = f(a + th), \qquad t \in [0, 1].$$

Si f est $n + 1$ fois différentiable dans U, v est $n + 1$ fois différentiable (différentiabilité d'une fonction composée), et on calcule immédiatement les dérivées de v :

$$v'(t) = f'(a + th) \cdot h$$

$$v''(t) = (f''(a + th) \cdot h) \cdot h,$$

qu'on a convenu de noter $f''(a + th) \cdot (h, h)$ [ne pas oublier que $f''(a + th)$ est une application bilinéaire, d'ailleurs symétrique, de $E \times E$ dans F]. D'une manière générale, on voit par récurrence sur n que

(5.6.1) $$v^{(n)}(t) = f^{(n)}(a + th) \cdot \underbrace{(h, \ldots, h)}_{n \text{ fois}}.$$

Pour abréger, on conviendra de noter $(h)^n$ l'élément $(h, \ldots, h) \in E^n$.

Dans les corollaires 5.5.2 et 5.5.3, remplacons v et ses dérivées par les expressions (5.6.1). On obtient :

THÉORÈME 5.6.1 (*«formule de Taylor avec reste intégral»). Soit* $f : U \to F$ *une application de classe* C^{n+1}. *Si le segment* $[a, a + h]$ *est contenu dans* U, *on a* :

(5.6.2) $$\boxed{f(a + h) = f(a) + f'(a) \cdot h + \frac{1}{2} f''(a) \cdot (h, h) + \ldots + \frac{1}{n!} f^{(n)}(a) \cdot (h)^n + \int_0^1 \frac{(1 - t)^n}{n!} f^{(n+1)}(a + th) \cdot (h)^{n+1} \, dt.}$$

THÉORÈME 5.6.2 (*«formule de Taylor avec reste de Lagrange»). Soit* $f : U \to F$ *une application* $n + 1$ *fois différentiable ; supposons*

(5.6.3) $$\|f^{(n+1)}(x)\| \leqslant M \quad \text{pour} \quad x \in U.$$

Alors :

(5.6.4) $$\boxed{\| f(a + h) - f(a) - f'(a) \cdot h - \ldots - \frac{1}{n!} f^{(n)}(a) \cdot (h)^n \| \leqslant M \frac{\|h\|^{n+1}}{(n + 1)!}.}$$

En effet,

$$\|v^{(n+1)}(t)\| = \|f^{(n+1)}(a + th) \cdot (h, \ldots, h)\| ;$$

par la propriété de la norme d'une application $(n + 1)$-linéaire continue (cf. (1.8.5)), ceci est majoré par

$$\|f^{(n+1)}(a + th)\| \cdot \|h\|^{n+1},$$

et par l'hypothèse (5.6.3) ceci est majoré par $M \cdot \|h\|^{n+1}$. Il suffit alors d'appliquer le corollaire 5.5.3 (où M serait remplacé par $M\|h\|^{n+1}$).

Telles sont les deux « formules de Taylor ». On en déduit une troisième « formule de Taylor » : dans (5.6.4), on voit que si h tend vers 0, le second membre est $o(\|h\|^n)$, donc le premier membre aussi. Mais ce résultat a été obtenu en supposant que f a une dérivée $f^{(n+1)}$ bornée au voisinage de a. En réalité, il vaut sous des hypothèses plus larges :

THÉORÈME 5.6.3. *Soit* $f : U \to F$ *une application* $n - 1$ *fois différentiable. Supposons que* f *soit* n *fois différentiable au point* $a \in U$. *On a alors*

$$\boxed{\|f(a + h) - f(a) - f'(a) \cdot h \ldots - \frac{1}{n!} f^{(n)}(a) \cdot (h)^n\| = o(\|h\|^n).} \tag{5.6.5}$$

Cette « formule de Taylor » exprime seulement une propriété « asymptotique » ; elle dit ce qui se passe *quand* h *tend vers zéro.*

DÉMONSTRATION. Pour $n = 1$, (5.6.5) ne dit rien de plus que la *définition* de la dérivée $f'(a)$:

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

On va procéder par récurrence sur n, en supposant (5.6.5) vrai pour $n - 1$ $(n \geqslant 2)$. Considérons la fonction

$$\varphi(h) = f(a + h) - f(a) - f'(a) \cdot h - \ldots - \frac{1}{n!} f^{(n)}(a) \cdot (h)^n. \tag{5.6.6}$$

Calculons sa dérivée. Pour cela, cherchons d'abord la dérivée de la fonction $h \mapsto f^{(n)}(a) \cdot (h)^n$; cette dérivée est, pour chaque valeur de h, un élément de $\mathscr{L}(E\,;F)$, c'est-à-dire une fonction linéaire de $k \in E$ à valeurs dans F. Comme $f^{(n)}(a)$ est une fonction n-linéaire $E \times \ldots \times E \to F$, la relation (2.4.3) nous donne sa dérivée pour la valeur $(h, \ldots, h)$ de la variable : c'est l'application linéaire

$$k \mapsto f^{(n)}(a) \cdot (k, h, \ldots, h) + f^{(n)}(a) \cdot (h, k, h, \ldots, h) + \ldots + f^{(n)}(a) \cdot (h, \ldots, h, k).$$

Comme $f^{(n)}(a)$ est une fonction *symétrique*, ceci fait $k \mapsto nf^{(n)}(a) \cdot (h, \ldots, h, k)$. On peut interpréter cela comme suit : considérons $f^{(n)}(a)$ comme la dérivée $(n - 1)$-ième de $f' : U \to \mathscr{L}(E\,;F)$; c'est une fonction $(n - 1)$-linéaire symétrique à valeurs dans $\mathscr{L}(E\,;F)$. Notons

$$f^{(n)}(a) \cdot \underbrace{(h, \ldots, h)}_{n - 1 \text{ fois}} = f^{(n)}(a) \cdot (h)^{n-1}$$

sa valeur sur le multivecteur $(h, \ldots, h)$; c'est un élément de $\mathscr{L}(E\,;F)$. Alors on voit que la dérivée de la fonction

$$h \mapsto \frac{1}{n!} f^{(n)}(a) \cdot (h)^n \quad \text{(application } E \to F\text{)}$$

est

$$h \mapsto \frac{1}{(n-1)!} f^{(n)}(a) \cdot (h)^{n-1} \quad \text{(application } E \to \mathscr{L}(E\,;F)\text{)}.$$

Avec ces explications, la dérivée de la fonction φ définie par (5.6.6) est

$$\varphi'(h) = f'(a+h) - f'(a) - \ldots - \frac{1}{(n-1)!} f^{(n)}(a) \cdot (h)^{n-1}.$$

Appliquons l'hypothèse de récurrence à la fonction f' ; on obtient :

$$\|\varphi'(h)\| = o(\|h\|^{n-1}).$$

Autrement dit, pour tout $\varepsilon > 0$ il existe $\eta > 0$ tel que

$$\|h\| \leqslant \eta \quad \text{entraîne} \quad \|\varphi'(h)\| \leqslant \varepsilon \|h\|^{n-1}.$$

L'inégalité des accroissements finis entraîne alors

$$\|\varphi(h) - \varphi(0)\| \leqslant \varepsilon \|h\|^n \quad \text{pour} \quad \|h\| \leqslant \eta.$$

D'ailleurs $\varphi(0) = 0$. Donc on a

$$\|\varphi(h)\| = o(\|h\|^n),$$

ce qui est justement la relation (5.6.5) à démontrer.

6. POLYNOMES

La formule de Taylor (cf. 5.6) introduit la fonction de $h \in E$:

$$h \mapsto \frac{1}{n!} f^{(n)}(a) \cdot \underbrace{(h, \ldots, h)}_{\text{n fois}}\,;$$

rappelons que $f^{(n)}(a)$ est une application multilinéaire symétrique $E^n \to F$. Ceci va nous conduire à la notion générale *d'application polynomiale homogène de degré* n *de* E *dans* F.

La question a d'abord un aspect purement algébrique, que nous allons développer pour commencer.

Polynomes homogènes de degré n 6.1

Dans ce numéro et les suivants, K désigne un corps commutatif de caractéristique zéro, c'est-à-dire contenant le corps $\mathbf{Q}$ des rationnels. Il n'est pas nécessaire de supposer que K soit $\mathbf{R}$ ou $\mathbf{C}$; en particulier, K pourrait être égal à $\mathbf{Q}$. Tous les espaces vectoriels considérés sont des espaces vectoriels sur K, de dimension finie ou infinie.