Calcul différentiel

Christophe Antonini[1], Olivier Teytaud[2], Pierre Borgnat[3], Annie Chateau[4], and Edouard Lebeau[5]

[1]Enseignant en CPGE, Institut Stanislas, Cannes
[2]Chargé de rechercher INRIA, Université d'Orsay, Orsay
[3]Chargé de recherche CNRS, ENS Lyon, Lyon
[4]Maitre de conférence, Université Montpellier-2, Montpellier
[5]Enseignant en CPGE, Lycée Henri Poincaré, Nancy

9 juillet 2022

Un chapitre consacré au calcul différentiel. On y introduit les notions de différentielle, de dérivées partielles. Le chapitre est ponctué par deux théorèmes fondamentaux : le théorème d'inversion locale et le théorème des fonctions implicites.

1 Calcul différentiel

Il est recommandé de bien maîtriser la partie **??** avant d'étudier cette partie, et notamment les espaces de Banach. On pourra alors dans ce chapitre, après l'introduction, plonger dans le théorème des accroissements finis, d'inversion locale, des fonctions implicites, dans aussi les dérivées d'ordre supérieur, avant un peu de zoologie pour finir.

1.1 Introduction

1.1.1 Généralités

DÉFINITION 0.1 **différentiable**
Soient E et F des espaces vectoriels normés et U un ouvert de E. Soit une application $f : U \to F$, on dit que f est **différentiable** (ou **dérivable**) en $x \in U$ s'il existe une application linéaire

continue ϕ de E dans F telle que

$$lim_{h\to 0}\frac{f(x+h)-f(x)-\phi(h)}{\parallel h \parallel}=0$$

On appelle ϕ la **différentielle** ou **dérivée** de f en x, on la note $Df(x)$.

f est dite **différentiable** si elle est différentiable en tout point de U. Si E est euclidien, on appelle **gradient** de f en x, lorsque $Df(x)$ existe et lorsque F est un espace vectoriel réel de dimension 1 le vecteur ∇f_x tel que $Df(x)(u) =< \nabla f_x | u >$.

PROPOSITION 0.1

•Si f est dérivable en x, alors f est continue en x.

•La dérivée de f est unique et

$$Df(x)(h) = lim_{t\to 0}\frac{f(x+t.h)-f(x)}{t}$$

•La notion de dérivée ne dépend que des topologies et pas des normes (du moment qu'elles définissent la même topologie) ; si deux normes sont équivalentes, alors une fonction différentiable pour l'une est différentiable pour l'autre, et la différentielle est la même.

•$D(\lambda f + \mu g)(x) = \lambda Df(x) + \mu Dg(x)$

•Si $E = K$ corps associé aux espaces vectoriels E et F, alors la différentiabilité équivaut à l'existence de la limite pour $t \to 0$ de $\frac{f(x+t)-f(x)}{t}$. On note alors cette limite $f'(x)$, et $Df(x)(t) = t.f'(x)$.

•L'application qui à une application différentiable en x_0 associe sa différentielle en x_0 est une application linéaire de l'espace vectoriel des applications de E dans F différentiables en x_0 dans l'espace vectoriel des applications linéaires de E dans F.

•Une application linéaire continue f est différentiable en tout point x_0 et $Df(x_0)(h) = f(h)$.

Démonstration *Un peu laborieux mais rien de bien difficile, en notant $\epsilon(x,h) = f(x+h) - f(x) - Df(x)(h)$, pour h suffisamment petit pour que $x+h$ appartienne à U.*

DÉFINITION 0.2 **de classe C^1**

Soient E et F des espaces vectoriels normés et U ouvert de E. f de U dans F est **de classe C^1** si elle est différentiable et si l'application qui à x associe la différentielle de f en x est continue (voir **??** pour un rappel de la topologie usuelle sur $\mathcal{L}(E,F)$).

PROPOSITION 0.2

Quelques exemples de fonctions C^1 :

•Si f est constante, alors f est C^1 et sa dérivée est nulle partout.

•Si ϕ de E dans F est linéaire continue, alors ϕ est C^1 avec $D\phi(x) = \phi$, pour tout x.

•Si f de $E_1 \times E_2 \times ... \times E_n$ dans F est multilinéaire continue, alors f est C^1, et on a

$$Df(x_1,...,x_n)(h_1,...,h_n) = \sum_{i=1}^{n} f(x_1,x_2,...,x_{i-1},h_i,x_{i+1},...,x_n)$$

Théorème 0.3 **Différentielle de fonctions composées**

Soit E,F et G des espaces vectoriels normés , et U et V des ouverts de E et F respectivement. Si f de U dans V est différentiable en x et g de V dans G est différentiable en $f(x)$, alors la composée $g \circ f$ est différentiable en x et a pour différentielle

$$D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$$

Si g et f sont C^1 alors $g \circ f$ est C^1.

Démonstration *on écrit comme pour d'autres preuves* $f(x+h) = f(x) + Df(x)(h) + \epsilon(h) \parallel h \parallel$, *et de même* $g((f(x)+k)$, *et on calcule...*

Pour voir que la composée est C^1, *il suffit de voir que la différentielle est la composée de* 3 *applications continues.*

Définition 0.3 **Isomorphisme d'espaces normés**

Un **isomorphisme de l'espace vectoriel normé** E **sur l'espace vectoriel normé** F est une application $\phi : E \to F$ linéaire continue et bijective d'inverse continue. On note $Isom(E,F)$ le sous-ensemble de $\mathcal{L}(E,F)$ formé des isomorphismes de E dans F.

Théorème 0.4

Soient E et F des espaces de Banach. Le sous-ensemble $Isom(E,F)$ est ouvert dans $\mathcal{L}(E,F)$. L'application $inv : Isom(E,F) \to Isom(F,E)$ qui à u associe u^{-1} est C^1 avec $Dinv(u)(v) = -u^{-1}.v.u^{-1}$.

Démonstration *Il suffit bien sûr de traiter le cas où* $Isom(E,F)$ *n'est pas vide.*

Soit u_0 *un isomorphisme de* E *vers* F ; *alors* $u_0 + v = u_0.(Id + u_0^{-1}.v)$. *Si* $\parallel v \parallel < \parallel u_0^{-1} \parallel^{-1}$, *on a* $\parallel u_0^{-1}.v \parallel < 1$, *et donc la série*

$$\sum_{i=0}^{+\infty} (-u_0^{-1}.v)^i$$

est convergente dans $\mathcal{L}(E)$ *et sa somme donne un inverse à* $(Id + u_0^{-1}.v)$.

Donc pour $\parallel v \parallel < \parallel u_0^{-1} \parallel^{-1}$, *l'inverse de* $u_0 + v$ *est* $\sum_{i=0}^{+\infty} (-u_0^{-1}.v)^i.u_0^{-1}$.

On a alors $(u_0+v)^{-1} = u_0^{-1} - u_0^{-1}.v.u_0^{-1} + \sum_{i=2}^{+\infty} (-u_0^{-1}.v)^i.u_0^{-1}$. *La norme de cette somme est majorée par* $\sum_{i=2}^{+\infty} \parallel -u_0^{-1}.v \parallel^i \parallel u_0^{-1} \parallel = \frac{\|v\|^2 \|u_0^{-1}\|^3}{1 - \|v\| . \|u_0^{-1}\|}$, *qui est négligeable devant* $\parallel v \parallel$. *L'application* $Dinv$ *est continue par composition.*

Proposition 0.5 **Liens entre différentiabilité sur $\mathbb{R}$ et sur $\mathbb{C}$**

Un $\mathbb{C}$-espace vectoriel peut aussi être considéré comme un $\mathbb{R}$-espace vectoriel ; il suffit de restreindre le produit par un scalaire à un produit par un scalaire réel. En remplaçant E et F en tant que $\mathbb{C}$-espaces vectoriels par E et F en tant que $\mathbb{R}$-espaces vectoriels , une fonction différentiable pour $\mathbb{C}$ est différentiable pour $\mathbb{R}$; par contre la réciproque n'est pas garantie dans le cas général ; il faut que la différentielle sur $\mathbb{R}$ soit définie <u>et</u> que la différentielle sur $\mathbb{R}$ soit linéaire en tant qu'application entre $\mathbb{C}$-espaces vectoriels (c'est-à-dire appartienne à $\mathcal{L}_{\mathbb{C}}(E,F)$).

Démonstration *La preuve, sans grande difficulté, est laissée au lecteur.*

1.1.2 Applications à valeurs dans un produit d'espaces vectoriels normés

PROPOSITION 0.6
Soit f une application de E dans $F_1 \times F_2$, avec E, F_1 et F_2 des espaces vectoriels normés , et soient f_1 et f_2 ses composantes. Alors f est différentiable en x si et seulement si f_1 et f_2 sont différentiables en x, et $Df(x)(h) = (Df_1(x)(h), Df_2(x)(h))$.

En outre, f est C^1 si et seulement si f_1 et f_2 sont C^1.

Démonstration *Il suffit de voir que les projections canoniques de F sur F_1 et F_2 sont C^1 car linéaires et continues, et que l'injection canonique de F_1 dans F ou de F_2 dans F sont linéaires continues, donc elles aussi C^1.*

COROLLAIRE 0.7 **Formule de Leibnitz**

Si $f_1 : U \to F_1$ et $f_2 : U \to F_2$ sont différentiables en x et si $B : F_1 \times F_2 \to G$ est bilinéaire continue, alors $B(f_1, f_2) : x \mapsto B(f_1(x), f_2(x))$ est différentiable en x et

$$DB(f_1, f_2)(x)(h) = B(f_1(x), Df_2(x)(h)) + B(Df_1(x)(h), f_2(x))$$

En outre si f_1 et f_2 sont C^1 alors $B(f_1, f_2)$ est C^1.

1.1.3 Applications de plusieurs variables et dérivées partielles

PROPOSITION 0.8 **Définition des dérivées partielles**

Soit U un ouvert du produit $E_1 \times E_2$ de deux espaces vectoriels normés , soit $f : U \to F$, avec F espace vectoriel normé , et f différentiable en $a = (a_1, a_2)$. Alors les deux applications partielles $x_1 \mapsto f(x_1, a_2)$ et $x_2 \mapsto f(a_1, x_2)$ sont différentiables respectivement en a_1 et a_2. On note les deux différentielles obtenues respectivement $D_1 f(a_1, a_2)$ et $D_2 f(a_1, a_2)$, ou bien $\frac{\delta f}{\delta x_1}$ et $\frac{\delta f}{\delta x_2}$, et on les appelle respectivement **première dérivée partielle** et **deuxième dérivée partielle**. On a alors

$$Df(a_1, a_2)(h_1, h_2) = D_1 f(a_1, a_2)(h_1) + D_2 f(a_1, a_2)(h_2)$$

On peut généraliser de même à un produit fini d'espaces vectoriels normés ; si f est différentiable en $(a_1, a_2, ..., a_n)$, alors pour tout i dans $[\![1, n]\!]$ $x \mapsto f(a_1, ..., a_{i-1}, x, a_{i+1}, ..., a_n)$ est différentiable en a_i, sa différentielle en a_i est noté $\frac{\delta f}{\delta x_i}(a)$, et

$$Df(a_1, ..., a_n)(h_1, ..., h_n) = \sum_{i=1}^{n} \frac{\delta f}{\delta x_i}(a)(h_i)$$

Pour la cohérence de la définition, il convient de vérifier que les fonctions partielles $x_1 \mapsto f(x_1, a_2)$ et $x_2 \mapsto f(a_1, x_2)$ sont différentiables.

⚠ *Attention 0.1* Il n'y a pas de réciproque dans le cas général ! Même si toutes les dérivées partielles sont définies la différentielle n'est pas nécessairement définie. Par contre si les différentielles partielles existent *et* sont continues alors on peut conclure que f est différentiable et même C^1 (voir partie 1.2.3).

THÉORÈME 0.9
Soit E un espace vectoriel normé, U un ouvert de E et $F_1, ..., F_n$ des espaces vectoriels normés. Soit f de U dans $F_1 \times ... \times F_n$.

On note $p_i(x_1, ..., x_n) = x_i$.

Alors f est différentiable en x si et seulement si chacune des application f_i de E dans F_i, $f_i : y \mapsto p_i(f(x))$ est différentiable en x et on a alors

$$Df(x)(h) = (Df_1(x)(h_1), Df_2(x)(h_2), ..., Df_n(x)(h_n))$$

Démonstration *•Le sens « seulement si » est clair ; une composée d'applications différentiables est différentiable.*

•Le sens « si » et l'égalité annoncée s'obtiennent en considérant

$$f = \sum_{i=1}^{n} u_i \circ f_i$$

avec $u_i(x) = (0, ..., 0, x, 0, , , 0)$. f*, combinaison linéaire de composées d'applications différentiables, est bien différentiable.*

DÉFINITION 0.4 **matrice jacobienne**

Un cas particulier important est le cas où U est un ouvert de $\mathbb{R}^n$ et $F = \mathbb{R}^m$; on peut alors noter la différentielle sous forme matricielle ; cette matrice est appelée **matrice jacobienne**. Elle est de la forme :

$$\begin{pmatrix} \frac{\delta f_1}{\delta x_1} & \frac{\delta f_1}{\delta x_2} & \cdots & \frac{\delta f_1}{\delta x_n} \\ \frac{\delta f_2}{\delta x_1} & \frac{\delta f_2}{\delta x_2} & \cdots & \frac{\delta f_2}{\delta x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta f_m}{\delta x_1} & \frac{\delta f_m}{\delta x_2} & \cdots & \frac{\delta f_m}{\delta x_n} \end{pmatrix}$$

Si $n = m$, la matrice jacobienne est carrée, on peut donc considérer son déterminant, appelé **jacobien** de f.

On peut noter que dans le cas $m = 1$, la jacobienne s'identifie au gradient.

On peut définir de même une matrice pour la dérivée seconde si l'application est à valeurs de $\mathbb{R}^n$ dans $\mathbb{R}$; cette matrice s'appelle la matrice hessienne, voir la section **??**.

1.2 Le théorème des accroissements finis

On présente ici l'important théorème des accroissements finis. Son intérêt sera montré par des sections d'applications (interversion de limite et de dérivation, et existence de dérivées à partir d'existence de dérivées locales), mais aussi par un grand nombre de références (‖), dont tout ce qui concerne les développements de Taylor.

1.2.1 Résultats principaux

DÉFINITION 0.5 **dérivée à droite**
Soient a et b dans $\mathbb{R}$, avec $a < b$, et F un espace vectoriel normé . Une application f de $[a, b]$ dans F est dite **dérivable** à droite en x appartenant à $[a, b[$ si la limite à droite $lim_{h \to 0, h>0} \frac{f(x+h)-f(x)}{h}$ existe ; on l'appelle alors **dérivée à droite** de f en x.

THÉORÈME 0.10 **Inégalité des accroissements finis**

Soient a et b dans $\mathbb{R}$ avec $a < b$, et F un espace de Banach. On suppose que les deux fonctions $f : [a, b] \to F$ et $g : [a, b] \to \mathbb{R}$ sont continues sur $[a, b]$ et dérivables à droite sur $[a, b] \setminus D$ avec D au plus dénombrable. Si, pour tout $t \in [a, b] \setminus D$ on a $\| f'_d(t) \| \leq g'_d(t)$, alors $\| f(b) - f(a) \| \leq g(b) - g(a)$.

Démonstration *La démonstration dans des cadres plus simples (i.e. avec f continue sur $[a, b]$ et dérivable sur $]a, b[$) et à valeurs dans $\mathbb{R}$ est très vite faite, en appliquant le théorème de Rolle* **??** *à $x \mapsto f(x) - \frac{f(b)-f(a)}{b-a}(x - a)$. Le cas général ci-dessus évoqué est nettement plus difficile et ne sera pas discuté ici (cf [1]).*

COROLLAIRE 0.11
On a le même résultat en remplaçant les dérivées à droite par les dérivées à gauche.

Démonstration *Il suffit de remplacer dans le résultat x par $-x$.*

COROLLAIRE 0.12
Une fonction continue de $\mathbb{R}$ dans un espace vectoriel normé , et dont la dérivée existe et est nulle sauf sur un ensemble au plus dénombrable est constante.

Démonstration *Application immédiate du théorème des accroissements finis.*

Application 0.2 outre l'application qu'est ce corollaire, on utilise le théorème des accroissements finis pour les théorèmes 0.30 (caractère symétrique de la différentielle seconde), 0.21 (construction de la différentielles à partir des différentielles partielles), 0.26 (théorème d'inversion locale), **??** (dérivée sous le signe somme), **??** (inégalité de Taylor-Lagrange), **??** (compacts de $H(\Omega)$).

On l'utilise aussi pour montrer qu'une fonction dérivable à dérivée bornée est lipschitzienne, ou bien qu'une fonction C^1 est localement lipschitzienne. Notons aussi la proposition qui suit :

Proposition 0.13
Une fonction f continue de $\mathbb{R}$ dans $\mathbb{R}$ dont la dérivée existe et est positive sauf sur un ensemble au plus dénombrable est croissante.

Démonstration *L'astuce réside dans le fait que la fonction f dont il est question ici doit jouer le rôle de la fonction g du théorème des accroissements finis ! On utilise pour f une fonction nulle, donc de dérivée nulle ; on considère une fonction g de dérivée positive, et le tour est joué.*

Corollaire 0.14

Soit f définie de l'ouvert U de l'espace vectoriel normé E et à valeurs dans l'espace vectoriel normé F. Si f est dérivable et si le segment $[x, y]$ est inclus dans U, alors

$$\| f(y) - f(x) \| \leq \| y - x \| .sup_{z\in[x,y]} \| Df(z) \|$$

Démonstration *Si le sup est infini il n'y a rien à prouver. Sinon on considère la fonction θ qui à un réel t compris entre 0 et 1 associe $\|y-x\|.(sup_{z\in[x,y]}\|Df(z)\|).t$; θ est dérivable, en tout point, de dérivée constante égale à $\|y-x\|.(sup_{z\in[x,y]}\|Df(z)\|)$, que l'on va noter C. L'application ϕ qui à $t \in [0, 1]$ associe $f((1-t).x + t.y)$ est dérivable en tout point de $[0, 1]$, de dérivée $Df((1-t).x + t.y)(y-x)$. La norme de cette dérivée est majorée par θ', donc par C. On peut donc majorer $\|\phi(1) - \phi(0)\|$ par $\theta(1) - \theta(0)$.*
Voyons trois corollaires :

Corollaire 0.15
Une application définie sur un ouvert U de l'espace vectoriel normé E à valeurs dans l'espace vectoriel normé F différentiable et de différentielle nulle est localement constante. Si U est connexe, f est constante.

Corollaire 0.16

En définissant la distance entre deux points d'un ouvert connexe comme la borne inférieure des longueurs de lignes brisées entre ces deux points (voir le chapitre de topologie pour vérifier qu'un ouvert connexe d'un espace vectoriel normé est connexe par arcs et que toute paire de points dans un tel ensemble peut être reliée par une ligne brisée), et en supposant que f est une application de cet ouvert dans un espace vectoriel normé différentiable telle que pour tout x $\|Df(x)\| \leq k$, alors $\|f(b) - f(a)\|$ est inférieur ou égal à k fois la distance de a à b.

Définition 0.6 **localement lipschitzienne**
Une application **localement lipschitzienne** est une application entre espaces métriques telle que pour tout x il existe un voisinage de x sur lequel la restriction de f est lipschitzienne.

Corollaire 0.17
Une application de classe C^1 entre espaces métriques est localement lipschitzienne.

1.2.2 Applications : interversion de limite et de dérivation

Définition 0.7 **Convergence uniforme, rappel**

Une suite d'applications f_n de X dans Y avec X et Y espaces métriques converge uniformément vers f application de X dans Y si

$$lim_{n\to+\infty} sup_{x\in X} d(f_n(x), f(x)) = 0$$

Proposition 0.18

Si les f_n sont continues et convergent uniformément vers f alors f est continue.

Démonstration *Soit x dans X. Pour tout y dans X, on a :*

$$d(f(x), f(y)) \leq d(f(x), f_n(x)) + d(f_n(x), f_n(y)) + d(f_n(y), f(y)).$$

Soit $\epsilon > 0$. Prenons n assez grand pour avoir : $sup_{x\in X} d(f_n(x), f(x)) < \epsilon/3$. Par continuité de f_n, pour y suffisamment proche de x, on a : $d(f_n(x), f_n(y)) < \epsilon/3$.

En sommant, on obtient : $f(x) - f(y) \leq \epsilon$ pour y suffisamment proche de x.

Application 0.3 Ce résultat servira par exemple pour le théorème **??** (limite d'une fonction elle-même limite uniforme d'une suite de fonctions), ou le théorème 0.37 (construction d'une fonction continue partout et nulle part dérivable).

Théorème 0.19
On suppose E et F des espaces vectoriels normés , U ouvert de E, f_n une suite d'applications de U dans F différentiables, f_n convergeant simplement vers f, les Df_n convergeant uniformément vers une certaine application g de U dans $\mathcal{L}(E,F)$, alors :

- f est différentiable et $Df = g$.
- Pour tout C convexe et borné inclus dans U la convergence de $f_{n|C}$ vers $f_{|C}$ est uniforme.
- Si les f_n sont C^1 alors f est C^1.

Démonstration *laborieuse, mais pas vraiment difficile ; il suffit d'écrire $\epsilon_n = sup_{x\in U}\|Df_n(x) - g(x)\|$, avec ϵ_n tendant vers 0 lorsque n tend vers l'infini, et de montrer que $sup_{y\in C}\|f(y) - f_n(y)\| \leq \|f(x) - f_n(x)\| + \epsilon_n.D$ avec D le diamètre de C pour voir la deuxième propriété ; la première propriété se montre facilement à partir de là, et la troisième est un corollaire de la proposition 0.18.*

Corollaire 0.20

On suppose U connexe ouvert de E et f_n de U dans F différentiable ; E et F sont des espaces vectoriels normés , et F est complet (F est un Banach). On suppose qu'il existe x_0 tel que $f_n(x_0)$ converge, et que pour tout x il existe V_x voisinage de x tel que la suite des $Df_{n|V_x}$ soit de Cauchy pour la distance d définie par

$$d(f,g) = sup_{z \in V_x} \|f(z) - g(z)\|$$

(c'est-à-dire que la suite des Df_n converge normalement sur un certain voisinage de tout point)

Alors la suite des f_n converge simplement vers une application f différentiable de U dans F telle que tout x de U possède un voisinage V_x sur lequel les convergences de f_n et Df_n sont uniformes. De plus, si les f_n sont C^1, alors f aussi.

Démonstration *Soit l'ensemble A des z tels que la suite des $f_n(z)$ converge. Étant donné x dans U, on considère un voisinage V_x de x convexe et vérifiant l'hypothèse sur le critère de Cauchy (on peut toujours imposer V_x convexe en le restreignant à une boule).*

On définit alors $\alpha_{n,m} = sup_{z \in V_x} \|Df_n(z) - Df_m(z)\|$; par hypothèse, $\alpha_{n,m}$ tend vers 0 quand n et m tendent vers l'infini. Si $V_x \cap A \neq \emptyset$, alors soit x' dans $V_x \cap A$. Par convexité de V_x, en utilisant le corollaire 0.14, on peut écrire pour tout y dans V_x :

$$\|[f_m(y) - f_n(y)] - [f_m(x') - f_n(x')]\| \leq \alpha_{n,m}.\|y - x'\|$$

or $f_n(x')$ est une suite de Cauchy (comme toute suite convergente dans un métrique), donc $f_n(y)$ est une suite de Cauchy, et donc converge. Donc si $V_x \cap A \neq \emptyset$, $V_x \subset A$. Donc tout point de A possède un voisinage V_x contenu soit dans A soit dans A^c, ce qui montre que A et A^c sont ouverts. A étant non vide et U étant connexe, $A = U$. On a donc montré que la suite des f_n converge simplement.

$\mathcal{L}(E,F)$ est complet pour la norme uniforme puisque F l'est ; donc la suite des dérivées sur V_x converge uniformément. En appliquant le théorème précédent, on voit que f est dérivable de dérivée la limite des dérivées ; en supposant V_x borné on a alors $f_{n|V_x} \to f_{|V_x}$ uniformément, toujours par le théorème précédent.

1.2.3 Applications : dérivées partielles et dérivées

Proposition 0.21

E_1, E_2, ... , E_n et F des espaces vectoriels normés ; U un ouvert de $E = \Pi_i E_i$, f une application de U dans F. Si les $\frac{\delta f}{\delta x_i}$ existent sur un voisinage de x et sont continues en x, alors f est différentiable en x.

Démonstration

•Il est suffisant de montrer que

$$\|f(x_1, ..., x_n) - f(a_1, ..., a_n) - \sum_{i=1}^{n} \frac{\delta f}{\delta x_i}(a).(x_i - a)\| = o(\|x - a\|)$$

pour tout $(a_1, ..., a_n)$ de U.

•Pour cela on décompose $f(x_1,...,x_n)-f(a_1,...,a_n)-\sum_{i=1}^n \frac{\delta f}{\delta x_i}(a).(x_i-a)$ *en*

$$\begin{aligned}
& f(x_1,...,x_n)-f(a_1,x_2,...,x_n)-\frac{\delta f}{\delta x_1}(a).(x_1-a_1) \\
+ \quad & f(a_1,x_2,...,x_n)-f(a_1,a_2,x_3,...,x_n)-\frac{\delta f}{\delta x_2}(a).(x_2-a_2) \\
+ \quad & ... \\
+ \quad & f(a_1,a_2,...,a_i,x_{i+1},...,x_n)-f(a_1,...,a_{i+1},x_{i+2},...,x_n) \\
& -\frac{\delta f}{\delta x_{i+1}}(a).(x_{i+1}-a_{i+1}) \\
+ \quad & ...+ \quad f(a_1,a_2,...,a_{n-1},x_n)-f(a_1,...,a_n)-\frac{\delta f}{\delta x_n}(a).(x_n-a_n)
\end{aligned}$$

•Il suffit ensuite de montrer que pour x_i *tendant vers* a_i, $f(a_1,a_2,...,a_{i-1},x_i,...,x_n)-f(a_1,...,a_i,x_{i+1},...,x_n)-\frac{\delta f}{\delta x_i}(a).(x_i-a_i)$ *est un* $o(x_i-a_i)$. *(ensuite il suffira de sommer)*

•Le théorème des accroissements finis donne un $o(x_i-a_i)$ *pour* $f(a_1,...,a_{i-1},x_i,...,x_n)-f(a_1,...,a_i,x_{i+1},...,x_n)-\frac{\delta f}{\delta x_i}(a_1,...,a_i,x_{i+1},...,x_n).(x_i-a_i)$ *et la continuité des dérivées partielles montre que* $\frac{\delta f}{\delta x_i}(a_1,...,a_i,x_{i+1},...,x_n).(x_i-a_i)$ *est proche de* $\frac{\delta f}{\delta x_i}(a).(x_i-a_i)$.

Théorème 0.22

E_1, E_2, ... , E_n et F des espaces vectoriels normés ; U un ouvert de $E=\Pi E_i$, alors une application f de U dans F est C^1 si et seulement si les dérivées partielles $\frac{\delta f}{\delta x_i}$ existent et sont continues sur U.

Démonstration *Il est clair que si* f *est* C^1, *alors les dérivées partielles existent et sont continues. La réciproque, utilisant la proposition précédente, ne présente pas de difficulté majeure.*

Application 0.4 On pourra par exemple trouver une application dans la partie **??**.

1.3 Théorème d'inversion locale et fonctions implicites

On va voir ici d'importants théorème : le théorèmes d'inversion globale, d'inversion locale, de fonctions implicites. Ces théorèmes sont notamment fondamentaux pour

— calculer des différentielles d'inverse, ce qui peut être très pratique pour mettre à jour des solutions de problèmes inverses (théorème d'inversion locale).

— établir qu'un C^1-difféomorphisme de classe C^n est en fait un C^n-difféomorphisme (conséquence du théorème d'inversion globale).

— établir que toute application d'un espace euclidien dans lui-même dont la différentielle existe, est continue et est un endormophisme orthogonal en chaque point, est une isométrie affine (ce qui permettra d'étudier des propriétés métriques d'un sous-ensemble d'un espace euclidien en se ramenant à son image ou son image inverse par une telle application).

1.3.1 Théorème d'inversion globale

Le théorème d'inversion globale est notamment important pour construire le théorème d'inversion locale, mais il a aussi des corollaires propres distincts comme on va le voir.

⚠ *Attention 0.5* Attention, ne surtout pas dire que le théorème d'inversion globale est une généralisation ou une globalisation du théorème d'inversion locale ! Le théorème d'inversion globale concerne des applications linéaires, alors que le théorème locale concerne des fonctions C^1 !

DÉFINITION 0.8 **application contractante**
On appelle **application contractante** ou **contraction** une application lipschitzienne ayant un coefficient de Lipschitz < 1.

THÉORÈME 0.23 **Théorème de Banach du point fixe**

Soit X un espace métrique complet et h une contraction de X dans X. Alors :
•h admet un unique point fixe x_0
•$\forall x \; d(x, x_0) \leq \frac{1}{1-Lip(h)} d(x, h(x))$

Démonstration *Unicité :*
•*Si x_1 et x_2 sont deux points fixes de f, on a alors :*
$d(x_1, x_2) = d(h(x_1), h(x_2)) \leq Lip(h).d(x_1, x_2)$, donc $x_1 = x_2$
Existence :
•*Considérons x quelconque dans X, on va travailler sur la suite des $h^n(x)$.*
•*Supposons $n \leq m$, alors*

$$d(h^m(x), h^n(x)) \leq \sum_{i=m}^{n-1} d(h^i(x), h^{i+1}(x))$$

$$\leq \sum_{i=m}^{+\infty} Lip(h)^i.d(x, h(x)) \leq \frac{Lip(h)^m}{1 - Lip(h)}.d(x, h(x))$$

On en déduit facilement les deux résultats annoncés.

⚠ *Attention 0.6* Il faut que f soit une contraction, c'est-à-dire une application lischitzienne de constante de Lipschitz < 1 ; avec un rapport 1 cela ne marche pas, ni même avec $d(f(x), f(y)) < d(x, y)$. Par exemple, $f : x \mapsto x + e^{-x}$ définit une application de $\mathbb{R}^+$ dans $\mathbb{R}^+$, $\mathbb{R}^+$ est bien complet, et on a bien $d(f(x), f(y)) < d(x, y)$, et pourtant f n'admet pas de point fixe.

Application 0.7 Applications : théorème d'inversion locale 0.26, théorème de Cauchy-Lipschitz **??**, la résolution de l'équation de Volterra (voir [2]).

D'autres théorèmes de points fixes existent : par exemple le théorème du point fixe de Brouwer **??** (avec pour application le corollaire **??** quant au champ rentrant dans la sphère), le théorème de Kakutani, le théorème de Schauder, et même, pour ceux qui connaissent un peu la calculabilité, un théorème de point fixe que l'on trouvera dans le livre « Théorie de la récursion pour la métamathématique », de R. Smullyan (Masson, 1995), avec pour application le théorème de Rice et ses multiples conséquences (attention, il faut connaître un peu le domaine pour pouvoir se lancer dans ce genre d'originalités...).

Lemme 0.24
Soient U et V des ouverts des espaces normés E et F. On se donne h de U dans V bijective, dérivable en x_0. Alors h^{-1} est dérivable en $h(x_0)$ si et seulement si les deux conditions suivantes sont vérifiées :

- $Dh(x_0)$ est un isomorphisme de E sur F ;
- Il existe $K \geq 0$ et un voisinage W de $h(x_0)$ dans F tels que

$$\forall y \in W, \ \|h^{-1}(y) - x_0\| \leq K\|y - h(x_0)\|.$$

Démonstration *Tout d'abord montrons que ces deux conditions sont nécessaires. Pour cela on suppose qu'effectivement* h^{-1} *est dérivable en* $h(x_0)$, *et on procède comme suit :*

- *On dérive les deux expressions*

$$h^{-1} \circ h = Id_E \text{ et } h \circ h^{-1} = Id_E$$

et on montre bien que $Dh(x_0)$ *est un isomorphisme.*

- *Par définition de la dérivée, la quantité ci-dessous tend vers* 0 *pour* $y \to h(x_0)$ *:*

$$\frac{\|h^{-1}(y) - h^{-1}(h(x_0)) - D(h^{-1})(h(x_0))(y - h(x_0))\|}{\|y - h(x_0)\|}$$

Donc pour y *dans un certain* W *cette quantité est plus petite que* 1, *et donc pour* $y \in W$ *on a*

$$\|h^{-1}(y) - h^{-1}(h(x_0))\| \leq K\|y - h(x_0)\|$$

avec $K = 1 + \|D(h^{-1})(h(x_0))\|$.

Il reste à prouver la réciproque, c'est-à-dire que les conditions sont suffisantes.

- *Par définition, on a*

$$h(x) - h(x_0) = Dh(x_0)(x - x_0) + \|x - x_0\|\epsilon(x)$$

avec $\epsilon(x)$ *tendant vers* 0 *pour* $x \to x_0$.

En composant avec $Dh(x_0)^{-1}$ *(dont les hypothèses garantissent l'existence), on obtient*

$$x - x_0 = Dh(x_0)^{-1}(h(x) - h(x_0)) - \|x - x_0\|.(Dh(x_0)^{-1}(\epsilon(x)))$$

avec $y = h(x)$ *(tout* y *de* V *peut s'écrire ainsi) et* $y_0 = h(x_0)$, *on obtient alors*

$$h^{-1}(y) - h^{-1}(y_0) - Dh(x_0)^{-1}(y - y_0) = -Dh(x_0)^{-1}.\epsilon(h^{-1}(y)).\|h^{-1}(y) - h^{-1}(y_0)\|$$

On sait par hypothèse que pour y *assez proche de* y_0 *on a* $\|h^{-1}(y) - h^{-1}(y_0)\| \leq K.\|y - y_0\|$, *donc*

$$\frac{Dh(x_0)^{-1}.\epsilon(h^{-1}(y)).\|h^{-1}(y) - h^{-1}(y_0)\|}{\|y - y_0\|} \leq K.\|Dh(x_0)^{-1}\|.\|\epsilon(h^{-1}(y))\|$$

or $h^{-1}(y)$ *tend vers* x_0 *quand* $y \to y_0$ *donc cette quantité tend vers* 0 *; ce qui permet de conclure.*

Théorème 0.25 **Théorème d'inversion globale**

Soit A une application linéaire continue de E dans F, avec E un espace de Banach et F un espace normé, telle que A^{-1} existe et est continue (A est un homéomorphisme linéaire). Soit ϕ une application lipschitzienne de E dans F telle que $Lip(\phi) < \|A^{-1}\|^{-1}$. Alors :

•$h = A + \phi$ est inversible.

•h^{-1} est lipschitzienne, avec $Lip(h^{-1}) \leq \frac{\|A^{-1}\|}{[1-\|A^{-1}\|.Lip(\phi)]}$

•Si h est C^1 sur U ouvert de E, et si $\forall x \in U,\ Dh(x) \in Isom(E,F)$, alors h^{-1} est C^1 sur l'ouvert $h(U)$, et la différentielle de h^{-1} est donnée par

$$\forall x \in U,\ D(h^{-1})(h(x)) = (Dh(x))^{-1}.$$

Démonstration *Raisonnement en plusieurs étapes :*

•Étant donné y dans F, et on considère l'équation

$$h(x) = y$$

équivalente à

$$x = A^{-1}.y - A^{-1}(\phi(x))$$

L'application $x \mapsto A^{-1}(\phi(x))$ est Lipschitzienne de rapport < 1, donc l'application $x \mapsto A^{-1}(y) - A^{-1}(\phi(x))$ aussi. Donc par le théorème du point fixe de Banach, cette équation a une solution unique x.

h est donc inversible.

•Avec $y = h(x)$ et $y' = h(x')$, on peut écrire

$$x = A^{-1}(y) - A^{-1}(\phi(x))$$

$$x' = A^{-1}(y') - A^{-1}(\phi(x'))$$

et en déduire

$$\|x - x'\| \leq \|A^{-1}\|.\|y - y'\| + \|A^{-1}\|.Lip(\phi).\|x - x'\|$$

en utilisant l'hypothèse sur $Lip(\phi)$, on a alors

$$\|x - x'\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\|.Lip(\phi)}\|y - y'\|$$

D'où le résultat sur la constante de Lipschitz de h^{-1}.

•Supposons maintenant h C^1 sur U ouvert de E. $h(U)$ est un ouvert puisque h est un homéomorphisme. Par le lemme précédent, h^{-1} est dérivable sur U. En dérivant

$$h^{-1} \circ h = Id_E \text{ et } h \circ h^{-1} = Id_E$$

on obtient que $D(h^{-1})(h(x)) = (Dh(x))^{-1}$. En écrivant

$$D(h^{-1})(y) = Dh(h^{-1}(y))^{-1}$$

et en rappelant que $Inv : Isom(E,F) \to Isom(F,E), h \mapsto h^{-1}$ est continue on constate qu'en outre $D(h^{-1})$ est continue (voir la remarque qui suit le corollaire 0.27 pour plus d'informations sur Inv).

1.3.2 Théorème d'inversion locale

Le théorème d'inversion locale est important en soi (il permet de calculer des différentielles d'inverses) mais aussi en tant qu'outil pour construire le théorème des fonctions implicites.

DÉFINITION 0.9 **Difféomorphisme C^1**

Une application h de U dans V avec U ouvert d'un espace vectoriel normé et V ouvert d'un espace vectoriel normé est un **difféomorphisme** C^1 si h est bijective et de classe C^1 et de réciproque de classe C^1. Plus généralement, avec $k \geq 1$, une application h de U dans V avec U ouvert d'un espace vectoriel normé et V ouvert d'un espace vectoriel normé est un **difféomorphisme** C^k si h est bijective et de classe C^k et de réciproque de classe C^k.

Attention 0.8 Une application bijective et C^1 n'est pas un difféomorphisme C^1 ; il faut aussi que la réciproque soit C^1 !

Démonstration *Facile en distinguant trois cas (x dans la boule et pas y, x et y dans la boule, x et y hors de la boule de rayon r).*

Remarque Dans le cas d'une norme induite par un produit scalaire le coefficient de Lipschitz est en fait 1 ; et que le cas d'un coefficient de Lipschitz égal à 2 s'obtient notamment dans le cas de la métrique sur $\mathbb{R}^2$ définie par la norme $\|(x_1, x_2)\| = max(|x_1|, |x_2|)$ (considérer par exemple $(r + \epsilon/2, r + \epsilon/2)$ et $(r, r + \epsilon)$).

THÉORÈME 0.26 **Théorème d'inversion locale**

Soit h de U dans F une application C^1, avec U ouvert de E, et E et F des espaces de Banach. Si la différentielle $Dh(x_0)$ est bijective de E dans F pour un certain x_0 de U, alors il existe U_0 voisinage de x_0 dans E et un voisinage ouvert V_0 de $f(x_0)$ dans F tels que h induit un difféomorphisme C^1 de U_0 dans V_0. On a alors

$$D(h^{-1})(h(x)) = (Dh(x))^{-1} \text{ pour tout } x \text{ dans } U_0.$$

Application 0.9 Voir le théorème 0.29 (fonctions implicites) et le corollaire 0.27 (ci-dessous). Le théorème d'inversion locale permettra aussi de montrer l'équivalence des différentes définitions des variétés de $\mathbb{R}^n$, voir définition 0.13.

Démonstration *On pourra se référer par exemple à [3].*

Attention 0.10 Dans le cas de la dimension finie, le fait que la différentielle soit bijective implique que les dimensions des espaces soient les mêmes, et que le jacobien (défini puisque les dimensions sont les mêmes) est non nul.

COROLLAIRE 0.27

Soit f une application C^1 d'un ouvert U d'un espace de Banach E dans un espace de Banach F ; si f injective et si sa différentielle $f'(x)$ est un isomorphisme en tout $x \in U$, alors f est un C^1-difféomorphisme de U sur $f(U)$, qui est alors un ouvert de F.

Une extension est possible : modulo le résultat selon lequel l'application Inv qui à f dans $Isom(E,F)$ associe f^{-1} est C^∞ (que l'on trouvera par exemple dans [1]), on montre que si f est un C^1-difféomorphisme de classe C^n, alors f est un C^n-difféomorphisme.

Corollaire 0.28
Soit f une application de classe C^1 d'un espace euclidien E dans lui-même. On suppose que pour tout x, la différentielle $Df(x)$ est un endomorphisme orthogonal. Alors f est une isométrie affine.

Démonstration *Notons d'abord que f est 1-lipschitzienne. En effet, si a et b sont deux points, on a :*

$$\begin{aligned} & \|f(a)-f(b)\| \\ = \quad & \|\int_0^1 Df((1-t)a+tb).(a-b)dt\| \\ \leq \quad & sup_t\|Df((1-t)a+tb)\|.\|a-b\| = \quad \|a-b\|. \end{aligned}$$

Soit x un point de E. Comme $Df(x)$ est inversible, d'après le théorème d'inversion locale, il existe une boule B centrée en x telle que f induit un C^1-difféomorphisme de B sur $f(B)$. Pour tout y dans $f(B)$, la différentielle $D(f^{-1})(y) = (Df(f^{-1}(y)))^{-1}$ est un endomorphisme orthogonal, si bien que la réciproque de f de $f(B)$ vers B est également 1-lipschitzienne. Par double inégalité, la restriction de f à B est une isométrie.

En particulier, ceci montre que Df est localement constante, donc constante. Par intégration, f est une isométrie affine.

1.3.3 Théorème des fonctions implicites

Théorème 0.29

f application C^1 de U un ouvert de $E_1 \times E_2$ dans F, avec E_1, E_2 et F des espaces de Banach. Si f est différentiable par rapport à la deuxième variable en (a,b) et si cette différentielle est un isomorphisme alors il existe ϕ C^1 définie sur un ouvert U_1 de E_1 contenant a_1 et à valeurs dans un ouvert U_2 de E_2 contenant a_2 telle que pour tout (y_1,y_2) dans $U_1 \times U_2$ on ait

$$f(y_1,y_2) = f(a,b) \iff y_2 = \phi(y_1)$$

et pour tout x dans U_1 on a

$$D\phi(x) = -[D_2f(x,\phi(x))]^{-1} \circ D_1f(x,\phi(x))$$

Application 0.11 Voir par exemple la partie 1.5.5 sur les surfaces de $\mathbb{R}^3$ (montrant par exemple que les quadriques ou de manière générale les solutions d'équations dont la différentielle a certaines propriétés sont des surfaces régulières).

Démonstration

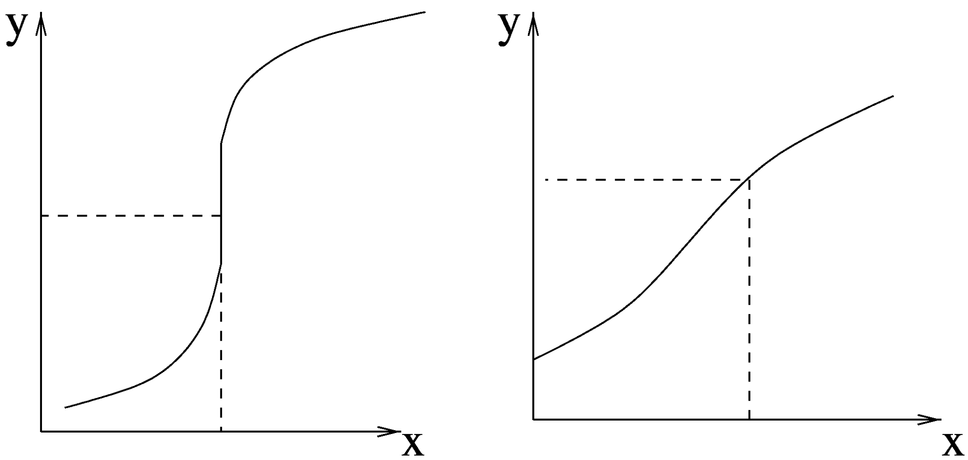

FIGURE 1 –
Commentaire : Illustration du théorème des fonctions implicites dans le cas d'une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$. La courbe est une courbe de niveau. A gauche la différentielle par rapport à y n'est pas un isomorphisme ; on comprend intuitivement qu'on ne peut pas donner y en fonction de x sur un voisinage. A droite c'est un isomorphisme ; donc on peut.

•*On va chercher à utiliser le théorème d'inversion locale. Pour cela il faut construire une application différentiable* C^1, *de différentielle bijective.*
•*On définit l'application* μ *de* U *dans* $E_1 \times E_2$ *dans* $E_1 \times G$, *définie par* $\mu(y_1, y_2) = (y_1, f(y_1, y_2))$.
•$D\mu(a,b)(h,k) = (h, D_1f(a,b).h + D_2f(a,b).k)$
•$D\mu(a,b)$ *est bijective (facile),* μ *est* C^1
•*On applique le théorème d'inversion locale à* μ *en* (a,b).
•*La fonction réciproque associe clairement à un couple* (q,r) *un couple* $(q, \phi(q,r))$, *et en fixant* r *on en déduit ce que l'on veut.*

La figure 1 illustre ce théorème.

1.4 Dérivées d'ordre supérieur

Les dérivées d'ordre élevé permettent de nombreuses choses ; par exemple, les algorithmes d'optimisation basés sur la hessienne ont des ordres de convergence largement supérieurs aux méthodes basées sur le seul gradient. De même, les schémas d'intégration numérique sont bien plus rapides théoriquement en ayant un ordre élevé (i.e. en utilisant une dérivée 3ème, 4ème). Il faut bien noter toutefois qu'une dérivée troisième, quatrième, cinquième, est bien instable et difficile à calculer. Ainsi, sur les machines de précision finie, on n'utilisera jamais, pour l'optimisation, une dérivée 3ème car elle est trop chère à calculer, et en intégration numérique (où la dérivée k-ième est utilisée pour démontrer les convergences mais n'est pas explicitement calculée) on s'arrête en général à 4 car le calcul devient instable ensuite. Ce problème de dérivée inutilisable ou trop chère est tel que de nombreuses méthodes d'optimisation se passent désormais de dérivées (méthodes évolutionnaires, méthodes de recuit simulé), et de même des méthodes de calcul d'intégrale, notamment quand la dimension augmente, se contentent de moyenner des valeurs ayant toutes exactement le même poids (i.e. $\int f$ est approximé par $\frac{1}{n}\sum_{i=1}^{n} f(x_i)$, typiquement avec des méthodes de Monte-Carlo ou Quasi-Monte-Carlo, au lieu de $\sum_{i=1}^{n} w_i f(x_i)$ avec $\sum w_i = 1$ et w_i parfois négatif lorsqu'on utilise des schémas numériques d'intégration basés sur les développements de Taylor).

1.4.1 Généralités

Définition 0.10 **de classe C^n**
Étant donnée une application f d'un ouvert U d'un espace de Banach E dans un espace de Banach F, on dit que f est **de classe C^n** (on dit aussi n **fois continûment différentiable**) si f est différentiable et si sa différentielle est de classe C^{n-1}. L'application est dite C^∞ si elle est C^n pour tout n ; on dit alors qu'elle est **indéfiniment différentiable**.

On note alors $f^{(1)}(a)$ l'application $Df(a)(x)$, et par récurrence $f^{(n)}(a)$ l'application $Df^{(n-1)}(a)$.

Étant donné une application f d'un ouvert U d'un espace de Banach E dans un espace de Banach F, on dit que f est n fois différentiable en x appartenant à U si et seulement si f est de classe C^{n-1} sur un voisinage de x et si la n-ième différentielle de f sur ce voisinage est différentiable en a.

Étant donné une application f d'un ouvert U d'un espace de Banach E dans un espace de Banach F, on note $\frac{\delta^2 f}{\delta x_i \delta x_j}$ l'application $\frac{\delta \frac{\delta f}{\delta x_j}}{\delta x_i}$.

Notez l'ordre dans la définition : on a besoin de différentielle (n fois) sur un ouvert pour définir la différentielle en un point $n+1$ fois.

1.4.2 Dérivées secondes

Théorème 0.30

Soit f une application de U, ouvert d'un espace de Banach , dans F, espace de Banach . Si f est deux fois différentiable en x_0, alors $f''(a)(h)(k) = f''(a)(k)(h)$. Via l'identification du théorème **??** $f''(a)$ est une forme bilinéaire symétrique.

Démonstration

$\bullet$*Il s'agit donc de montrer que* $f''(a)(h)(k) = f''(a)(k)(h)$.

$\bullet$*On introduit la fonction* $\mu(h,k) = f(a+h+k) - f(a+h) - f(a+k) + f(a)$

$\bullet\mu(h,k) = \mu(k,h)$ *clairement*

$\bullet$*On montre maintenant que* $\mu(h,k)$ *approche* $f''(a)(h)(k)$ *(on pourra plus loin en déduire le résultat souhaité, car* $f''(a)(h)(k)$ *approchera alors* $f''(a)(k)(h)$*)*

$\bullet\|\mu(h,k)-f''(a)(h)(k)\| \leq \|\mu(h,k)-f'(a+k)(h)+f'(a)(h)\|+\|f'(a+k)(h)-f'(a)(h)-(f''(a).k).h\|$

$\bullet$*Le second terme est petit par la définition de la différentielle* $f''(a)$*, le premier est petit par définition de la différentielle de* $f'(a)$*, et par utilisation des accroissements finis.*

$\bullet$*On arrive ainsi à montrer que* $\mu(h,k) - f''(a)(k)(h)$ *est un* $o(\|h\|^2 + \|k\|^2)$ *; par symétrie on a aussi le fait que* $\mu(h,k) - f''(a)(k)(h)$ *est un* $o(\|h\|^2 + \|k\|^2)$ *; on en déduit* $\|f''(a)(k)(h) - f''(a)(h)(k)\| = o(\|h\|^2 + \|k\|^2)$.

$\bullet\|f''(a)(\lambda.k)(\lambda.h) - f''(a)(\lambda.h)(\lambda.k)\| \leq \epsilon.(\|\lambda.h\|^2 + \|\lambda.k\|^2)$ *pour tout* ϵ*, pour* λ *suffisamment petit*

$\bullet\lambda^2\|f''(a)(k)(h) - f''(a)(h)(k)\| \leq \epsilon.\lambda^2(\|h\|^2 + \|k\|^2)$ *pour tout* ϵ *et pour* λ *assez petit*

$\bullet\|f''(a)(k)(h) - f''(a)(h)(k)\| \leq \epsilon(\|h\|^2 + \|k\|^2)$ *pour tout* ϵ *!*

$\bullet f''(a)(k)(h) = f''(a)(h)(k) = 0$ *d'où le résultat.*

THÉORÈME 0.31
Soit f deux fois différentiables en $a \in U$, avec f définie de U ouvert de $E = E_1 \times E_2 ... \times E_n$ (des espaces de Banach) dans F (un espace de Banach). Alors

$$(f''(a)(h_1, ..., h_n))(k_1, ..., k_n) = \sum_{i,j} \left(\frac{\delta^2 f}{\delta x_i \delta x_j}(a).h_i \right).k_j$$

Démonstration *On applique simplement deux fois la proposition 0.8, à* f *et* f'.

COROLLAIRE 0.32
Soit f deux fois différentiables en $a \in U$, avec f définie de U ouvert de $E = E_1 \times E_2 ... \times E_n$ (des espaces de Banach) dans F (un espace de Banach). Alors

$$\frac{\delta^2 f}{\delta x_i \delta x_j}(a)(h)(k) = \frac{\delta^2 f}{\delta x_j \delta x_i}(a)(k)(h)$$

Démonstration *Il s'agit simplement du théorème 0.30 après quelques manipulations...*

COROLLAIRE 0.33 **Théorème de Schwartz**

Soit f une application de $\mathbb{R}^n$ dans un espace de Banach F deux fois différentiable en x, alors $\frac{\delta^2 f}{\delta x_i \delta x_j}(x) = \frac{d^2 f}{\delta x_j \delta x_i}(x)$.

Démonstration *Il s'agit d'une reformulation dans le cas de* $E_i = \mathbb{R}$ *du corollaire précédent !*

Proposition 0.34 **Existence de la dérivée seconde**

Soi f une application de U ouvert de $E = E_1 \times E_2 ... \times E_n$ (des espaces de Banach) dans F (un espace de Banach). Alors si les $\frac{\delta f}{\delta x_i}$ existent et sont continues sur un voisinage de x, et si les $\frac{\delta^2 f}{\delta x_i \delta x_j}$ existent sur un voisinage de x et sont continues en x, alors f est deux fois différentiable en x.

Démonstration *Il suffit d'appliquer deux fois la proposition 0.21.*

1.4.3 Généralisations à la dérivée n-ième

On pourra si nécessaire réviser le chapitre **??** (algèbre multilinéaire) avant de plonger ici. Dans les applications informatiques en optimisation non-linéaire, souvent on s'arrête à la dérivée première, la dérivée seconde (la hessienne) étant approchée à partir de valeurs successives de la dérivée première (le gradient). Mais même sans calcul explicite des dérivées n-ième, on utilise parfois implicitement ces dérivées, par exemple pour estimer les erreurs de schémas d'intégrations numériques (par exemple, dérivée 6-ième pour la méthode dite de Boole-Villarceau, 4-ième pour Newton-Cotes).

Théorème 0.35 **Généralisation du théorème 0.30**

Si f est une application n fois différentiable de E dans F, où E et F sont des espaces de Banach , alors pour tout x de E, $f^{(n)}(x)$ est une application n-linéaire symétrique sur E à valeurs dans F.

Démonstration *On procède par récurrence. Pour $n = 1$, c'est clair. Pour $n = 2$, c'est le théorème 0.30. Supposons maintenant le résultat prouvé jusqu'au rang $n - 1$, et montrons le pour le rang n, avec $n \geq 3$.*

•Il suffit de montrer que si l'on permute deux variables consécutives parmi les h_i on ne change pas la valeur $f^{(n)}(x)(h_1, ..., h_n)$.

•$f^{(n-1)}$ est symétrique, donc on peut permuter sans rien changer h_i et h_{i+1} pour $i > 1$

•Pour $i = 1$ il suffit de rappeler que $f^{(n)} = (f^{(n-2)})''$ et d'utiliser 0.30.

1.5 Zoologie du calcul différentiel

On présentera ici la convexité (fondamentale pour l'optimisation numérique et pour beaucoup d'autres choses), une fonction pathologique continue partout et dérivable nulle part, une autre fonction pathologique dérivable dans toutes les directions mais non continue, avant de passer à l'étude de variétés et de surfaces régulières.

1.5.1 Fonctions convexes

Définition 0.11 **convexe**

Une fonction f définie sur un convexe U d'un espace vectoriel à valeurs dans $\mathbb{R}$ est dite **convexe** (resp. **strictement convexe**) si

$$\forall (u, v, t) \in U^2 \times [0, 1] f(tu + (1 - t)v) \leq tf(u) + (1 - t)f(v)$$

$$(\text{resp.})\forall(u,v,t) \in U^2\times]0,1[u \neq v \Rightarrow f(tu+(1-t)v) < tf(u)+(1-t)f(v)$$

Dans la suite de cette section, on suppose que U est un convexe d'un espace vectoriel et que f est une application de Ω dans $\mathbb{R}$, avec Ω ouvert contenant U. Les liens entre dérivabilité et convexité sont les suivants :

THÉORÈME 0.36
Soit f une fonction C^1. Alors f est convexe (resp. strictement convexe) sur U si et seulement si

$$\forall(u,v) \in U^2,\ f(v) \geq f(u)+f'(u)(v-u)$$

$$(\text{resp.})\forall(u,v) \in U^2 u \neq v \Rightarrow f(v) > f(u)+f'(u)(v-u)$$

f C^2 est convexe si et seulement si

$$\forall(u,v) \in U^2,\ f''(u)(v-u,v-u) \geq 0$$

Enfin, une implication dans un seul sens :

$$\forall(u,v) \in U^2,\ u \neq v \Rightarrow f''(u)(v-u,v-u) > 0$$

entraîne que f est strictement convexe.

Intuition Ce résultat généralise le cas standard d'une dérivée seconde positive comme critère de convexité pour $U =]a,b[\subset \mathbb{R}$. $f''(u)$ est simplement de la forme $t \mapsto at^2$, avec a la dérivée seconde.

Intuition L'hypothèse de convexité de U est nécessaire. Un exemple de fonction non convexe ayant sa différentielle seconde strictement positive est : $f(x) = x^2$ si $x \in [-1,1]$ et $f(x) = (x-3)^2$ si $x \in [2,4]$.

Application 0.12 On pourra voir **??** pour les applications de la convexité à la recherche d'extréma, **??** pour les applications de l'inégalité de Jensen, le lemme **??** (et par suite l'inégalité de Hölder), l'inégalité de Minkovski.

1.5.2 Fonction continue partout dérivable nulle part

Cet exemple, élaboré par M. Van der Waerden, est extrait du livre [4] auquel on se réfèrera pour plus de précisions que la brève présentation qui est faite ici.

THÉORÈME 0.37

Soit T la fonction définie sur $\mathbb{R}$ par $T(x) = min(x-E(x), E(x)+1-x)$ (c'est-à-dire que $T(x)$ est la distance de x à l'entier le plus proche de x).

La fonction f définie sur $\mathbb{R}$ par $f(x) = \sum_{n=0}^{\infty} \frac{T(10^n x)}{10^n}$ est continue partout et dérivable nulle part.

Démonstration

•Il convient tout d'abord de vérifier que f est bien définie et continue : pour cela, il suffit de voir que f est limite uniforme d'une suite de fonctions continues (voir proposition 0.18).

•La non-dérivabilité est plus dure à établir :

- *T, et donc f, est périodique, de période 1.*
- *On se limite donc à montrer la non-dérivabilité sur $[0,1[$*
- *On note les développements décimaux en excluant les développements illimités ne comportant que des 9 à partir d'un certain rang*[1]
- *Soit donc $x \in [0,1[$, on montre la non-dérivabilité de f en x.*
- *Soit x_n la n-ième décimale de x.*
- *définissons $h_m = -10^{-m}$ si $x_m = 4$ ou $x_m = 9$, $h_m = 10^{-m}$ sinon.*
- *Calculons maintenant*

$$\Delta_m = \frac{f(x+h_m) - f(x)}{h_m}$$

$$\Delta_m = \sum_{n=1}^{\infty} \pm 10^{m-n} (T(10^n x \pm 10^{n-m}) - T(10^n x)).$$

Pour $n \geq m$, $T(10^n x \pm 10^{n-m}) - T(10^n x)$ est nul, car $10^n x \pm 10^{n-m}$ et 10^{nx} ont la même partie fractionnaire.

Pour $n < m$, l'entier le plus proche est le même pour $10^n x \pm 10^{n-m}$ et $10^n x$ car le choix de h_m dans le cas où x_m vaut 4 ou 9 empêche de franchir un entier ou un demi-entier en passant de l'un à l'autre. On en déduit : $\pm 10^{m-n}(T(10^n x \pm 10^{n-m}) - T(10^n x))$ vaut 1 ou -1. Ainsi, Δ_m est une somme de m nombres valant 1 ou -1, ce qui est un entier de même parité que m. La suite des Δ_m n'a donc pas de limite.

On en profite pour montrer ce dont est capable Maple. Le dessin se trouve en figure 2.

Exemple Maple

$> T := x \to \min(x - \text{floor}(x), \text{floor}(x) + 1 - x)$

$T := x \to \min(x - \text{floor}(x), \text{floor}(x) + 1 - x)$

$> g := x- > sum(T(2^n * x)/2^n, n = 0..17);$

$$g := x \to \sum_{n=0}^{17} \frac{\mathrm{T}(2^n\, x)}{2^n}$$

1.5.3 Fonction dérivable dans toutes les directions mais non continue

DÉFINITION 0.12 **différentiable en $x \in U$ suivant la direction $u \in E$**
Soit f une application d'un ouvert U d'un espace vectoriel normé E dans un espace vectoriel normé F, alors f est dite **différentiable en $x \in U$ suivant la direction $u \in E$** si l'application $g : \mathbb{R} \to F$ $t \mapsto f(x+tu)$ est différentiable en 0. La différentielle de g en 0 est alors appelée **différentielle de f en x suivant u.**

1. Au profit de l'équivalent obtenu en remplaçant ...243999999999... par ...2440000....

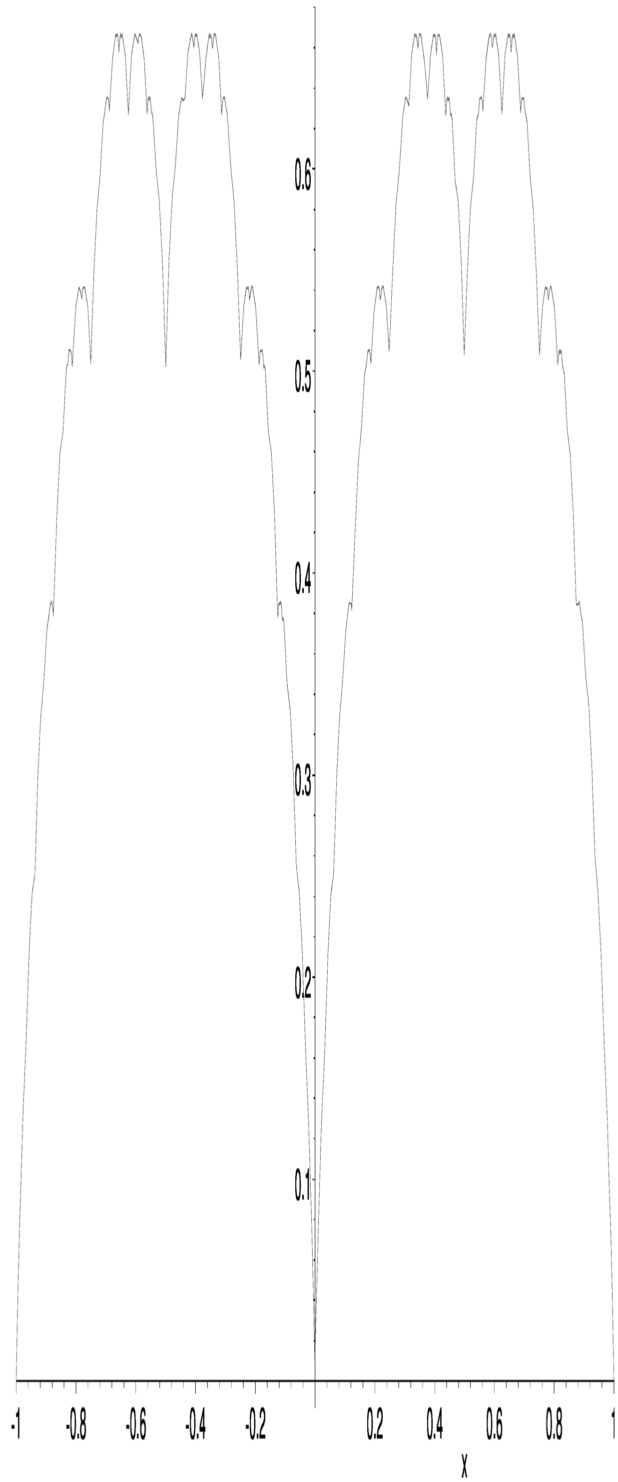

Figure 2 – Tracé d'une courbe continue dérivable nulle part, établie par Van der Waerden.

PROPOSITION 0.38
Il existe une application f différentiable dans toutes les directions en x et qui n'est pas continue en x.

Démonstration *[4] propose la fonction $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto \frac{x^5}{(y-x^2)^2+x^8}$, avec $f(0,0) = 0$. On constate la non continuité de f en regardant la limite de $x \mapsto f(x, x^2)$ en 0. La différentiabilité suivant toutes les directions est vite vue (distinguer différents cas, suivant $u = (a, b)$, cas a et b non nuls, cas a seulement nul, ou cas b seulement nul).*

On peut aussi regarder la fonction f définie par ses coordonnées polaires par

$$f(r\cos(\theta), r\sin(\theta)) = e^{-\frac{(|\theta|-r)^2}{r^4}}, \quad (r > 0, \theta \in [-\pi, \pi])$$

Avec $f(0,0) = 0$, on a bien une fonction différentiable dans toutes les directions (avec des différentielles nulles !), et on constate sans le moindre calcul que $f(x\cos(x), x\sin(x))$ est constant égal à 1 pour $x > 0$.

1.5.4 Sous-variétés de $\mathbb{R}^n$

DÉFINITION - PROPOSITION 0.13 **1**
Sous-variétés de $\mathbb{R}^n$

Soit M une partie de $\mathbb{R}^n$, x un point de M. Soit p un entier > 0 et k un entier > 0.

M est par définition une **variété de dimension p et de classe C^k au voisinage de x** si l'une des conditions équivalentes suivantes est vérifiée :

(i) Il existe V voisinage ouvert de x tel qu'il existe un C^k difféomorphisme de V sur $W \subset \mathbb{R}^n$ tel que $g(x) = 0$ et $g(M \cap V) = W \cap P$, avec P l'ensemble des $(y_1, ..., y_n)$ tels que $y_{p+1} = 0$, $y_{p+2} = 0$, ... , $y_n = 0$.

(ii) Après une permutation pertinente des coordonnées $(x_1, ..., x_n)$, il existe V voisinage ouvert de x et ϕ application C^k de $\mathbb{R}^p$ dans $\mathbb{R}^{n-p}$ tel que pour tout y dans V

$$y \in M \iff \phi(y_1, ..., y_p) = (y_{p+1}, ..., y_n)$$

(iii) Il existe un voisinage V ouvert de x, Ω un voisinage ouvert de 0 dans $\mathbb{R}^p$, f une application C^k de Ω dans $\mathbb{R}^n$, tels que f induise un homéomorphisme de Ω sur $M \cap V$, $f(0) = x$ et $f'(0)$ de rang p [2]. Cette application est appelée **représentation paramétrique locale** de la variété.

(iv) Il existe un voisinage V ouvert de x, Ω un voisinage ouvert de 0 dans $\mathbb{R}^p$, f une application C^k de Ω dans $\mathbb{R}^n$, tels que f induise un homéomorphisme de Ω sur $M \cap V$, $f(0) = x$ et $f'(y)$ de rang p pour tout y dans Ω [3].

Une **variété** de dimension p et de classe C^k est une partie qui est une variété de dimension p et de classe C^k au voisinage de chacun de ses points.

Démonstration

On notera pendant cette preuve $x_{|I}$, avec $I = \{i_1, ..., i_m\}$ et $i_1 < i_2 < \cdots < i_m$ un sous-ensemble de $[1, n]$ et x un élément de $\mathbb{R}^n$, $(x_{i_1}, \ldots, x_{i_m})$.

•L'équivalence entre (iii) et (iv) est claire ; bien sûr (iv) implique (iii), et réciproquement en supposant (iii) par continuité de la différentielle et continuité du déterminant d'une matrice extraite, on peut trouver un voisinage de 0 dans lequel la différentielle a le même rang. Il suffit alors de se restreindre à ce voisinage.

•Voyons maintenant que (iii) implique (ii).

Supposons (iii). La matrice de la différentielle de f en 0 est de rang p ; modulo une bonne permutation des coordonnées, on peut donc supposer que la matrice extraite de la différentielle pour les indices en ligne et en colonne inférieurs ou égaux à p est inversible.

En se restreignant aux p première coordonnées, f est alors C^k, de différentielle en 0 de rang plein. On peut donc appliquer le théorème d'inversion locale 0.26, et f ainsi restreint est donc un C^k difféomorphisme au voisinage de 0. En prenant ϕ la composée de f et de l'inverse de la restriction de f aux p-premières coordonnées, on obtient une fonction satisfaisant (ii).

•Voyons maintenant (ii) implique (iii).

Supposons (ii) vérifiée. Définissons alors $f(y) = (y + x_{|[1,p]}, \phi(y + x_{|[1,p]})$[4]. f convient.

•Montrons maintenant que (iii) implique (i).

Supposons (iii) vérifiée.

Définissons alors $g(y) = (y_{|[1,p]} - x_{|[1,p]}, y_{|[p+1,n]} - \phi(y_{|[1,p]})$... g convient pour (i).

•Il ne reste plus qu'à vérifier que (i) implique (iii).

Supposons donc (i) vérifiée.

Alors soit $f(y) = g^{-1}(y_{|[1,p]}, 0, \ldots, 0)$.

La différentielle de g est injective, donc la restriction à $\mathbb{R}^p$ est injective aussi. Donc f vérifie bien (iii).

1.5.5 Surfaces régulières de $\mathbb{R}^3$

En général, on peut travailler dans $\mathbb{R}^n$, mais pour la clarté du propos on se restreint ici aux surfaces de $\mathbb{R}^3$.

DÉFINITION 0.14 **surface régulière de $\mathbb{R}^3$**

On appelle **surface régulière de** $\mathbb{R}^3$ de classe C^k (ou plus simplement **surface** ; ici toutes les surfaces seront régulières) une variété de dimension 2 et de clase C^k dans $\mathbb{R}^3$. On dit qu'une surface S est **réglée** si par tout point de S passe au moins une droite incluse dans S. On dit qu'une surface S est **de révolution** si elle est invariante par rotation autour d'une certaine droite D ; on dit alors que $S \cap P$ est un **méridien** pour tout plan P contenant D ; un cercle situé sur un plan P orthogonal à D et contenu dans S est appelé un **parallèle**.

On appelle **plan tangent** à S surface régulière de classe C^k avec $k \geq 1$ en $x \in S$ le plan affine passant par X et de plan vectoriel directeur engendré par les dérivées partielles $\frac{\delta f}{\delta x}$ et $\frac{\delta f}{\delta y}$ par rapport à x et y d'une représentation paramétrique $(x, y) \mapsto f(x, y)$. Une surface S de classe au moins C^1 est dite **orientable** s'il existe une application n de S dans $\mathbb{R}^3$ telle que $n(s)$ soit de norme 1 et orthogonal à la direction du plan tangent de S en s. n est alors une **orientation** de S, et (S, n) est une **surface orientée**.

On note les points suivants :

— Le plan tangent ne dépend pas de la représentation paramétrique choisie.

— Pour tout arc à valeurs dans la surface, le vecteur vitesse (la dérivée) est à valeurs dans la direction du plan tangent.

— Toute solution d'une équation $f(x, y, z) = 0$ où f est C^k et de différentielle non nulle sur la ligne de niveau $f^{-1}(0)$ est une surface régulière de $\mathbb{R}^3$ (par le théorème des fonctions implicites 0.29).

4. On « recolle » ainsi un élément de $\mathbb{R}^p$ et un élément de $\mathbb{R}^{n-p}$ pour obtenir un élément de $\mathbb{R}^n$.

— La sphère de centre 0 et de rayon 1 n'admet pas de représentation paramétrique globale. Autrement dit, il n'existe pas d'homéomorphisme dont la différentielle est de rang 2 d'un ouvert de $\mathbb{R}^2$ dans $\mathbb{R}^3$, dont l'image soit exactement la sphère. Par contre, la sphère admet une représentation implicite $x^2 + y^2 + z^2 = 1$.

— On parle de surface régulière avec singularité lorsque la surface est une variété de classe C^k sauf en un nombre fini ou dénombrable de points, appelés **singularités**.

DÉFINITION 0.15 **Position par rapport au plan tangent**

On peut toujours, étant donnée une surface régulière S de classe au moins C^1 avec $0 \in S$ et tangente au plan $z = 0$ choisir une représentation paramétrique locale f en 0 telle que pour un certain $(x, y) \mapsto g(x, y)$ on ait $f(x, y) = (x, y, g(x, y))$ (ceci s'étend évidemment en tout point $\neq 0$ de S et pour tout plan tangent - on préfère translater et tourner la surface pour se ramener à ce cas plus « visuel »). Avec Q la forme quadratique qui a mêmes dérivées premières et secondes que g, on a alors nécessairement $Q(x, y) = rx^2 + ty^2 + 2sxy$ pour certains r, t, s (**notation de Monge**). On dit que 0 est :

— Un point **elliptique** si $rt > s^2$, Q est alors soit définie positive soit définie négative. Localement, la surface est entièrement d'un même côté du plan $z = 0$, et seul le point 0 est intersection du plan et de la surface (attention : *localement*).

— Un point **hyperbolique** si $rt < s^2$. La surface passe, localement, des deux côtés du plan tangent.

— Un point **parabolique** si $rt = s^2$ et $(r, s, t) \neq (0, 0, 0)$.

— Un point **plat** si $r = s = t = 0$.

La quantité $rt - s^2$ est appelée **courbure de Gauss** ou **courbure totale** de la surface S en 0.

On appelle **première forme fondamentale** d'une surface S au moins C^1 en un point $x \in S$ l'application qui à deux vecteurs appartenant à la direction du plan tangent à S en x associe leur produit scalaire. La première forme fondamentale est une forme bilinéaire, dépendant de x ; on la notera par la suite, dans une base de la direction du plan tangent, $g((x_1, x_2), (y_1, y_2)) = \sum_{(i,j)\in\{1,2\}^2} g_{i,j} x_i y_j$.

Application 0.13 L'utilisation des formes fondamentales (et de leurs préservations) permet de montrer que certaines formes ne sont *pas* homéomorphes. Elle permet aussi de définir l'aire. En effet, l'aire définie comme suit coïncide bien avec l'idée intuitive qu'on en a :

DÉFINITION 0.16 **aire**
Soit une variété paramétrée $S = \{f(x, y); (x, y) \in \Omega \subset \mathbb{R}^2\}$ avec $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ linéairement indépendants.

Alors, on définit l'aire comme suit

$$Aire(S) = \int_\Omega \sqrt{det fff(x, y)}$$

avec fff la forme fondamentale par rapport aux vecteurs $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$.

Cette définition équivaut à l'intégrale

$$Aire(S) = \int_\Omega \|\frac{\partial f}{\partial x} \wedge \frac{\partial f}{\partial y}\| \text{ (norme du produit vectoriel)}$$

Avec l'une ou l'autre de ces définitions, on peut montrer (laissé en exercice au lecteur) que l'aire est invariante par changement de variable, i.e. en remplaçant f par $f \circ \varphi$ et Ω par Ω', où $\varphi : \Omega' \to \Omega$ est un $\mathcal{C}^1$-difféomorphisme.

Références

[1] H. Cartan, *Calcul différentiel*, Hermann, 1977.

[2] A. Chambert-Loir, S. Fermigier, *Exercices de mathématiques pour l'agrégation, Analyse 3*, Masson, 1996.

[3] S. Lang, *Real analysis*, Addison-Wesley Publishing company, 1969.

[4] J. Vauthier, J.J. Prat, *Cours d'analyse mathématiques de l'intégration*, Masson, 1994.