Calcul différentiel

Jean-Paul Calvi

0.6.0

Université de Toulouse

CALCUL DIFFÉRENTIEL 0.6.0 UNIV.JEANPAULCALVI.COM

Première mise en ligne, à la version 0.4.0, le 28 mars 2011, sur le site jeanpaulcalvi.com

Développements 1 Insertion des démonstrations / non programmé

D'abord, l'Italie que tu crois déjà proche — et tu te prépares, ignorant, à pénétrer près d'ici dans ses ports —, une longue route déroutante, bordant de longues terres, t'en sépare bien loin.

Virgile, Énéide

Préfaces

0.4

Ce texte est issus d'un cours que j'ai donné dans le cadre du module de calcul différentiel de la licence de mathématiques appliquée à l'ingénierie (MAPI) de l'université Paul Sabatier. Il contient les exercices traités en travaux dirigés ainsi que les énoncés des contrôles. Pour l'essentiel, le contenu correspond à un celui d'un cours d'introduction classique. Le cadre est celui des espaces vectoriels normés abstraits mais, dans toutes les applications, ceux-ci seront de dimension finie. Les aspects computationnels sont privilégiés. Les possibilités qu'offrent les logiciels de calculs formels seront signalées et des exemples seront proposés utilisant le logiciel MAXIMA librement téléchargeable sur la page http ://maxima.sourceforge.net
Les graphes de fonctions de deux variables et certains autres calculs sont effectués à l'aide du logiciel SCILAB aussi librement téléchargeable sur http ://www.scilab.org
Foix, Mars 2011, JPC.

0.5 et 0.6

J'ai apporté quelques améliorations typographiques et augmenté le texte de quelques exercices, spécialement dans le chapitre sur les équations différentielles. J'ai commencé la rédaction des démonstrations des résultats fondamentaux. Foix, Avril 2013, JPC

Table des matières

1

Différentielles

Les lettres E, F, G désigneront toujours des espaces vectoriels normés. Lorsque les normes devront être spécifiées, nous écrirons $\|\cdot\|_E$, $\|\cdot\|_F$ et $\|\cdot\|_G$. Les lettres Ω et U seront réservées aux sous-ensembles ouverts de l'espace vectoriel normé dans lequel ils se trouvent.

§ 1. Introduction et définition

Le **nombre dérivé** d'une fonction $f : \mathbb{R} \to \mathbb{R}$ en un point x_0 est habituellement introduit comme la limite, lorsque h tend vers 0, du **taux d'accroissement** $(f(x_0 + h) - f(x_0))/h$. Si cette définition reste valable pour les fonctions $f : \mathbb{R} \to F$ qui prennent leurs valeurs dans un espace vectoriel normé quelconque F, elle ne peut pas être étendue aux fonctions dont la variable est un vecteur, pour la simple raison que la division par un vecteur h n'a plus de sens. Lorsque les vecteurs sont des éléments de $\mathbb{R}^n$, l'idée classique consiste à fixer toutes les variables sauf une puis à dériver la fonction par rapport à la variable restante. Autrement dit, on étudie la dérivée de la fonction

$$f_i : t \in \mathbb{R} \to f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n) \in F,$$

et une telle dérivée s'appelle une **dérivée partielle**. Malheureusement, une fonction peut admettre des dérivées partielles en un point sans pour autant avoir un comportement régulier en ce point (voir l'exemple 2). D'un point de vue pratique, cependant, l'emploi des dérivées partielles est dans bien des cas suffisant. On verra que si toutes les dérivées partielles sont des fonctions continues alors la fonction elle-même a les propriétés de régularité attendues (théorème 12). Pour cette raison, on a pendant longtemps pu se limiter à cette notion de dérivée partielle et beaucoup d'ingénieurs, encore aujourd'hui, ne connaissent et n'utilisent qu'elle. L'inconvénient principal de se limiter à l'étude des dérivées partielles est le suivant. Pour pouvoir parler de dérivées partielles, il faut d'abord disposer de variables ; lorsque l'on travaille sur $\mathbb{R}^n$ avec des vecteurs $x = (x_1, \ldots, x_n)$ les variables x_i paraissent s'imposer. Pourtant, nous savons qu'en géométrie le choix d'un repère adapté peut considérablement simplifier la solution d'un problème. De même il est tout à fait possible que dans un problème donné, il soit plus utile de travailler avec une base (v_i) de $\mathbb{R}^n$ différente de la base canonique et, avec des vecteurs qui s'écrivent $x = \sum_{i=1}^n c_i(x) v_i$, les variables naturelles sont les nombres $c_i(x)$ (qui se déduisent des x_i par une application linéaire bijective). Il faut alors vérifier que les énoncés des théorèmes construits en utilisant les dérivées partielles ne dépendent pas du choix préalable des variables c'est-à-dire du repère et ceci finit par conduire à la théorie que nous allons présenter. Celle-ci est intrinsèque, c'est-à-dire indépendante de la manière par laquelle on représente la variable. Un troisième inconvénient est qu'il est possible de considérer des dérivées partielles maniables (dans le sens donné ci-dessus) uniquement dans la mesure où nous manipulons des fonctions définies sur un espace de dimension finie et cette approche masque la généralisation de la notion de dérivée aux fonctions définies sur un espace vectoriel de dimension infinie. Cette difficulté est sans doute moins déterminante que la précédente car le calcul différentiel sur des espaces vectoriels de dimension infinie possède un champ d'application relativement limité qui ne justifierait pas en lui-même l'enseignement de cette notion à un niveau élémentaire.

1.1 La définition d'une différentielle

Pour trouver la meilleure généralisation de la notion de dérivée, nous devons nous concentrer sur une autre propriété que celle du taux d'accroissement. La propriété caractéristique généralisable du nombre dérivé est qu'il permet de construire une approximation locale de la fonction f au point x_0 ; autrement dit, la fonction affine $T_{x_0} : x \to f(x_0) + f'(x_0)(x - x_0)$ fournit une approximation locale de la fonction f et cette approximation est précise à l'ordre 1 dans le sens où

$$f(x) - T_{x_0}(x) = (x - x_0)\,\epsilon(x - x_0) \quad \text{avec} \lim_{x \to x_0} \epsilon(x - x_0) = 0. \tag{1.1}$$

Dans l'expression 'approximation locale', l'adjectif 'local' est important. Il signifie que la seule information dont nous disposons est une information 'à la limite', en particulier la seule formulation (1.1) ne permet aucune estimation (quantitative) de l'erreur entre $f(x)$ et $T_{x_0}(x)$ en dehors de x_0 (où elle est évidemment égale à 0). Les informations purement locales ont peu d'intérêt, mais comme nous l'a déjà appris l'analyse des fonctions de la variable réelle, des informations locales en un ensemble ouvert de points conduisent, sous certaines hypothèses, a des informations globales.

Voyons si la relation (1.1) garde un sens lorsque la fonction f est définie sur un ouvert Ω d'un espace vectoriel normé E à valeurs dans un autre espace vectoriel normé F. Il y a un problème superficiel, le produit $(x - x_0)\,\epsilon(x - x_0)$ n'aura pas de sens en général dans E mais nous pouvons, sans rien changer à sa signification, récrire (1.1) sous la forme

$$|f(x) - T_{x_0}(x)| = |x - x_0|\,\epsilon(x - x_0) \quad \text{avec} \lim_{x \to x_0} \epsilon(x - x_0) = 0 \tag{1.2}$$

où la nouvelle fonction ϵ est la valeur absolue de la précédente. Cette dernière relation se laisse étendre en

$$\|f(x) - T_{x_0}(x)\|_F = \|x - x_0)\|_E\,\epsilon(x - x_0) \quad \text{avec} \lim_{x \to x_0} \epsilon(x - x_0) = 0. \tag{1.3}$$

Il reste à voir quel sens nous pouvons donner à l'application T_{x_0}. La définition d'une **application affine** de E dans F ne cause aucune difficulté : c'est une application de la forme $A(x) + b$ où A est une application linéaire de E dans F et $b \in F$. L'application $T_{x_0}(x)$ devra donc être recherchée sous la forme $T_{x_0}(x) = f(x_0) + A(x - x_0)$ et c'est alors l'application linéaire A qui va servir de généralisation du nombre dérivé.

Definition 1.1. Soit $f : \Omega \subset E \to F$ et $a \in \Omega$. Nous dirons que f est **différentiable** en a s'il existe une application linéaire continue A de E dans F, i.e. $A \in \mathcal{L}(E, F)$, et une fonction réelle ϵ définie sur un voisinage de l'origine dans E, de limite nulle en 0 et telle que

$$\|f(a + h) - f(a) - A(h)\|_F = \|h\|_E\,\epsilon(h). \tag{1.4}$$

Soulignons le fait que l'application linéaire A doit être continue. L'utilité de cette hypothèse apparaîtra dès la démonstration du théorème 2. Naturellement, d'après le théorème 4 de l'annexe, la vérification de la continuité de A est superflue lorsque E est de dimension finie. Notons aussi que, dans la pratique, on reformule souvent la définition (1.4) en faisant intervenir une fonction d'erreur allant de U vers F. Ceci est expliqué plus bas, au point 1.3. Soulignons encore que la condition que ϵ soit de limite nulle dépend, en général, des normes utilisées sur E et sur F. Cependant lorsque E et F sont de dimensions finies, toutes les normes sont équivalentes (Théorème 1) de l'appendice 5) et nous sommes libres de choisir celles qui nous apparaitront les plus commodes.

Théorème 1 (et définition). *Il existe au plus une application $A \in \mathcal{L}(E, F)$ satisfaisant (1.4). Lorsqu'elle existe cette application linéaire est notée $df(a)$ et est appelée la **différentielle** de f en a.* ■

Démonstration. Supposons que A et B soient deux applications linéaires satisfaisant la condition. Prenons $h \in E \setminus 0$, en soustrayant les relations de définition avec $A(h)$ et $B(h)$, nous obtenons

$$\|A(h) - B(h)\|_F = \|h\|_E\,\epsilon_{A-B}(h)$$

où ϵ_{A-B} est la différence de la fonction epsilon correspondant à la définition pour A et de celle pour B. Fixons un élément u de E et posons $h = tu$ avec $t \in \mathbb{R}^*$. La relation ci-dessus donne, en tenant compte de la linéarité de A et B,

$$\|t(A-B)(u)\|_F = |t|\|u\|_E \epsilon_{A-B}(tu) \implies \|(A-B)(u)\|_F = \|u\|_E \epsilon_{A-B}(tu).$$

En faisant $t \to 0$ dans cette relation, puisque $\epsilon_{A-B}(tu) \to 0$, nous obtenons $(A-B)(u) = 0$. Puisque le raisonnement est valable pour tout $u \in E$, nous avons $A-B = 0$ ce qui achève la démonstration du théorème. ■

Nous avons dit que $df(a)$ était la meilleure généralisation du nombre dérivé. Il y a cependant une différence conceptuelle profonde entre les deux objets, une différence qui est à l'origine des difficultés rencontrées par certains étudiants dans l'étude du calcul différentiel : si le nombre dérivé, comme son nom l'indique est un nombre, la différentielle $df(a)$, elle, est une application linéaire de l'espace vectoriel dans lequel se trouve Ω (et a) à valeurs dans l'espace vectoriel d'arrivée de f. Dans le cas classique d'une application de $U \subset \mathbb{R}$ dans $\mathbb{R}$, la différentielle est l'application linéaire de $\mathbb{R}$ dans $\mathbb{R}$ définie par $df(a)(h) = h \times f'(a)$ (où $\times$ est la multiplication dans $\mathbb{R}$),

$$df(a) : h \in \mathbb{R} \to f'(a) \times h \in \mathbb{R}. \tag{1.5}$$

Le **nombre dérivé** est donc le réel qui détermine complètement l'application linéaire $df(a)$. En fait $f'(a) = df(a)(1)$. La définition de la différentielle est illustrée par la figure 1.

Théorème 2. *Si f est différentiable en $a \in \Omega \subset E$ alors elle est aussi continue en a.* ■

Démonstration. Nous tirons de la définition l'inégalité

$$\|f(a+h) - f(a)\|_F \leq \|df(a)(h)\|_F + \|h\|_E \epsilon(h).$$

Le droite terme de droite tend vers 0 lorsque h tend vers 0 par continuité de $df(a)$ tandis que le second par définition de ϵ. Il suit que $\lim_{h\to 0} \|f(a+h) - f(a)\|_F = 0$ et c'est la propriété de continuité de f en a. ■

1.2 L'application tangente et l'hyperplan tangent

L'application $x \in E \to f(a) + df(a)(x-a) \in F$ s'appelle l'**application affine tangente** à f en a. L'**hyperplan tangent** H_a au graphe de f au point a est le graphe de l'application affine tangente. Si $f : U \subset \mathbb{R}^n \to \mathbb{R}$ alors cet hyperplan qui se trouve dans $\mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$ a pour équation

$$x_{n+1} = f(a) + df(a)(x_1 - a_1, \ldots, x_n - a_n). \tag{1.6}$$

Une visualisation du plan tangent pour une fonction réelle de deux variables se trouve dans l'illustration 2.

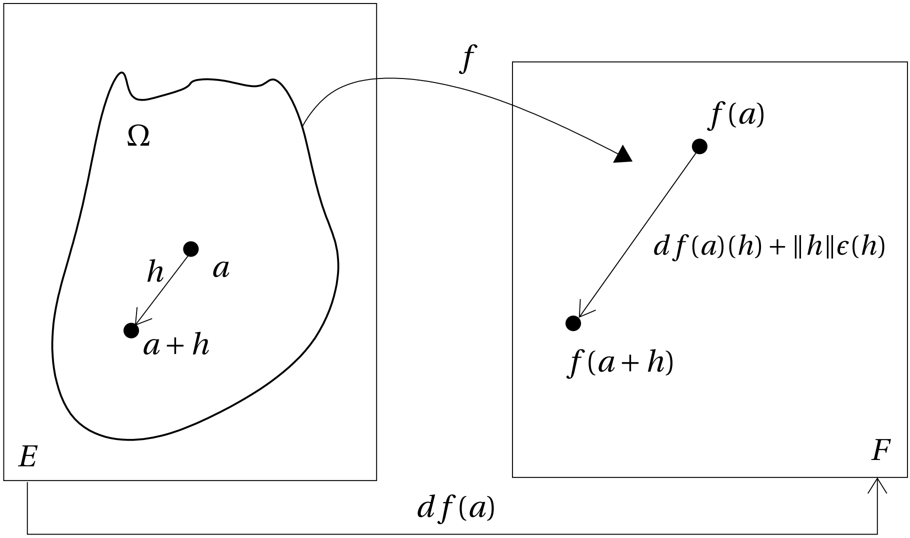


Figure 1 – Définition d'une différentielle.

E. 1. Vérifier que l'équation ci-dessus est toujours celle d'un hyperplan de $\mathbb{R}^{n+1}$, c'est-à-dire d'un sous-espace affine de dimension * n, justifiant ainsi la terminologie.

Dans l'étude du calcul différentiel des fonctions réelles d'une variable réelle il est facile, et très utile, de visualiser la plupart des notions sur le graphe des fonctions. Ce support n'est plus disponible lorsque nous étudions des fonctions définies sur un ouvert d'un espace vectoriel général. Il reste cependant la possibilité de visualiser les notions lorsque nous travaillons avec des fonctions réelles de deux variables réelles, c'est-à-dire définies sur un ouvert de $\mathbb{R}^2$. La figure 2 représente le graphe sur $[-4,4]\times[-4,4]$ de la fonction polynomiale f définie sur $\mathbb{R}^2$ par $f(x,y)=x^4+y^4$ ainsi que le plan tangent à son graphe au point $a=(2,2)$. Ce plan est directement déterminé par la différentielle de f en a, laquelle est définie, comme nous le verrons plus loin, par $df(a)(h_1,h,_2)=32h_1+32h_2$ de sorte que l'équation du plan est

$$\begin{aligned} z &= f(a)+df(a)((x,y)-(2,2)) && (1.7)\\ &= 32+32(x-2)+32(y-2). && (1.8)\end{aligned}$$

Notons l'étroite similarité entre la notion de droite tangente déjà connue et celle de plan tangent.

Il faut prendre garde que la représentation graphique d'une fonction réelle de deux variables réelles implique l'usage d'une perspective et d'un point de vue. Nous

*. Un sous-espace affine de dimension p est l'image par une translation ($x\to x+u$) d'un sous-espace vectoriel de dimension p.

verrons plus loin (exemple 2) qu'une perspective et un point de vue mal choisis peuvent conduire à des erreurs d'appréciation sur les propriétés d'une fonction.

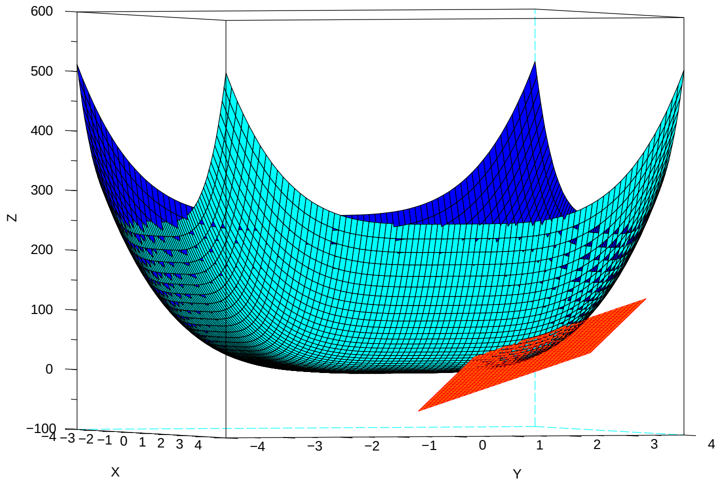


Figure 2 – Graphe de la fonction différentiable $f(x,y) = x^4 + y^4$ et du plan tangent au point $(2,2)$.

Code Scilab 1. — La figure 2 est obtenue avec le code suivant.

```
deff('z=f(x,y)','z=((2+x)^4+(2+y)^4)-(32+32*x+32*y)');
x=-0.1:0.01:0.1;
y=x ;
clf();
fplot3d(x,y,f,alpha=5,theta=20)
```

La table 1 donne une estimation de l'erreur entre le fonction f et sa fonction affine tangente au point $(2,2)$, donnée par $|f(2+h_1,2+h_2) - df(2,2)(h_1,h_2)|$ pour $h = (h_1,h_2) \in [-\delta,\delta]$ avec différentes valeurs de δ.

Code Scilab 2. — Les calculs de la Table 1 sont obtenus avec Scilab en utilisant le code suivant.

```
deff('z=f(x,y)','z=((2+x)^4+(2+y)^4)-(32+32*x+32*y)');
x=-0.1:0.01:0.1 ;y=x ;
max(abs(f(x,y)))
xx=-0.05:0.005:0.05 ;yy=xx;
max(abs(f(xx,yy)))
xxx=-0.01:0.001:0.01 ;yyy=xxx;
max(abs(f(xxx,yyy)))
```

δ	Erreur
10^{-1}	$5\,.\,10^{-1}$
$5\,.\,10^{-2}$	$12\,.\,10^{-2}$
10^{-2}	$4\,.\,10^{-3}$

Table 1 – Estimation de l'erreur entre la fonction $f(x,y) = x^4 + y^4$ et son application affine tangente au voisinage du point $(2,2)$.

1.3 Comment rechercher une différentielle à partir de la définition

Nous allons rapidement développer des outils de calculs des différentielles mais comme toujours en mathématiques certains exemples doivent être traités directement à partir de la définition. Pour appliquer cette définition à une fonction donnée $f : \Omega \subset E \to F$, il n'y a pas d'autres solutions que de calculer la différence $f(a+h) - f(a)$ et d'essayer de faire apparaître à l'aide de manipulations algébriques ou analytiques une application $A(h)$ et une **erreur** ou **fonction d'erreur** $R(h)$ telles que

$$f(a+h) - f(a) = A(h) + R(h) \tag{1.9}$$

avec $A \in \mathcal{L}(E;F)$ et R définie dans un voisinage de 0 avec la propriété que

$$\lim_{h\to 0} \frac{\|R(h)\|_F}{\|h\|_E} = 0. \tag{1.10}$$

Nous avons alors $df(a) = A$ et la fonction ϵ est définie par $\epsilon(0) = 0$ et

$$\epsilon(h) = \|R(h)\|_F / \|h\|_E, \quad \text{pour } \|h\|_E \text{ assez petit non nul.}$$

Exemple 1. Toute application affine $\mathcal{A} : x \in E \to A(x) + b$ avec $A \in \mathcal{L}(E)$ est différentiable en tout point a de E et

$$d\mathcal{A}(a) = A, \quad a \in E. \tag{1.11}$$

E. 2. Démontrer le résultat.

Exemple 2. Soient E un espace de Banach et B une forme bilinéaire symétrique continue sur $E \times E$. La forme quadratique Q définie sur E par $Q(x) = B(x,x)$ est différentiable en tout point de E et sa différentielle $dQ(a)$ au point a est donnée par la relation

$$dQ(a)(h) = 2B(a,h), \quad a, h \in E. \tag{1.12}$$

E. 3. Démontrer le résultat. Que dire lorsque B n'est pas supposée symétrique ?

E. 4. Soit f une application trilinéaire continue sur $E_1 \times E_2 \times E_3$ à valeur dans F. Montrer que f est différentiable en tout point de $E_1 \times E_2 \times E_3$ et calculer sa différentielle. Étendre le résultat au cas des fonctions n-linéaires, $n \geq 4$.

1.4 Deux règles de calculs

Les deux résultats suivants sont des conséquences simples de la définition qui s'avèrent très utiles pour le calcul pratique des différentielles.

Théorème 3 (Linéarité de la différentiabilité). *Soient $a \in \Omega \subset E$, f et g deux fonctions définies sur Ω à valeurs dans F et $\lambda \in \mathbb{R}$. Si f et g sont différentiables en a alors $f + \lambda g$ est différentiable en a et*

$$d(f + \lambda\, g)(a) = df(a) + \lambda\, dg(a). \tag{1.13}$$

■

Démonstration. Nous utilisons la technique présentée au point 1.3 en utilisant R_f (resp. R_g) pour les fonctions d'erreur correspondant à f (resp. à g). Nous observons que

$$\begin{aligned}((f + \lambda\, g)(a+h) - (f + \lambda\, g)(a) - (df(a) + \lambda\, dg(a))(h) \\ = \{f(a+h) - f(a) - df(a)(h)\} - \lambda\{g(a+h) - f(a) - df(a)(h)\} \\ = R_f(h) + \lambda R_g(h).\end{aligned} \tag{1.14}$$

Pour montrer que $df(a)+\lambda\, dg(a)$, il nous suffit alors d'établir que la fonction $R(h) := R_f(h) + \lambda R_g(h)$ satisfait les conditions des fonctions d'erreur, à savoir

$$\lim_{h\to 0} \frac{\|R(h)\|_F}{\|h\|_E} = 0$$

mais cela résulte immédiatement du fait que R_f et R_g satisfont cette même propriété. ■

E. 5. Soit $f : U \subset E \to F$ différentiable en $a \in U$. Montrer que la fonction g définie sur U par $g(x) = f(x) - f(a) - df(a)(x - a)$ est différentiable en a et calculer sa différentielle.

Théorème 4 (Différentielle des fonctions à valeurs dans un espace produit). *Soit $f : \Omega \subset E \to F$ et $a \in \Omega$. Si F est un produit d'espaces vectoriels normés, $F = F_1 \times F_2 \times \cdots \times F_m$ *, alors, notant f_i, $i = 1, \ldots, m$, les composantes de f de sorte que*

$$f(x) = (f_1(x), \ldots, f_n(x)), \quad x \in \Omega, \tag{1.15}$$

la fonction f est différentiable en a si et seulement si toutes les fonctions f_i sont différentiables en a et

$$df(a) = (df_1(a), \ldots, df_m(a)) \tag{1.16}$$

*. Nous avons $\|x\|_F = \max_{i=1,\ldots,m} \|x_i\|_{F_i}$ pour $x = (x_1, \ldots, x_n)$ voir 2.2.

où $(df_1(a), \dots, df_n(a))$ *désigne l'application linéaire de* E *dans* F *définie pour* $h \in E$ *par*

$$\big(df_1(a), \dots, df_m(a)\big)(h) = \big(df_1(a)(h), \dots, df_n(a)(h)\big). \tag{1.17}$$

Démonstration. Prenons $h \in F_i$ et notons $[df(a)(h)]_i$ la composante de $df(a)(h)$ qui appartient à F_i. La différentiabilité de f_i se déduit de la différentiabilité de f grâce à l'inégalité suivante qui utilise que la norme de chacune des composantes d'un élément de f est majoré par la norme de cet élément,

$$\begin{aligned} \|R_{f_i}(h)\|_{F_i} &:= \|f_i(a+h) - f_i(a) - [df(a)(h)]_i\|_{F_i} \\ &\leq \|f(a+h) - f(a) - df(a)(h)\|_F = \|R_f(h)\|_F. \end{aligned} \tag{1.18}$$

Réciproquement, si chacune des f_i est différentiable alors l'égalité

$$\begin{aligned} \|R_f(h)\|_{F_i} &:= \|(\dots, f_i(a+h) - f_i(a) - df_i(a)(h), \dots)\|_F \\ &= \|(\dots, R_{f_i}(h), \dots)\|_F = \max_{i=1,\dots,m} \|R_{f_i}(h)\|_{F_i}. \end{aligned} \tag{1.19}$$

montre que la fonction R_f satisfait la condition des fonctions d'erreurs dès que chacune des fonctions d'erreur R_{f_i} les satisfait. ■

Exemple 3. Soient E un espace de Banach, Q une forme quadratique continue sur E et L une forme linéaire continue. L'application $f : E \to \mathbb{R}^2$ définie par $f(x) = (Q(x), L(x))$ est différentiable en tout point de E et $df(x)(h) = (2B(x,h), L(h))$, $h \in H$.

E. 6. Supposons que F soit la somme directe de ses sous-espaces F_i,

$$F = F_1 \oplus F_2 \oplus \cdots \oplus F_n,$$

de sorte que que pour tout $x \in \Omega \subset E$, l'élément $f(x)$ de F s'écrive $f(x) = f_1(x) + f_2(x) + \cdots f_n(x)$ avec $f_i(x) \in F_i$, $i = 1, \dots, n$. Quel est le lien entre la différentiabilité de la fonction f et celle des fonctions f_i ?

1.5 La fonction différentielle

Il peut arriver que $f : \Omega \subset E \to F$ soit différentiable en tout point a d'un sous-ensemble ouvert U de Ω. Dans ce cas la fonction

$$df : a \in U \subset E \to df(a) \in \mathcal{L}(E;\, F) \tag{1.20}$$

est bien définie. Puisque l'ensemble de départ U est un ouvert d'une espace vectoriel normé et l'ensemble d'arrivée est un espace vectoriel normé (par la norme des applications linéaires, voir le théorème 3) il est légitime d'étudier la continuité de l'application df sur U. Lorsque celle-ci est continue en tout point de U, nous disons que f est **continûment différentiable** sur U, ou encore, que f est de classe C^1 sur U, et nous écrivons $f \in \mathsf{C}^1(U)$, ou, s'il est nécessaire de faire apparaître l'espace d'arrivée, $\mathsf{C}^1(U,\, F)$.

E. 7. Les fonctions considérées aux exemples 1 et 2 sont-elles de classe C^1.

Des théorèmes 3 et 4, nous déduisons immédiatement le

Théorème 5. $\mathsf{C}^1(U,\, F)$ *est un espace vectoriel et lorsque* F *est un produit cartésien,* $F = F_1 \times \cdots \times F_m$*, une fonction* f *appartient à* $\mathsf{C}^1(U,\, F)$ *si et seulement si chacune de ses composantes* f_i *appartient à* $\mathsf{C}^1(U,\, F_i)$*,* $i = 1, \ldots, m$*.* ■

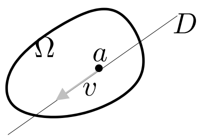
§ 2. Dérivées suivant un vecteur

2.1 Définition

Soient $f : \Omega \subset E \to F$, $a \in \Omega$ et $v \in E$. Nous dirons que f est dérivable suivant le (ou le long du) vecteur v si la limite suivante, notée $D_v f(a)$, existe

$$D_v f(a) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t}. \tag{2.1}$$

Nous prenons ici la limite d'une fonction définie sur un voisinage de 0 dans $\mathbb{R}$ et à valeurs dans F, à savoir la fonction $g : t \to (f(a+tv) - f(a))/t$.



Le fait que g soit effectivement définie sur un voisinage de 0 privé de 0 lui-même* provient du fait que a étant un élément de l'ouvert Ω, pour $|t|$ assez petit, $a + tv$ est encore un élément de Ω de sorte que $f(a+tv)$ est bien défini. Dans le calcul de $D_v(f)(a)$ seules interviennent les valeurs de la restriction de f à $D \cap \Omega$ où D est la droite passant par a et de vecteur directeur v. Il est naturel d'appeler $D_v f(a)$ la dérivée de f en a suivant le vecteur v car $D_v f$ n'est autre que la dérivée en 0 de la fonction de la variable réelle $t \to f(a+tv) \in F$. Remarquons que $D_0 f(a)$ existe toujours (et vaut 0).

E. 8. Soit $N : \mathbb{R}^2 \to R$ définie par $N(x, y) = \max\{|x|, |y|\}$. Calculer $D_v N(a)$ pour toutes les valeurs de a et v pour lesquelles c'est possible.

2.2 Les dérivées partielles ordinaires

Lorsque $E = \mathbb{R}^n$, $F = \mathbb{R}$ et v et le i-ième vecteur de la base canonique,

$$v = e_i = (0, \ldots, 0,\, 1,\, 0, \ldots,\, 0)$$

*. Un tel ensemble s'appelle un voisinage épointé de 0.

avec l'élément 1 à la i-ème coordonnée alors, lorsqu'il existe, le nombre $D_v f(a)$ n'est autre que la **dérivée partielle** de f au point a par rapport à la i-ème variable que nous noterons dans ce cours $\partial_i f(a)$ et qui est communément noté $\frac{\partial f}{\partial x_i}(a)$,

$$\partial_i f(a) = D_{e_i} f(a). \tag{2.2}$$

2.3 Le lien entre les dérivées directionnelles et la différentielle d'une fonction

Théorème 6. *Soit $v \in E$ et $f : \Omega \subset E \to F$. Si f est différentiable en $a \in \Omega$ alors la dérivée de f suivant v en a, $D_v f(a)$, existe et*

$$D_v f(a) = df(a)(v). \tag{2.3}$$

■

Démonstration. Si $v = 0$, la relation est évidente. Nous supposons $v \neq 0$. En appliquant la relation de différentiabilité avec R_f et $h = tv$, nous obtenons $f(a + tv) - f(a) = tdf(a)(v) + R_f(tv)$; en divisant par t et en faisant $t \to 0$, nous obtenons $D_v(f)(a) = df(a)(v)$ puisque $\|R_f(tv)/t\|_F = \|v\|_F \|R_f(tv)\|_F / \|tv\|_F \to 0$ lorsque $t \to 0$. ■

Ce résultat implique une relation simple entre la différentielle et les dérivées partielles sous l'hypothèse que la fonction soit différentiable. Supposons que E soit de dimension finie et que $(v_1, \dots, v_n)$ soit une base de E alors si $h = \lambda_1 v_1 + \cdots + \lambda_n v_n$ et $f = \Omega \subset E \to F$ est différentiable en $a \in \Omega$, nous avons

$$df(a)(h) = \lambda_1 D_{v_1} f(a) + \cdots + \lambda_n D_{v_n} f(a). \tag{2.4}$$

En particulier, si $E = \mathbb{R}^n$, $F = \mathbb{R}$ et $h = (h_1, h_2, \dots, h_n)$,

$$df(a)(h) = h_1 \partial_1 f(a) + \cdots + h_n \partial_n f(a). \tag{2.5}$$

On trouve souvent dans la littérature mathématique la formule (2.5) sous la forme

$$df(a) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)\, dx_j, \tag{2.6}$$

où dx_j est employé pour désigner la forme linéaire définie par $dx_j(h) = h_j$. À cause de cette relation, la différentielle est parfois appelé **dérivée totale**.

Exemple 1. La fonction $\|\cdot\|_2 = x \in \mathbb{R}^2 \to \|x\|_2 = (x_1^2 + x_2)^{1/2} \in \mathbb{R}$ n'est pas différentiable en $0 = (0, 0)$. Cet exemple n'est pas typique de d'analyse en plusieurs variables car dans, ce cas comme dans celui des fonctions d'une variable réelle, la non différentiabilité de la fonction en 0 s'observe par la présence sur le graphe (figure 3) du point anguleux $(0, 0, 0)$.

E. 9. Démontrer que la fonction $\|\cdot\|_2$ n'est pas différentiable en 0.

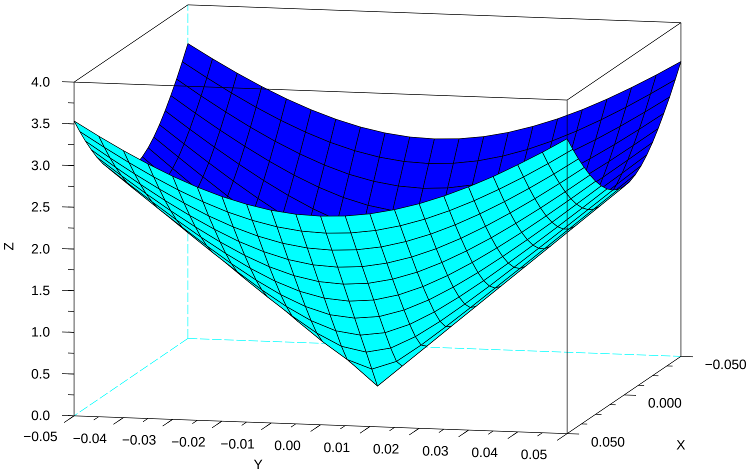


Figure 3 – Graphe de la fonction $\|\cdot\|_2$.

2.4 La matrice d'une différentielle

Supposons maintenant que $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ et notons f_i, $i = 1, \ldots, m$, les composantes de f de sorte que

$$f(x) = \big(f_1(x),\, f_2(x), \ldots,\, f_m(x)\big), \quad x = (x_1, \ldots, x_n).$$

L'utilisation conjointe du théorème 4 et de la relation (2.5) donne

$$df(a)(h) = \Big(df_1(a)(h),\, df_2(a)(h), \ldots,\, df_m(a)(h)\Big) \tag{2.7}$$

$$= \left(\sum_{j=1}^{n} \partial_j f_1(a) h_j, \sum_{j=1}^{n} \partial_j f_2(a) h_j, \ldots, \sum_{j=1}^{n} \partial_j f_m(a) h_j\right). \tag{2.8}$$

Puisque $df(a)$ est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$, nous pouvons déterminer sa matrice (dans les bases canoniques des espaces de départ et d'arrivée). Celle-ci se déduit immédiatement de la relation précédente qui s'écrit, en notant $df(a)(h) = (h'_1, \ldots, h'_n)$,

$$\begin{pmatrix} h'_1 \\ h'_2 \\ \dots \\ h'_i \\ \dots \\ h'_n \end{pmatrix} = \begin{pmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & \dots & \partial_n f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & \dots & \partial_n f_2(a) \\ \dots & \dots & \dots & \dots \\ \partial_1 f_i(a) & \partial_2 f_i(a) & \dots & \partial_n f_i(a) \\ \dots & \dots & \dots & \dots \\ \partial_1 f_m(a) & \partial_2 f_m(a) & \dots & \partial_n f_m(a) \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \dots \\ h_i \\ \dots \\ h_n \end{pmatrix}. \tag{2.9}$$

La matrice de $df(a)$ dans la base canonique est appelée la **matrice jacobienne** de f en a ou encore le jacobien de f en a, elle est notée $\mathbf{J}_f(a)$, et d'après la relation précédente,

$$\mathbf{J}_f(a) = (\partial_j f_i(a))_{i=1,\dots,m}^{j=1,\dots,n} \tag{2.10}$$

où, rappelons-le, i désigne la ligne et j la colonne. Lorsque $m = n$, nous pouvons calculer le déterminant de $\mathbf{J}_f(a)$. Celui joue un rôle très important en analyse et en géométrie. Certains auteurs réservent le mot **Jacobien** pour désigner $\det \mathbf{J}_f(a)$ plutôt que pour la matrice $\mathbf{J}_f(a)$.

E. 10. On considère l'application $f : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}^2$ définie par

$$f(r, \theta) = (r \cos\theta, r \sin\theta).$$

Montrer que cette fonction est différentiable en tout point de son domaine de définition et calculer $\det \mathbf{J}_f(a)$, $a \in \mathbb{R}^+ \times \mathbb{R}$.

2.5 Insuffisance des dérivées partielles

La seule existence des dérivées partielles, comme nous l'avions déjà indiqué dans l'introduction de ce chapitre, n'entraîne pas la différentiabilité.

Exemple 2. Considérons la fonction f définie sur $\mathbb{R}^2$ par

$$f(x, y) = \begin{cases} 0 & \text{si } (x, y) = (0, 0), \\ xy/\sqrt{x^2 + y^2} & \text{si } (x, y) \neq (0, 0). \end{cases} \tag{2.11}$$

L'application f est continue sur $\mathbb{R}^2$ et les dérivées en $(0, 0)$ existent. Elles sont données par $\partial_1(f)(0, 0) = 0 = \partial_2(f)(0, 0)$. Pourtant la fonction n'est pas différentiable en 0. Il n'est pas difficile de montrer que $D_v f(0, 0)$ existe pour tout v non nul. Nous donnons le graphe de cette fonction selon deux points de vue différents. Sur la première représentation, la singularité de la fonction à l'origine n'est pas du tout évidente. Elle semble le devenir sur la seconde où, au voisinage de l'origine, la représentation de la fonction évoque une 'nappe vrillée'. Mais les axes et la perspective de cette seconde représentation peuvent aussi fausser le jugement.

E. 11. Démontrer les assertions ci-dessus.

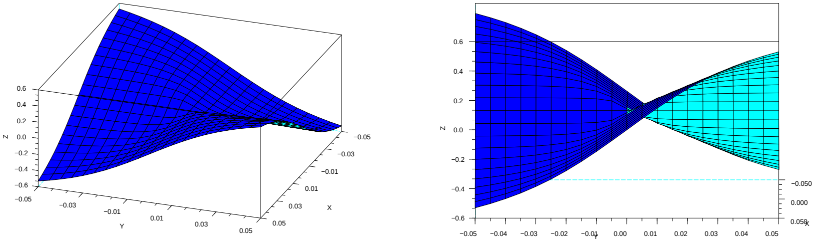


Figure 4 – Exemples de fonction non différentiable en un point : graphe de $xy/(\sqrt{x^2+y^2}$.

Nous verrons plus loin un résultat (théorème 12) auquel nous avons aussi déjà fait allusion dans l'introduction qui montre que l'existence de dérivées partielles continues impliquent en revanche que la fonctions est de classe C^1.

2.6 *Autres notations pour les différentielles.*

Il y a plusieurs notations concurrentes de celle que nous avons retenue. Là où nous écrivons $df(a)$ d'autres écrivent $Df(a)$ ou encore $f'(a)$. Cette dernière est probablement la plus commode. Nous y avons renoncé parce qu'elle peut créer une confusion chez les débutants : $f'(a)$ doit désigner une application linéaire alors qu'ils sont habitués à ce que $f'(a)$ désigne un nombre (ou un vecteur). Par ailleurs, pour éviter la multiplication des parenthèses, qui deviendra lourde lorsque nous utiliserons des fonctions composées, certains écrivent $df(a) \cdot h$ ou même $df(a)h$ à la place de $df(a)(h)$. Bien qu'ils soient utiles, nous essaierons ici d'éviter ces raccourcis et d'utiliser des parenthésages complets.

2.7 *Calcul automatique des dérivées partielles et des différentielles (avec Maxima)*

Le calcul d'une dérivée partielle d'une expression standard est très simple.

```
(%i1) f: x^2*y+x*cos(x+y);
```

(%o1) $x\cos(y+x) + x^2\, y$

```
(%i2) diff(f,x,1);
```

(%o2) $-x\sin(y+x) + \cos(y+x) + 2\,x\,y$

Le premier argument de la fonction *diff* est l'expression à dériver, le second est la variable par rapport à laquelle il faut dériver et le dernier est l'ordre de dérivation. Lorsque les deux derniers arguments ne sont pas mentionnés, le résultat est la différentielle de la fonction,

```
(%i16) diff(f);
```

(%o16) $\left(x^2 - x\sin(y+x)\right)\,\mathrm{del}(y) + \left(-x\sin(y+x) + \cos(y+x) + 2\,x\,y\right)\,\mathrm{del}(x)$

Ici les formes linéaires $dx : (h_1, h_2) \to h_1$ et $dy : (h_1, h_2) \to h_2$ sont notées $\mathrm{del}(x)$ et $\mathrm{del}(y)$. Observer que l'ordre des termes n'est pas nécessairement celui qui est attendu ou naturel.

Les variables peuvent être aussi nombreuses que nous voulons et la notation est libre,

```
(%i8)  g:x[1]*2^(x[1]+x[2]+x[3])+cos(x[2]*x[3]);
```

(%o8) $\cos(x_2\,x_3) + 2^{x_3+x_2+x_1}\,x_1$

```
(%i9)  diff(g,x[1]);
```

(%o9) $2^{x_3+x_2+x_1}\,x_1\log(2) + 2^{x_3+x_2+x_1}$

```
(%i10) diff(g);
```

(%o10) $\left(2^{x_3+x_2+x_1}\,x_1\log(2) - x_2\sin(x_2\,x_3)\right)\,\mathrm{del}(x_3)$
$+ \left(2^{x_3+x_2+x_1}\,x_1\log(2) - x_3\sin(x_2\,x_3)\right)\,\mathrm{del}(x_2)$
$+ \left(2^{x_3+x_2+x_1}\,x_1\log(2) + 2^{x_3+x_2+x_1}\right)\,\mathrm{del}(x_1)$

Le Jacobien est facilement calculable

```
(%i1)  f:exp(x)*cos(y);
```

(%o1) $e^x\cos(y)$

```
(%i2)  g:exp(x)*sin(y);
```

(%o2) $e^x\sin(y)$

```
(%i3)  j:jacobian([f,g],[x,y]);
```

(%o3) $\begin{pmatrix} e^x\cos(y) & -e^x\sin(y) \\ e^x\sin(y) & e^x\cos(y) \end{pmatrix}$

```
(%i4) jj:ev(j,[x=0,y=0]);
```

(%o4) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Observer l'utilisation de la fonction *ev* pour évaluer une expression, ici la matrice jacobienne lorsque $x = 0$ et $y = 0$. Le jacobien peut aussi servir à calculer une expression de $df(x)(h)$,

```
(%i5) jacobian([f],[x,y]).[h[1],h[2]];
```

(%o5) $h_1\, e^x \cos(y) - h_2\, e^x \sin(y)$

§ 3. Différentielle de la composée de deux fonctions

La formule établie ici est fondamentale. Pratiquement toutes les règles de calcul sur les différentielles s'en déduisent en l'associant à la définition, aux relations de linéarité et aux résultats pour les fonctions d'une seule variable. Certaines de ces règles sont établies à l'exercice 18.

Théorème 7 (Différentielle d'une composée). *Soit $f : \Omega \subset E \to F$ et $g : U \subset F \to G$. Supposons que $f(\Omega) \subset U$ de sorte que $g \circ f$ est bien définie sur Ω à valeurs dans G. Si f est différentiable en a et g est différentiable en $f(a)$ alors $g \circ f$ est différentiable en a et*

$$d\big(g \circ f\big)(a) = dg\big(f(a)\big) \circ df(a). \tag{3.1}$$

■

La construction du théorème est illustrée dans la figure 3.1.

Démonstration. Il s'agit de montrer que l'application linéaire (continue) $A := dg\big(f(a)\big) \circ df(a)$ satisfait les conditions de la définition, autrement dit,

$$(g \circ f)(a + h) - (g \circ f)(a) - A(h) = R(h), \quad \text{avec } \lim_{h \to 0} \|R(h)\|_G / \|h\|_E = 0.$$

Or, la différentiabilité de f en a puis celle de g en $f(a)$ nous donne

$$\begin{aligned} g(f(a+h)) &= g(f(a) + H), \quad (H = df(a)(h) + R_f(h)) \\ &= g(f(a)) + dg(f(a))\,(df(a)(h) + R_f(h)) + R_g\,(df(a)(h) + R_f(h))\,. \end{aligned} \tag{3.2}$$

Si bien que la fonction $R(h)$ s'exprime comme

$$R(h) = dg(f(a))\,(R_f(h)) + R_g\,(df(a)(h) + R_f(h))\,.$$

Cette fonction d'erreur satisfait la condition requise. Pour s'en assurer, il suffit d'observer d'une part que

$$\frac{\|dg(f(a))\,(R_f(h))\,\|_F}{\|h\|_E} \leq \|dg(f(a))\| \frac{\|R_f(h)\|_F}{\|h\|_E} \to 0 \ (h \to 0); \tag{3.3}$$

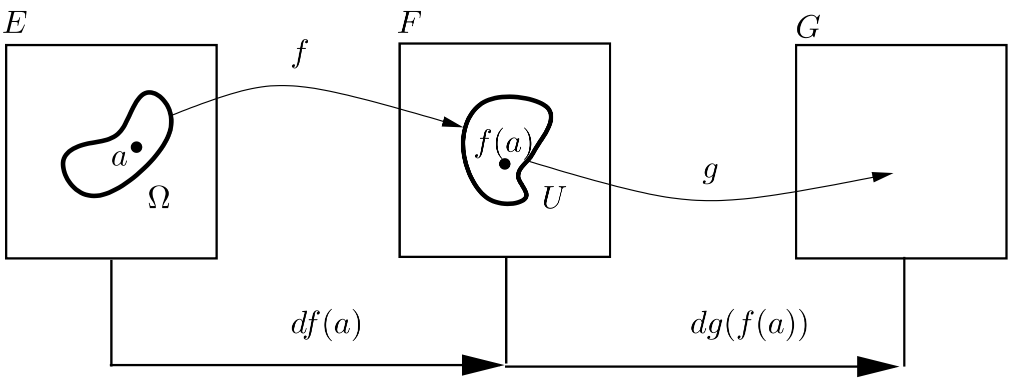
et, d'autre part

$$\frac{1}{\|h\|_E}\|R_g\,(df(a)(h) + R_f(h))\,\|_G = \frac{\|R_g\,(df(a)(h) + R_f(h))\,\|_G}{\|df(a)(h) + R_f(h)\|_F} \times \frac{\|df(a)(h) + R_f(h)\|_F}{\|h\|_E}.$$

Le premier terme sur la droite tend vers 0 par propriété de la fonction d'erreur R_g et le second tend vers 0 avec le même argument que dans (3.3).

■

E. 12. Quelle est la traduction en terme de Jacobiens de la formule précédente ? Expliciter la formule donnant la i-ème dérivée partielle d'une composée $u \circ v$.



$$d(g \circ f)(a) = dg(f(a)) \circ df(a)$$

Figure 5 – Différentielle d'une fonction composée

Exemple 1. Soit A une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$, $b \in \mathbb{R}^m$ et f une fonction différentiable en tout point de $\mathbb{R}^m$. Si g est définie sur $\mathbb{R}^n$ par $g(x) = f(A(x) + b)$ alors g est différentiable en tout point c de $\mathbb{R}^n$ et

$$dg(c)(h) = df(A(c) + b)(A(h)), \quad h \in \mathbb{R}^n.$$

Exemple 2. Voici comment calculer une dérivée partielle d'une fonction composée (compliquée) définie sur un sous-ensemble de $\mathbb{R}^2$ par la relation de la forme

$$\det \begin{pmatrix} u^2 & f^2(u,v,w) \\ f(u,v,w)+v & vf(u,v,w) \end{pmatrix},$$

où u et v sont deux fonctions de x et y.

```
(%i1) depends(f,[u,v,w]);
```

(%o1) $[\mathrm{f}\,(u, v, w)]$

```
(%i2) depends([u,v,w],[x,y]);
```

(%o2) $[\mathrm{u}\,(x, y)\,, \mathrm{v}\,(x, y)\,, \mathrm{w}\,(x, y)]$

```
(%i4) M:matrix([u^2,f^2],[f+v,f*v]);
```

(%o4) $\begin{pmatrix} u^2 & f^2 \\ v+f & f\,v \end{pmatrix}$

```
(%i5) diff(determinant(M),x,1);
```

(%o5) $-f^2\,\left(\left(\frac{d}{d\,w}\,f\right)\,\left(\frac{d}{d\,x}\,w\right)+\left(\frac{d}{d\,v}\,f\right)\,\left(\frac{d}{d\,x}\,v\right)+\frac{d}{d\,x}\,v+\left(\frac{d}{d\,u}\,f\right)\,\left(\frac{d}{d\,x}\,u\right)\right)$
$-2\,f\,(v+f)\,\left(\left(\frac{d}{d\,w}\,f\right)\,\left(\frac{d}{d\,x}\,w\right)+\left(\frac{d}{d\,v}\,f\right)\,\left(\frac{d}{d\,x}\,v\right)+\left(\frac{d}{d\,u}\,f\right)\,\left(\frac{d}{d\,x}\,u\right)\right)$
$+u^2\,v\,\left(\left(\frac{d}{d\,w}\,f\right)\,\left(\frac{d}{d\,x}\,w\right)+\left(\frac{d}{d\,v}\,f\right)\,\left(\frac{d}{d\,x}\,v\right)+\left(\frac{d}{d\,u}\,f\right)\,\left(\frac{d}{d\,x}\,u\right)\right)+f\,u^2\,\left(\frac{d}{d\,x}\,v\right)+2\,f\,u\,\left(\frac{d}{d\,x}\,u\right)\,v$

E. 13. Soit f une fonction différentiable sur $\mathbb{R}^n$ et g définie par

$$g(x_1,\ldots,x_n) = f(0,x_1,x_2,\ldots,x_n) \quad x \in \mathbb{R}^n.$$

Exprimer la différentielle de g en fonction de celle de f.

§ 4. Extremums

Soit $f : \Omega \subset E \to \mathbb{R}$. Nous dirons que $a \in \Omega$ est un **maximum local** (d'autres parlent de **maximum relatif**) s'il existe un voisinage U de a tel $f(x) \geq f(a)$ pour $x \in U$. Lorsque l'inégalité large '$\geq$' peut être remplacée par une inégalité stricte '$>$', nous parlons de **maximum local strict**. La définition de minimum local (resp. local strict) s'obtient de manière similaire en remplaçant $\geq$ par $\leq$ (resp. $>$ par $<$). Ces deux familles de points sont regroupées sous le terme d'**extremum local** (resp. local strict). Naturellement, l'étude des extremuns n'a de sens que pour les fonctions à valeurs réelles.

Théorème 8. *Soit $f : \Omega \subset E \to \mathbb{R}$. Si f admet un extremum local en $a \in \Omega$ et si f est différentiable en a alors $df(a) = 0$, autrement dit $df(a) \in \mathcal{L}(E,\mathbb{R})$ est l'application linéaire nulle.* ■

Démonstration. C'est un conséquence directe du critère connu pour les fonctions d'une variable réelle. Soit $v \in E$, $v \neq 0$. Considérons la fonction g définie sur un voisinage de 0 dans $\mathbb{R}$ par $g(t) := f(a + tv)$. La fonction f admettant un extremum en a, la fonction g admet un extremum en 0 et le théorème classique des fonctions d'une variable réelle nous assure que $g'(0) = 0$ mais un calcul montre que $g'(0) = df(a)(v)$. Nous avons donc $df(a)(v) = 0$ pour un élément quelconque v de E, autrement dit, $df(a) = 0$. ■

Pour s'assurer que $df(a)$ est l'application linéaire nulle, il suffit de vérifier qu'elle s'annule en tous les vecteurs d'une base quelconque. lorsque $E = \mathbb{R}^n$ et que l'on utilise la base canonique, la condition est équivalente à $\partial_i f(a) = 0$, $i = 1, \ldots, n$. Un point en lequel la différentielle s'annule s'appelle un **point critique**. Le théorème ci-dessus revient simplement à affirmer que *les extremums des fonctions différentiables sont atteints en des points critiques.* Il faut bien réaliser toutefois que la recherche des points critiques est un problème difficile. Dans le cas d'une fonction de $\Omega \subset \mathbb{R}^N \to \mathbb{R}$, il équivaut à la résolution d'un système de n équations (non linéaires) en n inconnues. Il est rarement possible d'obtenir des solutions exactes de tels systèmes et il faut recourir à des méthodes d'approximation. Une de ces méthodes est proposée à l'exercice 29.

Remarquons en outre que le théorème n'offre qu'une condition nécessaire pour que a soit un extremum. cependant, dans la pratique elle permet de réduire très notablement la grandeur de l'ensemble dans lequel il faut rechercher un maximum. Nous verrons plus loin une condition suffisante.

Dans la plupart des problèmes pratiques, y compris pour les fonctions d'une variable réelle, nous devons trouver des extremums globaux (l'inégalité doit avoir lieu non seulement sur un voisinage du point mais sur tout l'ensemble d'étude de la fonction). Là encore, un extremum global étant aussi local – dès lors que l'ensemble d'étude est ouvert–, les outils du calcul différentiel fournissent une aide considérable dans la solution des problèmes de recherche d'extremums. Lorsque, comme il plus courant, l'ensemble d'étude est compact, la théorie permet souvent de progresser vers la solution. Par exemple, si nous pouvons montrer que la différentielle ne s'annule jamais dans l'intérieur du compact alors les bornes de la fonctions sont à rechercher sur la frontière. Si cette frontière se laisse décrire par une équation du type $g(x) = 0$, régulière dans un sens à préciser, d'autres techniques issues du calcul différentiel permettent de considérablement réduire l'ensemble des points où peut se trouver un extremum. Une introduction à ces techniques sera donnée dans la partie 3.

§ 5. Théorème des accroissements finis

5.1 Le théorème élémentaire des accroissements finis

Si f est une fonction *réelle* continue sur un intervalle $[a, b]$ dérivable sur l'intérieur $]a, b[$ de cet intervalle alors il existe $c \in]a, b[$ tel que $f(b) - f(a) = (b - a)f'(c)$. Ce théorème est une forme plus générale (et une conséquence directe) du théorème de Rolle. Géométriquement, il signifie qu'il existe un point c à l'intérieur de l'intervalle en lequel la tangente au graphe de f est parallèle à la droite passant par les points $(a, f(a))$ et $(b, f(b))$. Un énoncé aussi précis ne peut pas être espéré dans le cas des fonctions à valeurs dans un espace de dimension supérieure. Le contre-exemple classique est celui de la fonction $f : x \in \mathbb{R} \to (\cos x, \sin x) \in \mathbb{R}^2$. Nous avons $f'(x) = (-\sin x, \cos x)$ qui ne s'annule jamais de sorte que l'égalité $f(0) - f(2\pi) = 2\pi f'(c)$ n'est satisfaite pour aucune valeur de c. Le résultat plus faible que nous pourrons étendre au cas des fonctions à valeurs dans un espace de dimension supérieure est celui de l'inégalité des accroissements finis qui dit, dans le cas classique, que sous les hypothèses précédentes et si, en outre, f' est bornée par M sur $]a, b[$ alors $|f(b) - f(a)| \leq M(b - a)$.

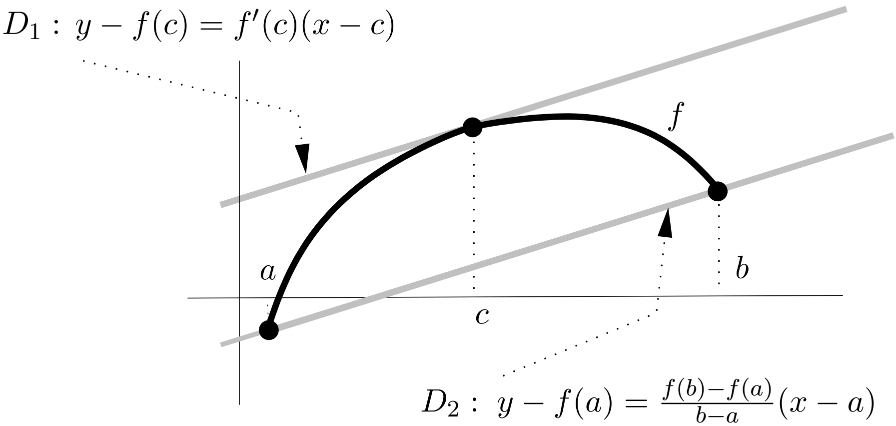


Figure 6 – Le théorème des accroissements finis classique ($D_1 \parallel D_2$ ou $f'(c) = (f(b) - f(a))/(b - a)$).

5.2 Enoncé du théorème

Si a et b sont deux éléments de E, le **segment** $[a, b]$ est par définition l'ensemble

$$[a, b] := \{ta + (1 - t)b : t \in [0, 1]\}. \tag{5.1}$$

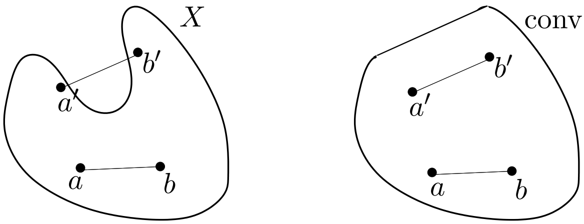


Figure 7 – Enveloppe convexe

Une modification immédiate de cette définition conduit à celles de $]a, b[$, $[a, b|$, etc. Rappelons qu'un sous-ensemble X de E est dit **convexe** lorsque il contient le segment $[a, b]$ chaque fois qu'il contient la paire $\{a, b\}$. Les boules sont des convexes. Les segments eux-mêmes sont aussi des ensembles convexes.

Dans la figure 7, l'ensemble X sur la gauche n'est pas convexe car le segment $[a', b']$ n'est pas inclus dans X. L'ensemble $\operatorname{conv}(X)$ sur la droite est convexe. C'est l'**enveloppe convexe** de X, c'est-à-dire le plus petit ensemble convexe contenant X.

Théorème 9 (des accroissements finis ou de la moyenne). *Soient $a, b \in \Omega \subset E$. Supposons que le segment $[a, b]$ soit contenu dans Ω et que la fonction $f : \Omega \to F$ soit différentiable en tout point de $]a, b[$. Si de plus, $\|df(x)\| \leq M$ pour tout $x \in]a, b[$ alors*

$$\|f(b) - f(a)\|_F \leq M\|b - a\|_E. \tag{5.2}$$

■

Démonstration. Nous démontrons le théorème dans le cas simple (mais courant) où la norme de F est donnée par un produit scalaire, $\|x\|_F = \sqrt{\langle x, x\rangle}$. Considérons alors la fonction g définie pour $t \in [0, 1]$ par

$$g(t) = \langle f(a + t(b - a), v\rangle, \quad t \in [0, 1],$$

où v est un élément quelconque de F. Puisque $a+t(b-a) = ta+(1-t)b$, l'hypothèse $[a, b] \subset \Omega$ assure que g est correctement définie sur $[0, 1]$. Elle est aussi dérivable sur $]0, 1[$ et sa dérivée est donnée par

$$g'(t) = \langle df(a)(b - a), v\rangle.$$

Pour obtenir cette relation il suffit de décomposer la fonction g comme

$$t \to f(a + t(b - a)) \to \langle f(a + t(b - a), v\rangle$$

et de calculer $g'(t) = dg(t)(1)$ par le théorème des fonctions composées. Ensuite une application du théorème des accroissement finis des fonctions d'une variable réelle donne l'existence de $c \in]0,1[$ (dépendant de a,b, f et v) tel que $g(1) - g(0) = g'(c)$ ou encore

$$\langle f(b) - f(a), v\rangle = \langle df(c)(b-a), v\rangle.$$

Choisissant $v = f(b) - f(a)$, nous obtenons

$$\|f(b) - f(a)\|_F^2 = \langle df(c)(b-a), f(b) - f(a)\rangle.$$

En utilisant l'inégalité de Cauchy-Schwarz pour le produit scalaire, nous arrivons à

$$\|f(b) - f(a)\|_F^2 = \|df(c)(b-a)\|_F \|f(b) - f(a)\|_F \leq M\|b-a\|_F \|f(b) - f(a)\|_F.$$

L'inégalité recherchée s'obtient finalement en divisant des deux côtés par $\|f(b) - f(a)\|$ étant entendu que lorsque $f(b) = f(a)$ la formule est évidemment satisfaite. ■

5.3 *Caractérisation des fonctions de différentielle nulle*

Théorème 10. *Si f une application différentiable sur un ouvert convexe U dont la différentielle est nulle en tout point de U alors f est constante sur U.* ■

Démonstration. Soit $a \in \Omega$. Nous montrons que $f = f(a)$ sur Ω. En effet prenant $x \in \Omega$, nous appliquer le théorème des accroissements finis à a et x avec $M = 0$ pour obtenir $f(x) = f(a)$. ■

Ce théorème est vraie sous l'hypothèse beaucoup plus faible que U soit connexe. Un ouvert U de E est dit **connexe** lorsque il n'est pas égal à la réunion de deux ensembles ouverts disjoints.

E. 14. Soit f une fonction différentiable sur $\mathbb{R}^n$, $n > 1$. On suppose que $\partial_1 f{=}0$ sur $\mathbb{R}^n$. Montrer que f est indépendante de x_1. Autrement dit, il existe une fonction différentiable g définie sur $\mathbb{R}^{n-1}$ telle que $f(x_1, x_2, \ldots, x_n) = g(x_2, \ldots, x_n)$ pour tout $x \in \mathbb{R}^n$.

5.4 *Obtention de la différentiabilité à partir des dérivées partielles*

Dans la plupart des cas que l'on rencontre dans la pratique la différentiabilité se déduit de l'existence de dérivées partielles suffisamment régulières. Cela résulte des théorèmes suivants qui ne s'appuient que sur le théorème élémentaire des accroissements finis.

Théorème 11. *Soient $f : U \subset \mathbb{R}^n \to \mathbb{R}$ et $a \in U$. Si f admet des dérivées partielles ($\partial_i f(x)$, $1 \leq i \leq n$) dans un voisinages de a et si ces dérivées partielles sont continues au point a alors f est différentiable en a.* ■

Dans ce cas, conformément à la relation (2.5), la différentielle de f en a est donnée par $df(a)(h) = \sum_{j=1}^{n} \partial_j f(a) h_j$.

Démonstration. Nous démontrons la théorème dans le cas $n = 2$ (le principe de la démonstration dans le cas général est identique) et nous choisissons la norme $\|(x_1, x_2) = \max\{|x_1|, |x_2|\}$. Le candidat $df(a)$ est $(h_1, h_2) \to \partial_1 f(a) h_1 + \partial_2 f(a) h_2$ et il s'agit d'établir que la fonction d'erreur R définie par

$$R(h) = f(a+h) - f(a) - (\partial_1 f(a) h_1 + \partial_2 f(a) h_2)$$

satisfait $\lim_{h\to 0} |R(h)|/\|h\| = 0$. Remarquons d'abord que

$$\begin{aligned} f(a+h) - f(a) & \\ &= f(a+h) - f(a+(h_1,0)) + f(a+(h_1,0)) - f(a) \\ &= h_2 \partial_2 f\big(a + (h_1, \theta_2(h_2)h_2)\big) + h_1 \partial_1 f\big(a + (\theta_1(h_1)h_1, 0)\big) \end{aligned} \tag{5.3}$$

où nous avons appliqué le théorème des accroissements finis des fonctions d'une variable, d'une part à la fonction $g_2(t) := t \to f(a + (h_1, t))$ pour estimer $g_2(h_2) - g_2(0)$ et d'autre part à la fonction $g_1(t) := t \to f(a + (t, 0))$ pour estimer $g_1(h_1) - g(0)$. Les seules informations dont nous disposons sur $\theta_1(h_1)$ et $\theta_2(h_1)$ c'est qu'ils sont tous deux compris entre 0 et 1. En employant l'expression obtenue dans la formule pour $R(h)$, nous obtenons

$$\begin{aligned} R(h) = h_2 \left\{\partial_2 f\big(a + (h_1, \theta_2(h_2)h_2)\big) - \partial_2 f(a)\right\} & \\ + h_1 \left\{\partial_1 f\big(a + (\theta_1(h_1)h_1, 0)\big) - \partial_1 f(a)\right\} & \end{aligned}$$

Nous en déduisons immédiatement que

$$\frac{|R(h)|}{\|h\|} \leq |\partial_2 f\big(a + (h_1, \theta_2(h_2)h_2)\big) - \partial_2 f(a)| + |\partial_1 f\big(a + (\theta_1(h_1)h_1, 0)\big) - \partial_1 f(a)|,$$

et le majorant tend vers 0 lorsque $h \to 0$ puisque $(h_1, \theta_2(h_2)h_2) \to 0$ et $(\theta_1(h_1)h_1, 0) \to 0$ et que, par hypothèse, les dérivées partielles sont continues au point a. ∎

E. 15. Quel est l'identité algébrique que nous devons substituer à (5.3) pour démontrer le théorème précédent dans le cas où n est quelconque ?

Théorème 12. *Soit $f = (f_1, \dots, f_m) : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$. Pour que f soit continûment différentiable sur Ω – c'est-à-dire $f \in \mathsf{C}^1(\Omega)$ – il faut et il suffit que les n dérivées partielles de ses m composantes, $\partial_j f_i$, (existent et) soient continues sur Ω.* ■

5.5 Suites de fonctions différentiables

Théorème 13. *Soit Ω un ouvert convexe de E et F un espace de Banach. Si $f_n = \Omega \subset E \to F$, $n \in \mathbb{N}$, est une suite de fonctions différentiables sur Ω vérifiant les deux hypothèses suivantes :*

(i) il existe $a \in \Omega$ telle que la suite $f_n(a)$ soit convergente,

(ii) la suite des applications $df_n : \Omega \to \mathcal{L}(E\ F)$ converge vers $A : x \in \Omega \to A(x) \in \mathcal{L}(E\ F)$ uniformément sur Ω ;

alors nous pouvons conclure que :

(C1) Il existe une fonction $f : \Omega \to F$ telle que pour tout $x \in \Omega$, $f_n(x) \to f(x)$ $(n \to \infty)$ et la convergence est uniforme sur une boule de centre x ;

(C2) la fonction limite f est différentiable sur Ω et $df(x) = A(x)$, $x \in \Omega$.

■

Démonstration. Notons $M_n := \sup_{x\in\Omega} \|df_n(x) - A(x)\|$. L'hypothèse (ii) nous dit que $M_n \to 0$ lorsque $n \to \infty$. Remarquons aussi que

$$\sup_{x\in\Omega} \|df_m(x) - df_n(x)\| \leq M_m + M_n.$$

La première étape consiste à exhiber la fonction limite f. Nous fixons x, $x \neq a$, et montrons que la suite $(f_n(x))$ converge ; sa limite définira la valeur $f(x)$.

Pour établir la convergence de $(f_n(x))$ nous utilisons le critère de Cauchy (c'est possible puisque l'espace d'arrivée F est complet). Fixons donc $\epsilon > 0$, nous cherchons $n_0 \in \mathbb{N}$ tel que $m, n > n_0$ entraine $\|f_m(x) - f_n(x)\|_F \leq \epsilon$.

En faisant intervenir le point a et en appliquant le théorème des accroissements finis à la fonction $f_m - f_n$, nous avons

$$\begin{aligned}\|f_m(x)-f_n(x)\|_F &\leq \|(f_m(x)-f_n(x))-(f_m(a)-f_n(a))\|_F+\|(f_m(a)-f_n(a))\|_F\\ &\leq (M_m + M_n)\|x-a\|_E + \|(f_m(a) - f_n(a))\|_F. \qquad (5.4)\end{aligned}$$

Puisque $M_n \to 0$, nous pouvons trouver n_0 tel que $n > n_1$ entraine $M_n \leq \epsilon/(4\|x - a\|_E$ et puisque $f_n(a) \to f(a)$, nous pouvons trouver n_2 tel que $m < n < n_2$ entraine $\|(f_m(a) - f_n(a))\|_F \leq \epsilon/2$ et, par suite, pour $m > n > n_0 := \max\{n_1, n_2\}$ nous avons $\|f_m(x) - f_n(x)\|_F \leq \epsilon$ ce qui achève la démonstration de l'existence de la fonction limite f.

Nous pouvons maintenant réécrire la relation (5.4) ci-dessus en remplaçant a par un point x_0 quelconque dans Ω. En faisant $m \to \infty$ dans cette nouvelle inégalité, il vient

$$\|f(x) - f_n(x)\|_F \leq M_n\|x - x_0\|_E + \|(f(x_0) - f_n(x_0))\|_F.$$

En particulier, si $B(x_0, r) \subset \Omega$ alors

$$\sup_{x\in B(x_0,r)} \|f(x) - f_n(x)\|_F \leq rM_n + \|(f(x_0) - f_n(x_0))\|_F \to 0 \quad (n \to \infty),$$

qui montre que (f_n) converge uniformément vers f sur un boule de centre x_0 lequel est un point quelconque de Ω (ce qui implique d'ailleurs que f est continue sur Ω).

Il reste à établir que f est différentiable en tout point x_0 de Ω et que sa différentielle en x_0 est $A(x_0)$. Posons $R(h) = f(x_0+h) - f(x_0) - A(x_0)(h)$. Nous voulons établir que $\|R(h)\|_F/\|h\|_E \to 0$ pour $h \to 0$. Nous commençons par établir une majoration intermédiaire. Le théorème des accroissements finis nous assure que

$$\|f_m(x_0 + h) - f_m(x_0) - f_n(x_0 + h) + f_n(x_0)\|_F \leq \|h\|_E(M_m + M_n),$$

et, en faisant $m \to \infty$,

$$\|f(x_0 + h) - f(x_0) - f_n(x_0 + h) + f_n(x_0)\|_F \leq \|h\|_E M_n.$$

Fixons alors $\epsilon > 0$ et choisissons n_0 de telle sorte que $M_{n_0} \leq \epsilon/3$. Nous avons

$$\begin{aligned} R(h) = f(x_0 + h) - f(x_0) - f_{n_0}(x_0 + h) + f_{n_0}(x_0) \\ - (A(x_0) - df_{n_0}(x_0))(h) \\ + f_{n_0}(x_0 + h) - f_{n_0}(x_0) - df_{n_0}(x_0)(h), \end{aligned} \tag{5.5}$$

et chacun des trois termes sur la gauche peut être estimé (en norme $\|\cdot\|_F$). Chacun des deux premiers est dominé par $\epsilon/3 \times \|h\|$ à cause du choix de n_0 et le troisième est encore dominé par $\epsilon/3 \times \|h\|$ pour $\|h\|_E$ assez petit par différentiabilité de f_{n_0} en x_0. Au total, pour $\|h\|_E$ assez petit, $\|R(h)\|_F \leq \|h\|_E\epsilon$ et cela achève la preuve de la différentiabilité de f en x. ■

E. 16. Montrer que dans l'énoncé ci-dessus, on peut remplacer l'hypothèse (ii) par

(ii bis) la suite des applications $df_n : \Omega \to \mathcal{L}(E\ F)$ est convergente vers $A : x \in \Omega \to A(x) \in \mathcal{L}(E\ F)$ et pour tout $x \in \Omega$ la convergence est uniforme sur une boule de centre x.

§ 6. Exercices et problèmes

17. Retrouver la différentielle de la forme quadratique de l'exemple 2 en écrivant Q comme la composée de la forme bilinéaire B et l'application $x \to (x, x)$.

18. *D'autres règles de calcul.*

I) Donner la différentielle d'un produit de deux fonctions $f, g : U \subset \mathbb{R}^n \to \mathbb{R}$ différentiables en a. Que dire de la différentiabilité des fonctions polynômes ?

II) Donner la dérivée d'une fonction de la forme $f \circ g$ où f est une fonction réelle dérivable de la variable réelle et g est une fonction différentiable de $U \subset E$ vers $\mathbb{R}$. Traiter le cas particulier des fonctions $1/g$ et $\log g$. Que dire de la différentiabilité des fractions rationnelles ?

III) Donner la différentielle d'une fonction de la forme $f \circ A$ où $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ et $A = \mathbb{R}^m \to \mathbb{R}^d$ est une application linéaire ou affine. Cas particulier des fonctions *ridge* (pour lesquelles $d = 1$). Voir l'exerice 27.

19. **Identité d'Euler**. Soit $h : \mathbb{R}^n \to \mathbb{R}$ un polynôme homogène de degré k. Cela signifie que pour tout $\lambda \in \mathbb{R}$ et $x \in \mathbb{R}^n$, on a $h(\lambda x) = \lambda^k h(x)$.

I) Démontrer que

$$\sum_{i=1}^{n} x_i \partial_i h(x) = kh(x). \tag{6.1}$$

II) En déduire que

$$|h(x)| \leq 1/\sqrt{k}\ \|x\|_2 \|\nabla h(x)\|_2, x \in \mathbb{R}^n,$$

où $\|\cdot\|_2$ désigne la norme euclidienne usuelle de $\mathbb{R}^n$ et $\nabla(x) = (\partial_1 h(x), \cdots, \partial_n h(x))$. Pour plus d'information sur ∇h voir l'exercice 26.

20. Soit $d \in \mathbb{N}^\star$. On définit la fonction f sur $\mathbb{R}^2$ à valeur dans $\mathbb{R}$ par $f(x,t) = (x-t)_+^d$ où

$$(x-t)_+^d = \begin{cases} (x-t)^d & \text{si } x \geq t \\ 0 & \text{si } x < t \end{cases}.$$

Etudier la différentiabilité de f.

21. Soit $\mathrm{GL}_n(\mathbb{R})$ le groupe des matrices inversibles de $M_n(\mathbb{R})$. Montrer que $\mathrm{GL}_n(\mathbb{R})$ est un ouvert de $M_n(\mathbb{R})$ et étudier la différentiabilité de l'application d'inversion $A \to A^{-1}$. Pour $A \in \mathrm{GL}_n(\mathbb{R})$ et $H \in M_n(\mathbb{R})$ de norme assez petite, on pourra calculer $(I + HA^{-1})(I - HA^{-1})$.

22. Soient Ω un ouvert de $M_n(\mathbb{R})$ et $f = \Omega \to M_n(\mathbb{R})$ une fonction différentiable. On définit g sur $M_n(\mathbb{R})$ par la relation

$$g(X) = X \cdot f(X),$$

($g(X)$ est le produit de la matrice X par la matrice $f(X)$).

(i) Montrer que g est différentiable sur Ω et exprimer $dg(X)$ en fonction de $df(X)$, $X \in \Omega$.

(ii) On suppose maintenant que Ω est l'ouvert formé des matrices inversibles et on pose $f(X) = X^{-1}$. Retrouver la différentielle de $df(X)$ à partir de la relation $X \cdot f(X) = Id$, $X \in \Omega$.

23. Soient E un espace vectoriel normé et B une forme bilinéaire symétrique sur $E \times E$. On suppose que $f : \mathbb{R} \to E$ une fonction dérivable sur $\mathbb{R}$ à valeurs dans E vérifiant

$$B(f'(t), f(t)) = 0, \quad t \in \mathbb{R}.$$

Montrer que la fonction g définie sur $\mathbb{R}$ pat $g(t) = B(f(t), f(t))$ est une fonction constante.

24. Montrer que la différentielle de l'application déterminant (det) sur $M_n(\mathbb{R})$ est donnée par la relation

$$d(\det)(A)(H) = \operatorname{Tr}\left({}^t\operatorname{comat}(A)\ H\right), \tag{6.2}$$

où $\operatorname{comat}(M)$ désigne la co-matrice (matrice des cofacteurs) de M et $\operatorname{Tr}(M)$ désigne la trace de la matrice M. On pourra utiliser le fait que $\det$ est un application multilinéaire des collonnes de la matrice pour établir d'abord que

$$d(\det)(A)(H) = \sum_{i=1}^{n} \det(A_i(H)), \tag{6.3}$$

où $A_i(H)$ est la matrice obtenue en substituant la i-ème colonne de H à la i-ème colonne de A. Ensuite, on pourra établir l'identité demandée en commençant par $A = I$ puis A inversible puis conclure par un argument de densité.

25. ($\leftarrow$ 24) Montrer que l'ensemble $\mathrm{GL}_n^+(\mathbb{R})$ formé des matrices de déterminant strictement positif est un ouvert de $M_n(\mathbb{R})$ et donner la différentielle de l'application $A \to \log\det A$.

26. Soit H un espace de Hilbert, Ω un ouvert de H et f une fonction différentiable sur Ω à valeurs dans $\mathbb{R}$. Pour tout $a \in \Omega$, $df(a)$ est une forme linéaire sur H continue. D'après le théorème de représentation de Riesz, il existe $u = u_a \in H$ tel que $df(a)(h) = \langle u_a, h\rangle$, $h \in H$ où $\langle\cdot,\cdot\rangle$ représente le produit scalaire de H. Cet élement u_a s'appelle de gradient de f en a et est noté $\nabla f(a)$. Donner des formules pour $\nabla(fg)$, $\nabla(f/g)$, $\nabla(f\circ g)$.

27. On note par $\langle x,\, y\rangle$ le produit scalaire ordinaire de deux vecteurs x et y de $\mathbb{R}^n$. Soient $\lambda \in \mathbb{R}^n$, $\lambda \neq 0$, et $f : \mathbb{R} \to \mathbb{R}$ une fonction réelle dérivable sur $\mathbb{R}$ tout entier. On définit F sur $\mathbb{R}^n$ par la relation $F(x) = f(\langle\lambda,\, x\rangle)$. Rappelons que ces fonctions sont appelées **fonctions ridge** ou parfois dans les textes anglo-saxons **plane waves**, une terminologie qui se comprend aisément à la vue du graphe donné à la figure 27. Montrer que F est différentiable en tout point de $\mathbb{R}^n$ et calculer une expression de sa différentielle en $x_0 \in \mathbb{R}^n$ en fonction de f et de λ. Le fonction F peut-elle admettre un extremum local strict en un point x_0 de $\mathbb{R}^n$?

28. E et F désignent des espaces vectoriels de dimension finie, Ω un ouvert de E et a un point de Ω. Si $f : \Omega \to F$ est une fonction différentiable en a on pose

$$P(f,a)(x) = f(a) + df(a)(x-a), \quad x \in E.$$

L'objet du problème est l'étude de quelques propriétés de la fonction $P(f,a)$.

I) Montrer que si f et g sont différentiables en a et $\lambda \in \mathbb{R}$ alors $P(f+\lambda g, a) = P(f,a) + \lambda P(g,a)$.

II) (a) Montrer que si f est une application linéaire alors $P(f,a) = f$.

(b) On suppose maintenant $F = \mathbb{R}$ et que $f(x) = B(x,x)$ où B est une forme bilinéaire symétrique (continue) sur $E \times E$. Montrer que $P(f,a)(x) = B(a,x)$, $x \in E$.

III) On suppose $F = E$, $f : \Omega \to E$ différentiable en $a \in \Omega$ et $g : U \to E$ différentiable en $f(a) \in f(\Omega) \subset U$.

(a) Montrer que $P(g\circ f, a) = P(g, f(a)) \circ P(f,a)$.

(b) On suppose maintenant que A est une application linéaire bijective de E dans E et que $g \circ A = g$. Montrer que $P(g, A(a)) = P(g,a) \circ A^{-1}$.

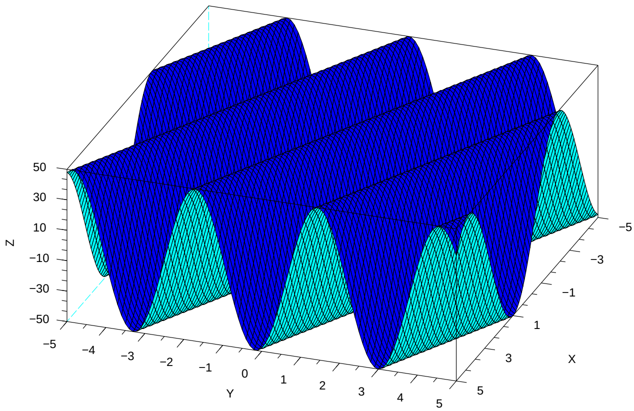


Figure 8 – Graphe de la fonction *ridge* $(x, y) \in \mathbb{R}^2 \to z = 50 \sin(x + 2y) \in \mathbb{R}$.

IV) Dans cette partie, on suppose que $F = \mathbb{R}$ et que f et g sont deux fonctions de Ω dans $\mathbb{R}$ différentiables en a. Montrer que

$$P(fg, a) = f(a)P(g, a) + g(a)P(f, a) - (fg)(a).$$

29 (Un algorithme d'optimisation). On dit qu'une fonction continue f de $\mathbb{R}^n$ dans $\mathbb{R}$ est **coercive** lorsque $\lim_{\|x\|\to\infty} f(x) = \infty$. L'espace $\mathbb{R}^n$ est muni de sa norme euclidienne associée au produit scalaire $\langle\cdot,\cdot\rangle$.

I) Montrer que toute fonction coercive admet un minimum sur $\mathbb{R}^n$.

II) On suppose maintenant que f est de classe C^1 sur $\mathbb{R}$ et que $df(a) \neq 0$. On note $\nabla f(a)$ l'unique vecteur de $\mathbb{R}^n$ tel que $df(a)(h) = \langle\nabla f(a), h\rangle$. (Quelles sont les coordonnées de $\nabla f(a)$?) Montrer que si $t > 0$ et assez petit alors $f(a - t\nabla f(a)) < f(a)$.

III) On suppose toujours que f est coercive et de classe C^1 sur $\mathbb{R}^n$. Considérons l'algorithme suivant pour déterminer un point critique de f :

(a) On prend $x_0 \in \mathbb{R}$ quelconque.

(b) Ayant construit $x_0, \ldots, x_k$
  - Si $df(x_k) = 0$ l'algorithme s'arrête : x_k est point critique.
  - Si $df(x_k) \neq 0$ alors on détermine (par un autre algorithme) $t_k \in]0, \infty[$ tel que

$$f(x_k - t_k\nabla f(x_k)) = \inf_{t>0} f(x_k - t\nabla f(x_k)),$$

  puis on pose $x_{k+1} = x_k - t_k\nabla f(x_k)$.

IV) Montrer que, sous les hypothèses de l'algorithme, t_k existe.

V) En utilisant le fait que $g'(t_k) = 0$ où g est la fonction définie sur $\mathbb{R}^+$ par $g(t) = f(x_k - t\nabla f(x_k))$ montrer que les vecteurs $\nabla f(x_{k+1})$ et $\nabla f(x_k)$ sont orthogonaux.

VI) Montrer que la suite (x_k) est bornée. On utilisera le fait que la suite $(f(x_k))$ est décroissante et le fait que f est coercive.

Théorème 14. *On suppose que l'algorithme ci-dessus ne s'arrête pas. Toute sous-suite convergente de (x_k) converge vers un point critique de f.* ■

VII) Démontrer le théorème. Supposant que X est limite d'une sous-suite de (x_k) avec $df(X) \neq 0$, on recherchera une contradiction.

30. Soit U un ouvert comme sur la figure 9. On suppose que f est différentiable sur U et que $\|df(x)\|$ est borné par M pour $x \in U$. Trouver une constante $d(a,b)$ telle que $\|f(b)-f(a)\| \leq M\, d(a,b)$. Donner une généralisation de ce résultat.

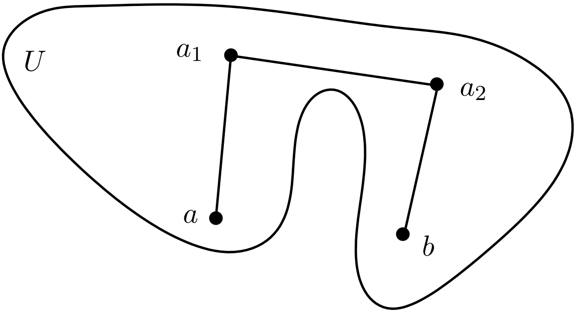


Figure 9 –

31. Soit $M_n(\mathbb{R})$ l'espace vectoriel des matrices carrées d'ordre n, $n \geq 2$, à coefficients réels que l'on munit d'une norme notée $\|\cdot\|$. Nous supposons que cette norme est multiplicative, cela signifie que pour A et B dans $M_n(\mathbb{R})$, on a $\|A \cdot B\| \leq \|A\|\,\|B\|$.

I) On définit l'application b sur $M_n(\mathbb{R}) \times M_n(\mathbb{R})$ par la relation $b(X,Y) = X \cdot Y$, autrement dit $b(X,Y)$ est le produit matriciel de X par Y. L'espace $M_n(\mathbb{R}) \times M_n(\mathbb{R})$ est muni de la norme produit : $\|(A,B)\| = \max(\|A\|, \|B\|)$.

(a) Montrer que b est une application bilinéaire sur $M_n(\mathbb{R}) \times M_n(\mathbb{R})$. Est-elle symétrique ?

(b) Montrer que b est différentiable en $(A,B) \in M_n(\mathbb{R}) \times M_n(\mathbb{R})$ et déterminer sa différentielle en (A,B).

(c) En déduire la différentielle en A de l'application $P_2 : X \in M_n(\mathbb{R}) \to X^2 \in M_n(\mathbb{R})$. On utilisera le fait que $P_2 = b \circ \Phi$ avec $\Phi : X \in M_n(\mathbb{R}) \to (X,X) \in M_n(\mathbb{R}) \times M_n(\mathbb{R})$.

II) Soit $k \geq 2$. On définit l'application $P_k : X \in M_n(\mathbb{R}) \to X^k \in M_n(\mathbb{R})$.

(a) Montrer, par récurrence sur k que si $A \in M_n(\mathbb{R})$ et $H \in M_n(\mathbb{R})$ alors

$$dP_k(A)(H) = \sum_{i=1}^{k} A^{i-1} H A^{k-i} \tag{6.4}$$

$$= HA^{k-1} + AHA^{k-2} + A^2HA^{k-3} + \cdots + A^{k-2}HA + A^{k-1}H. \tag{6.5}$$

(b) Trouver une majoration dépendant de k et de $\|A\|$ pour la norme de l'application linéaire $dP_k(A)$.

(c) Montrer que pour tout $k \geq 1$ et toutes matrices X, Y dans $M_n(\mathbb{R})$ on a

$$\|X^k - Y^k\| \leq k \max(\|X\|, \|Y\|)^{k-1} \|X - Y\|.$$

III) On note $S_k = Id + \sum_{i=1}^{k} P_i$. L'application S_k est donc une application de $M_n(\mathbb{R})$ dans $M_n(\mathbb{R})$.

(a) Soit $R \in]0, 1|$. Montrer que la suite S_k est uniformément convergente sur la boule $B_f(0, R) = \{X \in M_n(\mathbb{R}) \ : \ \|X\| \leq R\}$.

(b) Soit ϕ la limite de la suite S_k. Montrer que ϕ est différentiable sur la boule ouverte de centre 0 et de rayon 1 et déterminer la différentielle de ϕ.

(c) Calculer $S_k(X)(Id - X)$ et en déduire que pour toute matrice X avec $\|X\| < 1$ on a $\phi(X) = (Id - X)^{-1}$.

(d) Montrer que pour toutes matrices X et Y de normes < 1, on a

$$\|(Id - X)^{-1} - (Id - Y)^{-1}\| \leq \frac{2\|X - Y\|}{\Big(1 - \max(\|X\|, \|Y\|)\Big)^2}.$$

32. *Exponentielle matricielle.*

Soit $M_n(\mathbb{R})$ l'espace vectoriel des matrices carrées d'ordre n, $n \geq 2$, à coefficients réels que l'on munit d'une norme multiplicative notée $\| \cdot \|$ (voir le problème 31).

I) Montrer que la série de terme général $X^k/k!$, $k \geq 0$, est normalement convergente sur tout compact de $M_n(\mathbb{R})$. On note $\exp(X)$ sa limite, ie

$$\exp X = \sum_{k=0}^{+\infty} \frac{1}{k!} X^k.$$

II) Montrer que $\exp$ est de classe C^1 sur $M_n(\mathbb{R})$.

III) Soit $X \in M_n(\mathbb{R})$. Montrer que l'application $t \in \mathbb{R} \to \exp(tX) \in M_n(\mathbb{R})$ est dérivable et que

$$\frac{d}{dt} \exp(tX) = X \exp(tX).$$

En déduire que pour $A \in M_n(\mathbb{R})$ et $x \in \mathbb{R}^n$ donnés, l'application $\Phi : \mathbb{R} \to \mathbb{R}^n$ définie par : $\Phi(t) = \exp(tA)(x)$, vérifie

$$\Phi'(t) = A(\Phi(t)) \quad \text{et} \quad \Phi(0) = x.$$

IV) Soit $(e_1, \ldots, e_n)$ une base de $\mathbb{R}^n$. On pose pour $k = 1 \ldots, n$,

$$e_k(t) = \exp(tA)(e_k) \quad \text{et} \quad J(t) := \det(e_1(t), \ldots, e_n(t)).$$

Montrer que

$$J'(t) = \mathrm{Tr}(A) J(t).$$

Que peut on dire dans le cas où $\mathrm{Tr}(A) = 0$.

V) Soit $A : t \to A(t)$ une application continue de $\mathbb{R}$ dans $M_n(\mathbb{R})$. On suppose que la fonction $t \in \mathbb{R} \mapsto U(t) \in M_n(\mathbb{R})$ satisfait la relation $U' = A \cdot U$. Montrer que si $u = \det \circ U$ et $a := \operatorname{Tr} \circ A$ alors $u' = a\, u$.

33. Soient E un espace vectoriel normé de dimension finie dont la norme est notée $\| \cdot \|$ et $B : E \times E \to \mathbb{R}$ une forme bilinéaire symétrique continue et définie positive dont la norme est notée L de sorte que $|B(x,y)| \leq L\|x\|\,\|y\|$ pour tous $x, y \in E$. Rappelons que B définie positive signifie $B(x,x) > 0$ pour tout $x \neq 0$.

La résultat de la première question est utilisable dans la suite mais on peut traiter directement les questions suivantes sans avoir résolu la première.

I) Soit $\lambda : E \to \mathbb{R}$ et $f = E \to E$. On définit F sur E par la relation

$$F(x) = \lambda(x) f(x),$$

autrement dit, $F(x)$ est la multiplication du vecteur $f(x)$ par le réel $\lambda(x)$.
Montrer que si λ et f sont différentiables en $x \in E$ alors F est différentiable en x et exprimer $dF(x)$ à l'aide des différentielles de λ et f. (On pourra remarquer que l'application $(t, y) \in \mathbb{R} \times E \to ty \in E$ est bilinéaire.)

II) Dans cette partie on étudie la différentiabilité de l'application F définie sur $F(x) = \frac{1}{B(x,x)} x$.

(a) Quel est l'ouvert le plus grand sur lequel F est définie ? Cet ouvert sera noté Ω.

(b) Montrer que l'application $x \in \Omega \to \frac{1}{B(x,x)} \in \mathbb{R}$ est différentiable et déterminer ses différentielles.

(c) En déduire une expression pour les différentielles de F.

(d) Montrer que pour tout $x \in \Omega$ l'application $dF(x)$ est un automorphisme de E (une application linéaire bijective de E dans lui-même).

III) On pose $M(x,y) = \inf_{t \in [0,1]} B(tx + (1-t)y, tx + (1-t)y)$

(a) A quelle condition sur x et y a-t-on $M(x,y) > 0$?

(b) On suppose que U est un sous-ensemble de Ω tel que $M(x,y) > 0$ pour tous $x, y \in U$. Déterminer une constante K dépendant de x, y, L et $M(x,y)$ telle que

$$\|F(x) - F(y)\| \leq K\|x - y\|.$$

(c) Lorsque $M(x,y) > 0$, calculer $M(x,y)$ en fonction de $B(x,x)$, $B(y,y)$ et $B(x,y)$.

34. Norme des différentielles. Nous étudions les relations entre les limites des taux d'accroissement et les normes des différentielles. Dans cet exercice nous supposons que les espaces E et F sont de dimensions finies.

(A) Démontrer le théorème suivant.

Théorème 15. *Soit $f : \Omega \subset E \to F$ une fonction différentiable au point $a \in \Omega$. Nous avons*

$$\|df(a)\| = \limsup_{x \to a} \frac{\|f(x) - f(a)\|_F}{\|x - a\|_E}.$$

■

Rappelons qu'une limite supérieure peut se définir comme suit

$$\limsup_{x\to a} g(x) = \lim_{t\to 0} \sup_{x\in B(a,t)} g(x),$$

la limite sur la droite existant certainement (dans $\mathbb{R}\cup\{\infty\}$) puisque la fonction en t sur la droite est décroissante. Nous avons noté $B(a,t) = \{x \in E : \|x-a\| \leq t\}$. Une limite inférieure se définit de manière similaire en remplaçant le $\sup$ par un $\inf$ dans la relation ci-dessus.

(B) Montrer que

$$\liminf_{x\to a} \frac{\|f(x)-f(a)\|_F}{\|x-a\|_E} = \inf_{\|u\|_E=1} \|df(a)(u)\|_F.$$

Montrer que lorsque $df(a)$ est un isomorphisme alors le terme de droite est égal à $\|[df(a)]^{-1}\|^{-1}$. Lorsque $df(a)$ n'est pas un isomorphisme, que vaut le terme de droite ?

2

Différentielles secondes et supérieures

§ 1. Définition

1.1 Différentier l'application différentielle

Nous avons déjà remarqué que lorsque $f : \Omega \subset E \to F$ est différentiable en tous points d'un sous-ensemble ouvert U de Ω alors il est légitime d'étudier la continuité de l'application $df : a \in U \subset E \to df(a) \in \mathcal{L}(E, F)$. En réalité, nous pouvons non seulement étudier la continuité, mais aussi la différentiabilité. Lorsque cette application df est différentiable au point $a \in U$, nous pouvons construire sa différentielle $d(df)(a) = ddf(a)$ qui sera un élément de $\mathcal{L}(E;\, \mathcal{L}(E;\, F))$. D'après le théorème 6, à l'élément $ddf(a) \in \mathcal{L}(E;\, \mathcal{L}(E;\, F))$, correspond une unique application bilinéaire continue, c'est-à-dire une unique élément de $\mathcal{L}(E, E; F))$ que nous noterons $d^2(f)(a)$. En fait,

$$d^2(f)(a)(h,\, k) = [ddf(a)(h)](k), \quad h,\, k \in E. \tag{1.1}$$

L'application bilinéaire $d^2 f(a)$ est appelée la **différentielle seconde** de f en a. Lorsqu'elle existe, nous disons que la fonction f est deux fois différentiable en a. Le fait que $d^2 f(a)$ existe sous-entend donc que f est différentiable dans un voisinage de a.

E. 35 (Exemples de calcul de différentielles secondes).

I) Montrer que la différentielle seconde d'une application linéaire est nulle.

II) Déterminer la différentielle seconde d'une forme quadratique.

III) Montrer que la composée de deux fonctions deux fois différentiables en a est aussi différentiable en a et déterminer une expression de $d^2(f \circ g)$ en fonctions des différentielles de f et g.

1.2 Double dérivation suivant deux vecteurs

Si df est définie sur U, il en est de même de $D_v f : x \to D_v f(x) \in F$, $v \in E$, puisque, d'après le théorème 6, $D_v f(x) = df(x)(v)$. Il est alors légitime d'envisager le calcul de $D_w(D_v f)(x)$.

Théorème 1. *Si $f : \Omega \subset E \to F$ est deux fois différentiable en $a \in \Omega$ alors $D_w(D_v f)(a)$ existe et*

$$D_w(D_v f)(a) = d^2 f(a)(w, v), \qquad v, w \in E. \tag{1.2}$$

■

Théorème 2. *Lorsque $f : \Omega \subset E \to F$ est deux fois différentiable en $a \in \Omega$ alors pour tous vecteurs non nuls h et k dans E*

$$d^2 f(a)(k, h) = \lim_{t \to 0} \frac{1}{t^2} \left\{ f(a + t(h + k)) - f(a + th) - f(a + tk) + f(a) \right\}. \tag{1.3}$$

■

Théorème 3. *Lorsque $f : \Omega \subset E \to F$ est deux fois différentiable en $a \in \Omega$ alors l'application bilinéaire $d^2 f(a) \in \mathcal{L}(E, E; F)$ est symétrique, autrement dit*

$$d^2 f(a)(h, k) = d^2 f(a)(k, h), \qquad h, k \in E. \tag{1.4}$$

■

1.3 Dérivées partielles secondes

Supposons, comme dans 2.3, que E soit de dimension finie n et que $(v_1, v_2, \ldots, v_n)$ soit une base de E. Si $f = \Omega \subset E \to F$ est deux fois différentiable en a alors d'après le théorème 1 pour $h = \sum_{i=1}^n h_i v_i$ et $k = \sum_{i=1}^n k_i v_i$

$$\begin{aligned}
d^2 f(a)(h, k) &= D_h(D_k f)(a) && (1.5)\\
&= D_h\left(\sum_{i=1}^n k_i D_{v_i} f\right)(a) && (1.6)\\
&= \sum_{i=1}^n k_i D_h(D_{v_i} f)(a) && (1.7)\\
&= \sum_{j=1}^n \sum_{i=1}^n k_i h_j D_{v_j}(D_{v_i} f)(a). && (1.8)
\end{aligned}$$

Puisque $D_{v_j}(D_{v_i} f)(a) = d^2 f(a)(v_j, v_i) = d^2 f(a)(v_i, v_j) = D_{v_i}(D_{v_j} f)(a)$ l'expression peut être simplifiée lorsque $h = k$ en regroupant les termes identiques et nous obtenons

$$d^2f(a)(h,h) = \sum_{i=1}^{n} h_i^2 D_{v_i}(D_{v_i}f)(a) + 2\sum_{i<j} h_i h_j D_{v_i}(D_{v_j}f)(a). \tag{1.9}$$

Lorsque $E = \mathbb{R}^n$ est muni de base canonique v_i et $F = \mathbb{R}$, l'expression devient

$$d^2f(a)(h,h) = \sum_{i=1}^{n} h_i^2 \partial_{ii}f(a) + 2\sum_{i<j} h_j h_i \partial_{ij}f(a), \tag{1.10}$$

où ∂_{ij} désigne la dérivée partielles secondes ordinaire. La notation suivante est souvent rencontrée

$$\partial_{ii}^2 f(a) = \frac{\partial^2 f}{\partial x_i^2}(a), \quad \partial_{ij}^2 f(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a). \tag{1.11}$$

La matrice $H = H(f,a)$ dont les coefficients sont les éléments $\partial_{ij}f(a)$ s'appelle la **matrice hessienne** de f au point a ou simplement la **hessienne** de f en a,

$$H_{ij} = \partial_{ij}f(a), \quad 1 \leq i,j \leq n.$$

Si f est deux fois différentiables en a la matrice $H(f,a)$ est symétrique * et nous avons

$$d^2f(a)(h,h) = \begin{pmatrix} h_1 & \dots & h_n \end{pmatrix} H(f,a) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}, \quad h \in \mathbb{R}^n. \tag{1.12}$$

Exemple 1. La figure 1.3 montre un calcul de Hessienne. Noter l'emploi de la fonction *bfloat* pour avoir une valeur numérique approchée.

1.4 Fonctions de classe C^2

Une application $f : \Omega \subset E \to F$, deux fois différentiable sur un sous-ensemble ouvert U de Ω est dite de classe C^2 sur U et nous écrivons $f \in \mathsf{C}^2(U)$ lorsque l'application $d^2f : x \in U \to d^2f(x) \in \mathcal{L}(E,E;F)$ est continue. Pour la topologie de $\mathcal{L}(E,E;F)$, nous renvoyons à la section 5 de l'annexe. L'analogue direct du théorème 12 pour les fonctions de classe C^2 reste valable. Les différentielles d'ordre k, $k \geq 3$ (et les classes de fonctions $\mathsf{C}^k(U)$) se définissent d'une manière similaire à partir des différentielles d'ordre $k-1$.

∗. Il existe d'autres conditions assurant cette symétrie.

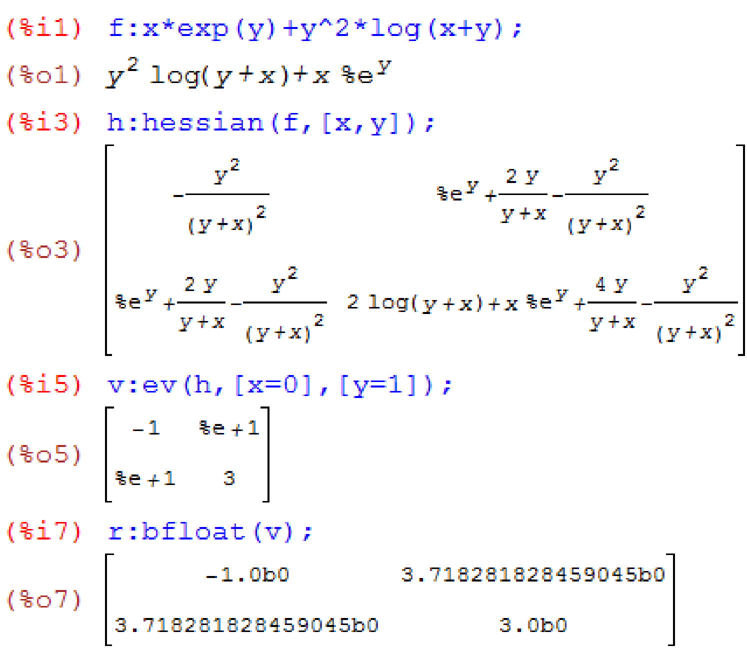


Figure 1 – Un calcul de hessienne avec Maxima.

§ 2. Formule de Taylor à l'ordre deux

Théorème 4. *Soient $f : \Omega \subset E \to F$ et $a \in \Omega$. Si f est deux fois différentiable en a alors il existe un fonction ϵ définie sur un voisinage de 0 dans E et vérifiant $\lim_{x\to a} \epsilon(x-a) = 0$ telle que*

$$f(x) = f(a) + df(a)(x-a) + \frac{1}{2}\, d^2 f(a)(x-a\,,\, x-a) + \|x-a\|_E^2\, \epsilon(x-a). \quad (2.1)$$

■

L'expression

$$\mathbf{T}_a^2(f)(x) = f(a) + df(a)(x-a) + \frac{1}{2}\, d^2 f(a)(x-a\,,\, x-a) \qquad (2.2)$$

s'appelle le **polynôme de Taylor** de f au point a à l'ordre 2. Nous utiliserons aussi la notation $\mathbf{T}_a^2(f;x)$.

Lorsque $E = \mathbb{R}^n$ et $F = \mathbb{R}$, il s'agit effectivement d'un polynôme ordinaire de n variables réelles. De manière précise, si $x = (x_1, x_2, \dots, x_n)$,

$$\mathbf{T}_a^2(f; x_1, \dots, x_n) = f(a) + \sum_{i=1}^n \partial_i f(a)(x_i - a_i) + \frac{1}{2}\sum_{i=1}^n \partial_{ii} f(a)(x_i - a_i)^2 + \sum_{1\le i<j\le n} \partial_{ij} f(a)(x_i - a_i)(x_j - a_j). \quad (2.3)$$

E. 36. Déterminer le polynôme de Taylor à l'ordre deux à l'origine de la fonction f définie sur $\mathbb{R}^2$ par $f(x, y) = \exp((x + \sin y)$.

Exemple 1. La figure 2 donne un exemple de calcul de polynôme de Taylor à l'ordre 2. On pourrait demander un polynôme de n'importe quel ordre. La troisième paramètre donne le point de la forme

$$a = (\alpha, \alpha, \dots, \alpha)$$

en lequel le polynôme est calculé, ici $\alpha = 0$ et le point a est l'origine. Nous pouvons toujours supposer que le point en question est l'origine. Cela provient de la relation

$$\mathbf{T}_a(f, 2)(x) = \mathbf{T}_0\big(f(\cdot + a), 2\big)(x - a).$$

§ 3. Application à la recherche des extremums

Jusqu'à présent, nous connaissons seulement une condition *nécessaire* pour l'existence d'un extremum local d'une fonction différentiable en un point a. C'est la condition d'annulation de la différentielle donnée par le théorème 8. La formule de Taylor nous permet d'obtenir une condition suffisante très utile dans la pratique.

Théorème 5. *Soient E un espace vectoriel normé de dimension finie et $f : \Omega \subset E \to \mathbb{R}$ deux fois différentiables en $a \in \Omega$ telle que $df(a) = 0$.*

(i) Si $d^2f(a)$ est définie positive alors a admet un minimum local strict en a.

(ii) Si $d^2f(a)$ est définie négative alors f est un maximum local strict en a.

■

Rappelons les définitions,

(i) $d^2f(a)$ est ***définie positive*** si $d^2f(a)(h, h) > 0$, $h \in E$, $h \neq 0$,

(ii) $d^2f(a)$ est ***définie négative*** si $d^2f(a)(h, h) < 0$, $h \in E$, $h \neq 0$,

```
(%i20) g:1/(1+(a*x+b*y+c*z));
```

(%o20) $\dfrac{1}{c\,z+b\,y+a\,x+1}$

```
(%i21) taylor(g,[x,y,z],0,2);
```

(%o21) $1+(-a\,x-b\,y-c\,z)+$

$(a^2\,x^2+(2\,a\,b\,y+2\,a\,c\,z)\,x+b^2\,y^2+2\,b\,c\,z\,y+c^2\,z^2)+...$

```
(%i23) rg:g^(1/2);
```

(%o23) $\dfrac{1}{\sqrt{c\,z+b\,y+a\,x+1}}$

```
(%i25) taylor(rg,[x,y,z],0,2);
```

(%o25) $1-\dfrac{a\,x+b\,y+c\,z}{2}+$

$\dfrac{3\,a^2\,x^2+(6\,a\,b\,y+6\,a\,c\,z)\,x+3\,b^2\,y^2+6\,b\,c\,z\,y+3\,c^2\,z^2}{8}+...$

Figure 2 – Détermination d'un polynôme de Taylor avec Maxima.

Lorsque les inégalités strictes sont remplacées par des inégalités larges, on dit simplement que f est positive (ou négative).

Insistons sur le fait que la condition du théorème n'est pas une condition nécessaire, mais seulement une condition suffisante comme le montre l'exemple suivant.

Exemple 1. Soient $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x_1, x_2) = x_1^2 + x_2^4$. Il est immédiat que $(0, 0)$ est un minimum (global) de f sur $\mathbb{R}^2$. Pourtant, il est facile de vérifier que $d^2 f(0, 0)(h, h) = h_1^2$ de sorte que $d^2 f(a)$ est positive sans être définie positive.

En fait, en utilisant la même technique de démonstration, il n'est pas difficile d'établir que, naturellement sous réserve que $df(a) = 0$, pour que a soit un maximum local, il est ***nécessaire*** que $d^2 f(h, h) \leq 0$, $h \in E$, et pour qu'il soit un minimum local il est nécessaire que $d^2 f(h, h) \geq 0$, $h \in E$, en particulier si la forme quadratique $h \to d^2 f(h, h)$ prend des valeurs de signes différents, a n'est pas un extremum local. La différence entre cette condition nécessaire et la condition suffisante énoncée dans le théorème est lié à la différence entre une inégalité stricte < 0 ou > 0 et une inégalité large ≤ 0 ou ≥ 0.

§ 4. Exercices et problèmes

E. 37. I) Soit f une fonction différentiable sur un ouvert convexe $\Omega \subset \mathbb{R}^n$ à valeur dans $\mathbb{R}^m$. Soient $a, b \in \Omega$. On considère la fonction g définie sur $[0, 1]$ par la relation $g(t) = f(a + t(b - a))$. Que vaut $g'(t)$? Montrer que si f est de classe C^1 sur Ω, on a

$$f(b) = f(a) + \int_0^1 df(a + t(b - a))(b - a)dt.$$

En déduire que

$$\|f(b) - f(a)\| \leq \max_{x \in [a,b]} |||df(c)||| \cdot \|b - a\|.$$

Est-ce une nouvelle démonstration du théorème des accroissements finis ?

II) On suppose maintenant que f est de classe C^2 sur Ω. Montrer que

$$f(b) = f(a) + df(a)(b - a) + \int_0^1 (1 - t)d^2 f(a + t(b - a))(b - a, b - a)dt.$$

En déduire une majoration pour $\|f(b) - f(a) - df(a)(b - a)\|$.

III) Comparer cette estimation aux valeurs numériques rapportées dans la table 1 .

38 (Polynôme de Taylor d'une fonction ridge).

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction *ridge* définie par $f(x) = g(\langle \lambda, x \rangle)$ où g est une fonction de classe $\mathbb{C}^2$ sur $\mathbb{R}$, $\lambda \in \mathbb{R}^n$ et $\langle \cdot, \cdot \rangle$ désigne le produit scalaire usuel dans $\mathbb{R}^n$. Démontrer que

$$\mathbf{T}_a^2(f; x) = T_{\langle \lambda, a \rangle}^2(g; \langle \lambda, x \rangle),$$

où le polynôme de Taylor dans le terme de droite est un polynôme d'une seule variable.

39. Soient u et v deux fonctions deux fois dérivables sur $\mathbb{R}$. On pose $f(x,y) = u(x) - v(y)$. (a) Déterminer en fonction des dérivée de u et de v la forme de la matrice hessienne de f en un point (x_0, y_0) en lequel $df(x_0, y_0)$ est nulle. La matrice hessienne est la matrice de l forme quadratique $h \to d^2f(x_0, y_0)(h, h)$. (b) Soient maintenant $P(x) = ax^2 + bx + c$ et $Q(x) = \alpha x^2 + \beta x + \gamma$ deux trinômes du second degré. Donner une condition suffisante portant sur les coefficients de P et Q pour que $f(x,y) = e^{P(x)} - e^{Q(y)}$ admette un minimum local strict.

40. Soient α et β deux paramètres réels. On définit $f : \mathbb{R}^2 \to \mathbb{R}$ par

$$f(x,y) = x^3/6 + \alpha x^2/2 + y^3/6 + \beta y^2/2.$$

Déterminer en fonction de α et β les extremums locaux de f.

41. Soient E un espace vectoriel normé de dimension finie et $B : E \times E$ une forme bilinéaire symétrique (continue) définie positive (on a donc $B(y,y) > 0$ pour tout vecteur y non nul). La forme quadratique associée à B est notée Q de sorte que $Q(y) = B(y,y)$. Soient a, b deux éléments distincts de E. Ces vecteurs servent de paramètres dans la fonction f définie sur E par $f(x) = Q(x-a) + Q(x-b)$, $x \in E$.

I) Montrer que f est différentiable en tout point de E et déterminer sa différentielle.

II) Recherche d'un minimum.

(a) Montrer qu'il existe un unique point $x \in E$ tels que $df(x) = 0$. Ce point sera noté x_0.

(b) Montrer que pour tout $h \in E$, $f(x_0 + h) = 2B(h,h) + f(x_0)$. Que peut-on en déduire pour x_0 ?

III)

42. Soit E un espace de Banach, $B : E \times E \to \mathbb{R}$ une forme bilinéaire symétrique continue et $L : E \to \mathbb{R}$ une forme linéaire continue. On considère l'application f définie sur E à valeurs dans $\mathbb{R}$ par la relation

$$f(x) = \frac{1}{2}B(x,x) - L(x).$$

I) Démontrer, en détaillant le raisonnement et le calcul, que si f admet un extremum local en $a \in E$ alors

$$\forall h \in E \quad B(a,h) = L(h). \tag{4.1}$$

II) On suppose maintenant que la forme bilinéaire B est positive. Cela signifie que $B(h,h) \geq 0$ pour tout $h \in E$. On suppose en outre que *la condition (4.1) est satisfaite.* (i) Montrer, à partir de la relation $B(a-h,\ a-h) \geq 0$, que

$$\forall h \in E \quad -B(a,a) \leq B(h,h) - 2L(h).$$

(ii) Montrer que f admet un minimum global en a, c'est-à-dire $f(a) \leq f(x)$ pour tout $x \in E$.

***

3

Inversion locale et théorème des fonctions implicites

Dans cette partie les espaces vectoriels normés sont supposés de dimension finie.

§ 1. Le théorème d'inversion locale

Nous savons bien que pour qu'un fonction continue f définie sur un intervalle I dans $\mathbb{R}$ définisse une bijection de I sur son image $f(I)$ il faut et il suffit qu'elle soit strictement monotone. Une condition aussi simple ne saurait exister pour les fonctions de plusieurs variables puisque la notion de fonction monotone n'a plus de sens. La dérivabilité de la fonction f ci-dessus toutefois n'entraîne pas nécessairement celle de sa fonction réciproque. L'exemple classique est celui de la fonction $x \mapsto x^3$ qui est une bijection strictement croissante de $\mathbb{R}$ de $\mathbb{R}$ mais dont la fonction réciproque n'est pas dérivable à l'origine. Pour que la fonction réciproque soit dérivable au point b il est nécessaire et suffisant que $f'(a) \neq 0$ ou $f(a) = b$ et, dans ce cas, $(f^{-1})'(b) = 1/f'(a)$. En particulier, si f est une fonction continûment dérivable sur un voisinage de $a \in I$ et $f'(a) \neq 0$ alors, pour x assez proche de a, disons $x \in I_a$, $f'(x)$ sera non nulle de signe constant. La fonction, par suite, sera strictement monotone et sa fonction réciproque f^{-1} sera dérivable en tout point de I_a. C'est ce résultat qu'il est possible de généraliser au cas des fonctions différentiables sur un ouvert d'un espace vectoriel normé. Observons qu'au contraire du cas des fonctions d'une seule variable, seul le recours au calcul différentiel nous permet de trouver une condition *suffisante* pour garantir qu'une fonction définit une bijection au voisinage d'un point donné.

Théorème 1. *Soit $f : \Omega \subset E \to F$ continûment différentiable sur Ω et $a \in \Omega$. Si $df(a)$ est un isomorphisme de E sur F alors f est un difféomorphisme local au voisinage de a. Cela signifie qu'il existe des voisinages ouverts U de a et V de $b = f(a)$ tel que f soit une* *bijection de* U *sur* V *et* $f^{-1} : V \to U$ *est différentiable sur* V*. En outre*

$$d(f^{-1})(b) = [df(a)]^{-1}. \tag{1.1}$$

■

E. 43. A propos de l'énoncé du théorème : est-il vrai que pour tout $x \in U$, $df(x)$ est inversible ?

E. 44. Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x, y) = (x + y, xy)$. En quels points (x, y), $df(x, y)$ est-il un isomorphisme ?

Exemple 1. Tant que des fonctions simples sont en jeu, le vérification de la condition $df(a)$ inversible ne cause aucune difficulté, voyez la figure 1. Le test s'effectue par le calcul du déterminant du jacobien (qui doit être non nul).

```
(%i2) p:x*y^3+x^2+y^2+x*z+z;
(%o2) x z+z+x y^3+y^2+x^2
(%i3) q:x^3+y^3+z^3;
(%o3) z^3+y^3+x^3
(%i4) r:x*y*z;
(%o4) x y z
(%i6) jac:jacobian([p,q,r],[x,y,z]);
      [ z+y^3+2 x   3 x y^2+2 y   x+1  ]
(%o6) [   3 x^2        3 y^2      3 z^2 ]
      [    y z          x z        x y  ]
(%i8) ev(determinant(jac),[x=1],[y=2],[z=1]);
(%o8) 189
```

Figure 1 – Exemple de la vérification de la condition du théorème d'inversion locale pour la fonction $f(x, y, z) = (x\,z + z + x\,y^3 + y^2 + x^2, z^3 + x^3 + y^3, xyz)$ au point $a = (1, 2, 1)$.

§ 2. Le théorème des fonctions implicites

2.1 La forme locale d'une courbe

Considérons le cercle de rayon 1 et de centre l'origine dans $\mathbb{R}^2$, $\mathcal{C} = \{(x,y) \in \mathbb{R}^2 \ : \ x^2+y^2 = 1\}$. C'est ensemble n'est pas le graphe d'une fonction d'une variable g car, toute droite parallèle à l'axe des y rencontre le graphe d'une fonction en au plus un point et c'est une propriété que ne possède évidemment pas le cercle. Pourtant, localement, $\mathcal{C}$ est bien le graphe d'une fonction. Ici, l'adverbe *localement* signifie que pour tout point $a = (a_1, a_2)$ de $\mathcal{C}$, il existe un voisinage de a, c'est-à-dire un ensemble de la forme $\mathcal{C} \cap B(a,R)$, $R > 0$, qui est le graphe d'une fonction autrement dit, l'équation $x^2+y^2 = 1$ se laisse décrire par une relation $y = g(x)$ ou $x = g(y)$ pour x proche de a_1 et y proche de a_2. Remarquons qu'ici, pour R assez petit, $\mathcal{C} \cap B(a,R)$ est un arc de cercle. Les fonctions g dépendent du point (a,b) considéré, comme il est expliqué sur la figure. 2

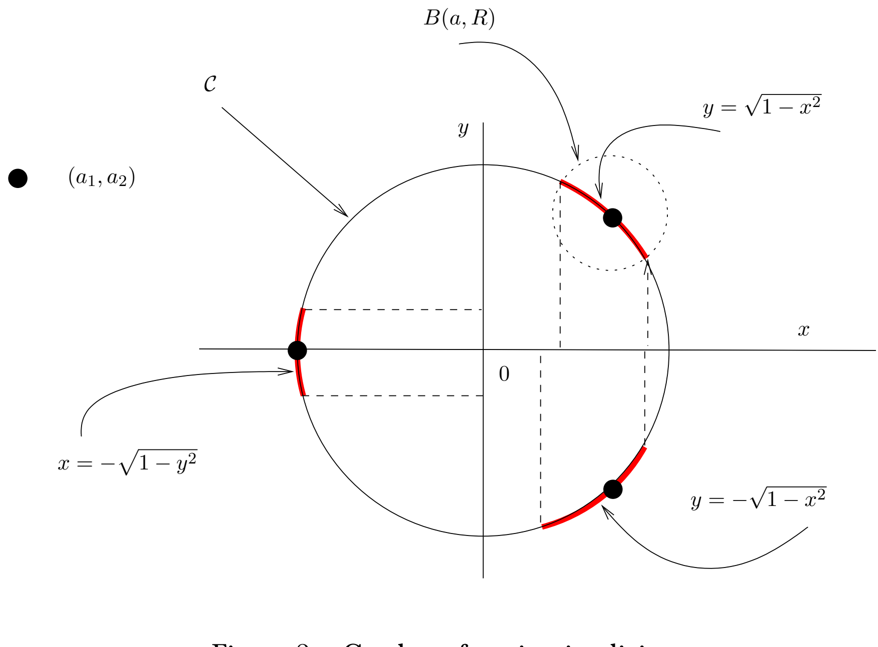


Figure 2 – Cercle et fonction implicite.

Nous voulons résoudre le même problème dans le cas où le cercle $\mathcal{C}$ est remplacé par un ensemble de la forme $\{x \ : \ f(x) = 0\}$ où $f : \Omega \subset \mathbb{R}^{n+1} \to \mathbb{R}$. Ce problème revêt une importance cruciale car beaucoup des ensembles couramment étudiés en géométrie sont de la forme $\{x \ : \ f(x) = 0\}$ où f est une fonction indéfiniment différentiable et même, dans les cas les plus importants, des polynômes. Pour donner une réponse valable dans toutes les dimensions, nous devrons supposer que f est de classe C^1. Comme le montre l'exemple $\mathcal{H}$ de la figure 3 une simple hypothèse de différentiabilité ne peut suffire. À une exception près, au voisinage de tous les points, comme illustré par le point a_r l'ensemble $\mathcal{H}$ se laisse décrire comme le graphe d'une fonction. L'exception est le point a_s de la figure. Il est facile d'expliquer ce résultat négatif. Quel que soit $R > 0$, petit, $\mathcal{H} \cap B(a_s, R)$ est formé de deux 'branches' qui se coupent en a_s et cet ensemble ne saurait être le graphe d'une fonction ; il est en réalité la réunion des graphes de deux fonctions.

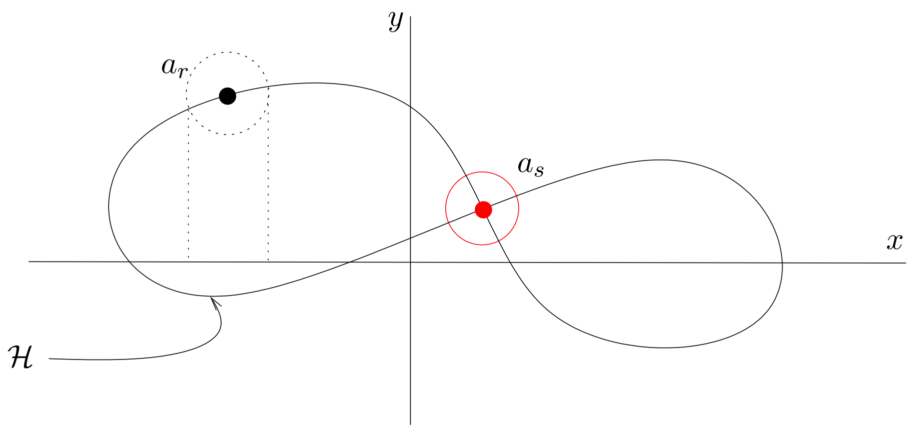


Figure 3 –

2.2 *Le théorème*

Le théorème suivant donne une réponse générale au problème. C'est un résultat fondamental qui sert de pont entre l'analyse et la géométrie.

Théorème 2 (des fonctions implicites). *Soit $f : \Omega \subset \mathbb{R}^{n+1} \to \mathbb{R}$ de classe C^1 sur Ω et*

$$a = (a_1,\, a_2,\, \ldots,\, a_n,\, a_{n+1}) \in \Omega,$$

vérifiant $f(a) = 0$. Alors, si $\partial_{n+1} f(a) \neq 0$, il existe un voisinage ouvert U de a, un voisinage ouvert V de $a' = (a_1,\, a_2, \ldots,\, a_n)$ et une application $g \in \mathsf{C}^1(V)$ tels que

$$\left\{ \begin{array}{l} x = (x_1, \ldots, x_{n+1}) \in U \\ f(x_1,\, x_2,\, \ldots, x_n,\, x_{n+1}) = 0 \end{array} \right\} \iff \left\{ \begin{array}{l} (x_1, \ldots,\, x_n) \in V \\ x_{n+1} = g(x_1,\, x_2,\, \ldots,\, x_n) = 0 \end{array} \right\}. \tag{2.1}$$

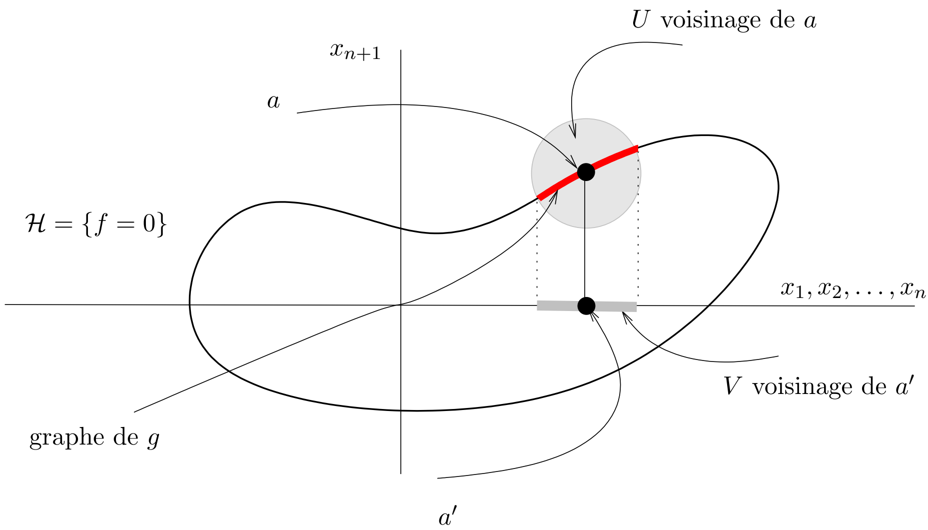


Figure 4 – Théorème des fonctions implicites

De plus, la différentielle de g en $a' = (a_1, \ldots, a_n)$, est donnée par

$$dg(a')(h_1, \ldots, h_n) = -\frac{1}{\partial_{n+1} f(a)} df(a)(h_1, \ldots, h_n, 0). \tag{2.2}$$

En particulier,

$$\partial_i g(a') = -\frac{\partial_i f(a)}{\partial_{n+1} f(a)}, \qquad i = 1, \ldots, n. \tag{2.3}$$

■

Bien entendu, si l'hypothèse $\partial_{n+1} f(a) \neq 0$ est remplacée par $\partial_j f(a) \neq 0$ alors la conclusion est que

$$x_j = g(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n).$$

De manière générale, nous pouvons construire autant de fonctions implicites qu'il y a de dérivées partielles non nulles.

Remarquons encore que la relation (2.2) implique que si f est de classe C^k sur U alors g est de classe C^k sur V. Cette relation permet aussi de calculer les polynômes de Taylor de g.

Les différents objets intervenant dans l'énoncé du théorème des fonctions implicites (a, U, a', V, f, g) et leurs relations sont illustrés dans la figure 4.

2.3 Hypersurfaces

Un ensemble $\mathcal{H} = \{x \in \Omega \ : \ f(x) = 0\}$ où $f : \Omega \subset \mathbb{R}^{n+1} \to \mathbb{R}$ est une fonction dans $\mathsf{C}^1(\Omega)$ s'appelle une hypersurface de $\mathbb{R}^{n+1}$ de classe C^1. Un point $a \in \mathcal{H}$ est dit **régulier** lorsque $df(a) \neq 0$. Nous pouvons alors construire au moins une fonction implicite g et nous définissons l'hyperplan tangent de $\mathcal{H}$ en a comme l'hyperplan tangent au graphe de g en a', comme défini dans la relation 1.6. L'hyperplan tangent ne dépend pas de la fonction implicite choisie, comme le montre le théorème suivant.

Théorème 3. *Soit $\mathcal{H} = \{x \in \Omega \ : \ f(x) = 0\}$ une hypersurface de $\mathbb{R}^{n+1}$ de classe C^1 et a un point régulier de $\mathcal{H}$. L'équation de l'hyperplan tangent à $\mathcal{H}$ en a est donnée par*

$$\partial_1 f(a)(x_1 - a_1) + \partial_2 f(a)(x_2 - a_2) + \cdots + \partial_{n+1} f(a)(x_{n+1} - a_{n+1}) = 0. \qquad (2.4)$$

■

E. 45. Déterminer l'équation du plan tangent en tout point de la sphère unité $\mathcal{S}$ dans $\mathbb{R}^n$, d'équation $\sum_{i=1}^n x_i^2 = 1$.

E. 46. Donner une expression pour un **vecteur normal unitaire** à l'hyperplan tangent à $\mathcal{H}$ en a.

Exemple 1. Il est assez facile de tracer des hypersurfaces de $\mathbb{R}^2$. La code code Maxima permet de tracer une hypersurface et la droite tangente en un point en cette hypersurface. Le résultat apparaît dans la figure 2.3.

```
(%i1) f:x^2-(4*y^3-3*y+1);d1:diff(f,x,1);d2:diff(f,y,1);
(%o1) -4*y^3+3*y+x^2-1
(%o2) 2*x
(%o4) 3-12*y^2
(%i5) a:ev(d1,x=1,y=0); b:ev(d2,x=1,y=0);
(%o5) 2
(%o5) 3
(%i6) load(implicit_plot);
(%o6) C:/PROGRA~1/MAXIMA~1.2/share/maxima/5.19.2/share/contrib
/implicit_plot.lisp
(%i7) implicit_plot ([f=0, a*(x-1)+b*y=0], [x, -4, 4], [y, -4, 4],
 [gnuplot_preamble, "set zeroaxis"]);
(%o7) done
```

Les deux dernières lignes de codes peuvent être substituées par les suivantes

```
(%i9)  load(draw);
(%o9) done
(%i12)  draw2d(grid = true,
             line_type = solid,
             key       = "f=0",
             implicit(f=0, x, -4,4, y, -4,4),
```

```
line_type = dots,
key       = "x^3+y^3 = 3*x*y^2-x-1",
implicit(a*(x-1)+b*y=0, x,-4,4, y,-4,4),
title     = "Two implicit functions" )
```

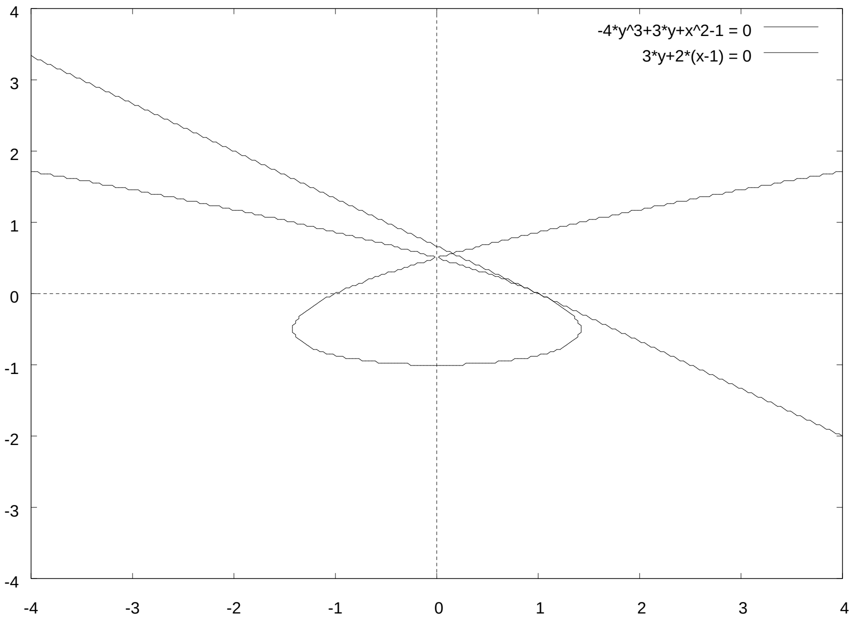


Figure 5 – L'hypersurface $\{x^2 = 4y^3 - 3y + 1\}$ dans $[-4, 4] \times [-4, 4]$ et de sa droite tangente au point $(1, 0)$.

Il est possible de tracer des hypersurfaces dans $\mathbb{R}^3$. Le code suivant produit la figure 6

```
(%i3) f:x^2*y+z^2-15*x*y*z-x; d1:diff(f,x,1);d2:diff(f,y,1);d3:diff(f,z,1);
(%o3) z^2-15*x*y*z+x^2*y-x
(%o4) -15*y*z+2*x*y-1
(%o5) x^2-15*x*z
(%o6) 2*z-15*x*y
(%i7) a:ev(d1,x=1,y=0,z=1);b:ev(d2,x=1,y=0,z=1);c:ev(d3,x=1,y=0,z=1);
(%o7) -1
(%o8) -14
(%o9) 2
(%i10) load(draw);
(%o10) C:/PROGRA~1/MAXIMA~1.2/share/maxima/5.19.2/share/draw/draw.lisp
(%i11) draw3d(color=blue,grid = true,
```

```
implicit(f=0,x,-2,2.5,y,-1,1,z,0,2),
surface_hide=true,color=red, implicit(a*(x-1)+b*y+c*(z-1)=0,
 x,-2,2,y,-2,2,z,-2,2));
(%o11) [gr3d(implicit,implicit)](%o9) [gr3d(implicit,implicit)]
```

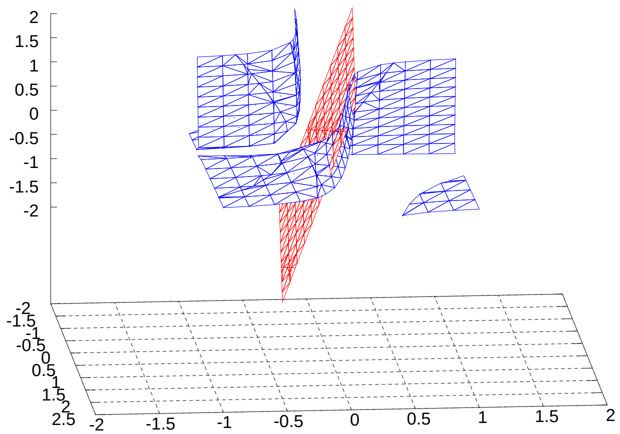


Figure 6 – Intersection de l'hypersurface $\{x^2y+z^2-15xyz-x=0\}$ avec $[-2,2]^3 \subset \mathbb{R}^3$ et tracé de l'hyperplan tangent au point $(1,0,1)$.

§ 3. Problème d'extremums liés

3.1 Hypersurfaces régulières

Soit f une fonction de classe C^1 sur $\Omega \subset \mathbb{R}^n$ à valeurs dans $\mathbb{R}$ et $\mathcal{H}$ l'hypersurface de classe C^1 définie par f, $\mathcal{H} = \{x \in \Omega \; : \; f(x) = 0\}$. Nous disons que $\mathcal{H}$ est régulière si $df(a) \neq 0$ en tout point a de $\mathcal{H}$.

Exemple 1. La sphère $S(a,R) = \{x \in \mathbb{R}^n \; : \; \|x-a\|_2 = R\}$ où $\|\cdot\|_2$ désigne la norme euclidienne est une hypersurface régulière.

E. 47. Soit A une application linéaire de $\mathbb{R}^n$ dans lui-même. On pose $S_A(a,R) = \{x \in \mathbb{R}^n \; : \; \|Ax-a\|_2 = R\}$. $S_A(a,R)$ est-elle une hypersurface régulière ?

Remarquons qu'une hypersurface $\mathcal{H} = \{x \in \Omega \ : \ f(x) = 0\}$ est toujours un sous-ensemble fermé de Ω comme image réciproque d'une fermé par une application continue. Cet ensemble par contre n'est pas nécessairement compact (prendre $f : x \in \mathbb{R}^n \to x_1 \in \mathbb{R}$). Si g est une fonction continue sur $\mathcal{H}$, nous pouvons cependant considérer le problème du calcul de $\sup_{x\in\mathcal{H}} g(x)$ ou $\inf_{x\in\mathcal{H}} g(x)$ en sachant bien que dans le cas où $\mathcal{H}$ est compact, le sup et l'inf sont atteints.

3.2 La condition des multiplicateurs

Le théorème suivant donne une condition *nécessaire* pour que le sup ou l'inf ci-dessus soit atteint en un point donné a. Cette condition permet très souvent de réduire à très peu de valeurs, le lieu où l'extremum est susceptible d'être atteint.

Théorème 4 (du multiplicateur de Lagrange). *Soient $\mathcal{H} = \{x \in \Omega \ : \ f(x) = 0\}$ une hypersurface de classe C^1 régulière et g une fonction différentiable sur un ouvert U contenant $\mathcal{H}$. Si a est un point de $\mathcal{H}$ satisfaisant $g(a) = \sup_{x\in\mathcal{H}} g(x)$ ou bien $g(a) = \inf_{x\in\mathcal{H}} g(x)$ alors les formes linéaires $df(a)$ (qui est non nulle) et $dg(a)$ sont colinéaires. Autrement dit, il existe $\lambda \in \mathbb{R}$ tel que $dg(a) = \lambda df(a)$.* ■

E. 48. Déterminer la plus grande valeur que peut prendre la fonction $f : \mathbb{R}^3 \to \mathbb{R}$ définie par $f(x, y, z) = xyz$ sur la sphère d'équation $x^2 + y^2 + z^2 = 1$. On précisera en quel(s) point(s) le maximum est atteint. Établir un résultat similaire sur la sphère de $\mathbb{R}^n$ d'équation $x_1^2 + x_2^2 + \cdots + x_n^2 = 1$.

§ 4. Exercices et problèmes

49. Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x, y) = (e^x \cos y, e^x \sin y)$.

I) Déterminer l'ensemble image de f.

II) Montrer que f est continûment différentiable en tout point de $\mathbb{R}^2$ et vérifier que pour tout (a_1, a_2), $df(a_1, a_2)$ est inversible de sorte que d'après le théorème d'inversion locale f est difféomorphisme local au voisinage de tous les points de $\mathbb{R}^2$.

III) La fonction f est-elle une bijection de $\mathbb{R}^2$ sur $f(\mathbb{R}^2)$.

50. On considère l'hypersurface de $\mathbb{R}^2$, $\mathcal{H} = \{y^3 + x^2 - 1 = 0\}$.

I) S'agit-il d'un hypersurface régulière ? compacte ?

II) Donner le polynôme de Taylor à l'ordre 2 au point $x = 1$ d'une fonction implicite $y = g(x)$ pour $\mathcal{H}$ au point $a = (1, 0)$.

51. Soient A et B deux matrices dans $M_n(\mathbb{R})$. On suppose que A est inversible et on définit $\mathcal{H} = \{x \in \mathbb{R}^n \ : \ \|Ax\|_2 = 1\}$ où $\|\cdot\|_2$ désigne la norme euclidienne ordinaire.

I) Montrer que $\mathcal{H}$ est une hypersurface régulière compacte.

II) En quels points $a \in \mathbb{R}^n$ peut-on avoir

$$\|Ba\|_2 = \max_{x\in\mathcal{H}} \|Bx\|_2 \quad \text{ou} \quad \|Ba\|_2 = \min_{x\in\mathcal{H}} \|Bx\|_2?$$

III) Résoudre complètement le problème dans le cas ou $n = 2$, $A = Id$ et

$$B = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}.$$

52 (Relations sur les fonctions implicites).

I) Soit f une fonction de classe C^1 sur $\mathbb{R}^2$. On suppose que $f(x_0, y_0) = 0$ et que

(i) $\partial_1 f(x_0, y_0) \neq 0$,

(ii) $\partial_2 f(x_0, y_0) \neq 0$.

D'après le théorème des fonctions implicites, l'hypothèse (i) entraîne l'existence d'une fonction ϕ_1 définie sur un voisinage de y_0 telle que, au voisinage de (x_0, y_0), la condition $f(x, y) = 0$ soit équivalente à $x = \phi_1(y)$. De la même manière l'hypothèse (ii) entraîne l'existence d'une fonction ϕ_2 définie sur un voisinage de x_0 telle que, au voisinage de (x_0, y_0), la condition $f(x, y) = 0$ soit équivalente à $y = \phi_2(x)$. Montrer que $\phi_1'(y_0)\phi_2'(x_0) = 1$.

II) Nous supposons maintenant que f une fonction de classe C^1 sur $\mathbb{R}^3$ et que $f(x_0, y_0, z_0) = 0$ et que (i) $\partial_1 f(x_0, y_0, z_0) \neq 0$, (ii) $\partial_2 f(x_0, y_0, z_0) \neq 0$ et (iii) $\partial_3 f(x_0, y_0, z_0) \neq 0$. Comme au dessus, cette condition permet de définir trois fonctions implicites de classe C^1 qui décrivent localement l'ensemble $f(x, y, z) = 0$, $x = \phi_1(y, z)$, $y = \phi_2(x, z)$ et $z = \phi_3(x, y)$. Montrer que

$$\partial_3\phi_1(y_0, z_0) \cdot \partial_2\phi_3(x_0, y_0) \cdot \partial_1\phi_2(x_0, z_0) = -1.$$

III) Existe-t-il une relation similaire pour une fonction f de n variables, $n \geq 4$?

53. Soit f_m la fonction réelle définie sur $\mathbb{R}^3$ par

$$f_m(x, y, z) = x^2 + 3y^2 + mz^2 - 1, \quad (x, y, z) \in \mathbb{R}^3,$$

où m est un paramètre réel. On définit

$$\mathcal{H}_m = \{(x, y, z) \in \mathbb{R}^3 \ : \ f_m(x, y, z) = 0\}.$$

I) Montrer que $\mathcal{H}_m$ est un sous-ensemble compact de $\mathbb{R}^3$ si et seulement si $m > 0$. (Indication : dans le cas $m \leq 0$, on pourra trouver une suite (x_n, y_n, z_n) telle que $(x_n, y_n, z_n) \in \mathcal{H}_m$ mais $\|(x_n, y_n, z_n)\|_2 \to \infty$, $n \to \infty$.)

II) Donner une condition nécessaire et suffisante sur m pour que $\mathcal{H}_m$ soit une hypersurface régulière.

III) On suppose dans cette partie que $m > 0$. Déterminer M_m défini par

$$M_m = \max_{(x,y,z)\in\mathcal{H}_m} xyz^2.$$

On pourra procéder comme suit :

(a) On montrera d'abord que M_n existe (dans $\mathbb{R}$).

(b) En utilisant le théorème de Lagrange on montrera que si (x, y, z) est un point en lequel le maximum est atteint alors $2x^2 = 6y^2$.

(c) On utilisant le fait que $(x, y, z) \in \mathcal{H}_m$, on en déduira que

$$x^2 = \frac{1}{2}(1 - mz^2) \quad \text{et} \quad y^2 = \frac{1}{6}(1 - mz^2).$$

(d) En utilisant un raisonnement similaire, on déterminera ensuite les valeurs possibles de z puis conclura le calcul de M_m.

54. On munit $\mathbb{R}^n$ de la norme euclidienne habituelle, notée $\|\cdot\|_2$, de sorte que

$$\|x\|_2^2 = \langle x, x\rangle = \sum_{i=1}^{n} x_i^2, \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Soit A une matrice réelle carrée d'ordre n. On définit la fonction f sur $\mathbb{R}^n$ par

$$f(x) = \langle Ax, x\rangle, \quad x \in \mathbb{R}^n.$$

I) Montrer que f est différentiable en tout point $a \in \mathbb{R}^n$ et exprimer $df(a)$ en fonction de A et de A^T. On rappelle que A^T vérifie la relation

$$\langle Ax, y\rangle = \langle x, A^T y\rangle, \quad x, y \in \mathbb{R}^n.$$

II) Montrer en calculant la différentielle de la fonction g définie par

$$g(x) = f(x) - \frac{1}{2}\langle (A + A^T)x, x\rangle, \quad x \in \mathbb{R}^n$$

que

$$f(x) = \frac{1}{2}\langle (A + A^T)x, x\rangle, \quad x \in \mathbb{R}^n.$$

Donner une preuve algébrique directe n'utilisant pas le calcul différentiel.

III) On définit

$$\mathcal{H}_A = \{x \in \mathbb{R}^n \ : \ f(x) = 1\}.$$

(a) Montrer que $\mathcal{H}_A$ est un sous-ensemble fermé de $\mathbb{R}^n$. A quelle(s) condition(s) sur A, cet ensemble est-il non vide.

(b) On suppose que la condition trouvée à la question précédente est satisfaite. Montrer que $\mathcal{H}_A$ est une hypersurface régulière de $\mathbb{R}^n$.

(c) Donner un exemple de matrice A pour laquelle $\mathcal{H}_A$ est un sous-ensemble non vide compact (respectivement, non vide non compact) de $\mathbb{R}^n$.

IV) On suppose maintenant qua la matrice $A + A^T$ est inversible. Soit u un élément non nul de $\mathbb{R}^n$. On cherche à déterminer

$$M(A, u) := \sup_{x \in \mathcal{H}_A} \langle u, x\rangle.$$

(a) Montrer que si $M(a, u)$ est atteint en $a \in \mathcal{H}_A$ alors x est nécessairement colinéaire à $(A + A^T)^{-1}u$. Combien y-a-t-il d'éléments $a \in \mathcal{H}_A$ vérifiant cette condition ?

(b) Déterminer $M(A, u)$ lorsque $n = 3$, $u = (a, b, c)$ et

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

55. On considère l'application $\phi : \mathbb{R}^2 \to \mathbb{R}^3$ définie par $\phi(x) = (\phi_1(x), \phi_2(x), \phi_3(x))$ avec

$$\phi_1(x) = \frac{1 - x_1^2 - x_2^2}{1 + x_1^2 + x_2^2},\ \phi_2(x) = \frac{2x_1}{1 + x_1^2 + x_2^2} \text{ et } \phi_3(x) = \frac{2x_2}{1 + x_1^2 + x_2^2} \quad x = (x_1, x_2). \quad (4.1)$$

I) Montrer que ϕ est différentiable sur $\mathbb{R}^2$ et déterminer le rang de l'application $D\phi(x)$ en tout point x de $\mathbb{R}^2$. (On rappelle que le rang d'une application linéaire est par définition la dimension de son ensemble image.)

II) On désigne par S la sphère de centre $0 = (0, 0, 0)$ et de rayon 1,

$$S = \{y = (y_1, y_2, y_3) \in \mathbb{R}^3 \ : \ y_1^2 + y_2^2 + y_3^2 = 1\} \quad (4.2)$$

et $P = (-1, 0, 0) \in S$. Montrer que pour tout $x = (x_1, x_2) \in \mathbb{R}^2$, $\phi(x)$ est l'intersection de S avec la droite passant par P et le point M_x de coordonnées $(0, x_1, x_2)$ et montrer que ϕ est une bijection de $\mathbb{R}^2$ sur $S \setminus \{P\}$ (la sphère privée du point P).

III) Soit Π la plan d'équation $x_1 = -1$ dans $\mathbb{R}^3$. On considère la fonction $\eta : \mathbb{R}^3 \setminus \Pi \to \mathbb{R}^2$ définie par

$$\eta(y_1, y_2, y_3) = \left(\frac{x_2}{1 + x_1}, \frac{x_3}{1 + x_1}\right). \quad (4.3)$$

(a) Étudier la différentiabilité de η.

(b) Montrer que pour tout $x \in \mathbb{R}^2$, on a $(\eta \circ \phi)(x) = x$.

IV) Trouver une fonction ϕ' et une fonction η' pour lesquelles le rôle précédemment joué par P est joué par $P' = (1, 0, 0)$.

V) Soit $a \in S$, $a \notin \{P, P'\}$, et $x_a = \eta(a)$. Montrer que $f = \eta' \circ \phi$ est un difféomorphisme au voisinage de x_a.

Les fonctions ϕ et ϕ' permettraient de définir sur la sphère S une structure de ***variété différentielle***. Cette structure permet à son tour d'étendre la notion de fonction différentiable à des fonctions définies sur des ensembles plus complexes que les ouverts des espaces $\mathbb{R}^n$; elle joue un rôle fondamental dans les mathématiques modernes.

***

4

Equations différentielles

Dans cette partie les espaces vectoriels normés considérés sont de dimension finie.

§ 1. Les théorèmes de Cauchy-Lipschitz

1.1 Le cas des fonctions lipschitziennes

Théorème 1 (de Cauchy-Lipschitz). *Soit I un intervalle de $\mathbb{R}$, $f : I \times E \to F$, et $y_0 \in F$. Si f est continue sur $I \times E$ et s'il existe $L \in \mathbb{R}^+$ satisfaisant*

$$\|f(t, y) - f(t, z)\|_F \leq L \|y - z\|_E, \quad t \in I,\ y,\ z \in E \tag{1.1}$$

alors l'équation différentielle

$$\begin{cases} y'(t) = f(t, y(t)), & t \in I, \\ y(t_0) = y_0. \end{cases} \tag{1.2}$$

admet une et une seule solution $t \to y(t)$ définie sur I. ■

Cet énoncé est simple mais il faut bien réaliser que les hypothèses sont fortes car la fonction f doit être Lipschitzienne sur l'espace vectoriel E tout entier (uniformément en t). Elle ne s'applique pas par exemple à l'étude de l'équation différentielle $y'(t) = y^2(t)$. Elle s'appliquera toutefois au cas des applications différentielles linéaires que nous allons voir plus bas.

E. 56. Vérifier que l'équation $y'(t) = y^2(t)$ sur $I =]0, +\infty[$ ne satisfait pas les condition du théorème. Trouver cependant une solution définie sur I.

E. 57. On considère une équation différentielle de la forme

$$Y^{(n)}(t) = g(t, Y(t), Y'(t), \ldots, Y^{(n-1)}(t)), \quad t \in I \text{ avec } Y(t_0) = Y_0, \tag{1.3}$$

où I est un intervalle et g est une fonction continue de $I \times \mathbb{R}^n$ à valeurs dans $\mathbb{R}$.

I) Montrer que les solutions de (1.3) sont en bijection avec les solutions d'une équation de la forme

$$y'(t) = f(t,\, y(t)), \quad t \in I \text{ avec } y(t_0) = y_0 \tag{1.4}$$

où $y : t \in I \to y(t) \in \mathbb{R}^n$ et $y_0 \in \mathbb{R}^n$ et $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ sont à déterminer en fonction de g et Y_0. (On pourra poser $y = (y_1, \ldots, y_n)$ avec $y_j(t) = Y^{(j-1)}(t)$.)

II) Donner une condition sur g assurant que l'équation (1.4) satisfait aux conditions du théorème de Cauchy-Lipschitz.

E. 58 (Changement de fonction inconnue).

On considère une équation différentielle de la forme

$$y'(t) = f\big(t,\, A(t)(y(t))\big), \quad t \in I, \tag{1.5}$$

où f est une fonction continue sur $I \times E$ et $t \in I \to A(t) \in M_n(\mathbb{R})$ est une application continue telle que pour tout $t \in I$, la matrice $A(t)$ est inversible. Quelle est l'équation différentielle satisfaite par la fonction $u(t) := A(t)y(t)$ si y est solution de (1.5).

1.2 Le cas des fonctions localement lipschitziennes

Le théorème suivant s'appuie sur des hypothèses de même type mais moins forte que celle du théorème 1 pour une conclusion moins précise. Il permet toutefois d'obtenir des informations sur une classe beaucoup plus grande d'équations différentielles.

Soient I un intervalle de $\mathbb{R}$ contenant t_0, $y_0 \in F$ et $f = I \times \Omega \to F$ une fonction continue. Nous dirons que le couple (J, B) où $J = [x_0 - \alpha,\, x_0 + \alpha]$ est un sous-intervalle de I de centre t_0 et $B = B(y_0,\, R) = \{z \,:\, \|z - y_0\|_E \leq R\}$ une boule fermée dans Ω constituent un **cylindre de sécurité** pour $(t_0,\, y_0,\, f)$ si les conditions suivantes sont satisfaites

(i) f est bornée par une constante M sur $J \times B$,

(ii) $\alpha\, M \leq R$.

Il est toujours possible de trouver un cylindre de sécurité sous la seule hypothèse de la continuité de f. Commençons par prendre une boule fermée $B(y_0, R)$ quelconque dans Ω, et $\alpha_0 > 0$ de sorte que $[x_0 - \alpha_0,\, x_0 + \alpha_0]$ est inclus dans I. Puisque f est continue sur le compact $[x_0 - \alpha_0,\, x_0 + \alpha_0] \times B(y_0, R)$, elle y est bornée par un certain $M > 0$. Puisque $\lim_{\alpha \to 0} \alpha M = 0$, nous pouvons choisir $\alpha < \alpha_0$ assez petit pour que $\alpha M \leq R$. Puisque

$$[x_0 - \alpha,\, x_0 + \alpha] \times B(y_0, R) \subset [x_0 - \alpha_0,\, x_0 + \alpha_0] \times B(y_0, R),$$

la fonction f est bornée par M sur $[x_0 - \alpha,\, x_0 + \alpha] \times B(y_0, R)$ et les deux conditions demandées sont donc vérifiées.

Théorème 2 (de Cauchy-Lipschitz (second)). *Soit $f = I \times \Omega \to F$ une fonction continue et (J, B) un cylindre de sécurité pour (t_0, y_0, f). S'il existe $L \in \mathbb{R}^+$ satisfaisant*

$$\|f(t,\, y) - f(t,\, z)\|_F \leq L\, \|y - z\|_E, \quad t \in J,\ y,\, z \in B \tag{1.6}$$

alors l'équation différentielle

$$\begin{cases} y'(t) = f(t, y(t)), & t \in I, \\ y(t_0) = y_0. \end{cases} \tag{1.7}$$

admet une et une seule solution y définie sur J et prenant ses valeurs dans B ($y(J) \subset B$). ■

§ 2. Equations différentielles linéaires

2.1 Définition

Les équations différentielles pour lesquelles la fonction $f(t, y)$ est de la forme $A(t)(y) + b(t)$, avec $A(t) \in \mathcal{L}(E)$, autrement dit, est une fonction affine en y avec une application linéaire et une constante associée dépendant de t, est appelée équation différentielle linéaire. Cette terminologie est impropre. Il aurait été préférable de la réserver à ce que la tradition a défini comme les équations différentielles linéaires homogènes pour lesquelles $b(t) = 0$ c'est-à-dire $f(t, y) = A(t)(y)$. Nous dirons aussi que cette dernière équation est l'équation sans second membre associée à l'équation différentielle linéaire $y'(t) = A(t)(y)(t) + b(t)$.

Théorème 3. *Soit A une application continue d'un intervalle $I \subset \mathbb{R}$ dans $\mathcal{L}(E)$ et b une application continue de I dans E. Si $(t_0, y_0) \in I \times E$ alors l'équation différentielle*

$$\begin{cases} y'(t) = A(t)(y(t)) + b(t), & t \in I, \\ y(t_0) = y_0. \end{cases} \tag{2.1}$$

admet une et une seule solution sur I. ■

2.2 Le cas particulier des équations homogènes

Soit A une application continue d'un intervalle $I \subset \mathbb{R}$ dans $\mathcal{L}(E)$. L'ensemble des solutions de l'équation différentielle

$$y'(t) = A(t)(y(t)), \quad t \in I, \tag{2.2}$$

est un espace vectoriel que nous noterons $\mathcal{S}$. Nous savons que si $(t_0, y_0) \in I \times E$, l'équation admet une et une seule solution vérifiant $y(t_0) = y_0$. Cette solution est notée $y[t_0, y_0, \cdot] : t \in I \to y[t_0, y_0; t] \in E$. Remarquons de l'ensembles des solutions de l'équation linéaire $y't) = A(t)(y(t)) + b(t)$ est complètement déterminé lorsque sont connues l'ensemble des solutions de l'équation sans second membre et une solution particulière y_p. Cela provient du fait que si y est solution de l'équation avec second membre alors $y - y_p$ est solution de l'équation homogène.

Théorème 4. *Avec les notations précédentes. Pour $s \in I$, nous avons :*

(i) L'application $v \in E \to y[s, v, \cdot]$ *est une application linéaire bijective entre* E *et* $\mathcal{S}$*,*

(ii) En particulier $\mathcal{S}$ *est un espace vectoriel dont la dimension est égale à celle de* E*.*

■

Voici quelques conséquences importantes de ce théorème.

(i) Si y est une solution de (2.2) qui s'annule en un point $s \in I$ – ie $y(s)$ est le vecteur nul de E – alors y est nécessairement la fonction nulle. Autrement dit, ou bien une solution ne s'annule en aucun point ou bien elle s'annule partout.

(ii) Supposons que $\dim E = n$. Si $f_1, f_2, \ldots, f_n$ sont n éléments de $\mathcal{S}$ alors pour montrer qu'ils forment une base de $\mathcal{S}$, il suffit de trouver $t \in I$ tel que les vecteurs $f_i(t)$, $i = 1, \ldots, n$ soient linéairement indépendants.

§ 3. Exercices et problèmes

59 (Résolvante d'un système linéaire homogène).

Soit I un intervalle de $\mathbb{R}$ et A une fonction continue de I dans $\mathcal{L}(E)$. Nous considérons l'équation

$$y'(t) = A(t)(y(t)), \quad t \in I.$$

L'espace vectoriel des solutions est noté $\mathcal{S}$. Pour tout $s \in I$, nous définissons ϕ_s de E dans $\mathcal{S}$ par la relation

$$\phi_s(v) = y[s, v, \cdot],$$

autrement dit, $\phi_s(v)$ est l'unique solution qui prend la valeur v au point s. Cette application ϕ_s est un isomorphisme linéaire de E sur $\mathcal{S}$. Si nous remplaçons s par un autre élément r de I, nous obtenons un autre isomorphisme ϕ_r. Il suit que $\phi_s^{-1} \circ \phi_r$ est un isomorphisme de E dans lui-même. Cet isomorphisme est noté $R(s, r)$. Cette définition est illustrée dans la figure 1. L'application $(r, s) \to R(r, s) \in \mathcal{L}(E)$ s'appelle la résolvante de l'équation différentielle $y'(t) = A(t)(y(t))$. L'objet de cet exercice est d'étudier quelques propriétés de la résolvante.

I) Propriétés algèbriques.

(a) $R(s, s) = Id$.

(b) $R(s_1, s_2) \circ R(s_2, s_3) = R(s_1, s_3)$ et $R(r, s) = (R(s, r))^{-1}$.

(c) Montrer que si $f \in \mathcal{S}$ alors $f(s) = R(s, r)(f(r))$.

(d) En déduire que pour tout $v \in E$ la fonction $t \in I \to R(t, r)(v)$ appartient à $\mathcal{S}$ puis que

$$\frac{d}{dt} R(t, s) = A(t) R(t, s). \tag{3.1}$$

(e) Nous supposons maintenant que $\dim E = n$ et que $f_1, f_2, \ldots, f_n$ forment une base de l'espace des solutions $\mathcal{S}$. Nous fixons une base $e = (e_i)$ de E. Montrer que

$$\mathrm{mat}_e(R(s, r)) = V(s) V(r)^{-1},$$

où $V(r)$ est la matrice dont la i-ème colonne est formée des coordonnées de $f_i(r)$ dans la base e. Notons que cette propriétés signifie que connaître la résolvante est équivalent à connaître une base de $\mathcal{S}$.

$$\begin{tikzcd}
E \arrow[dr, "\phi_r"] \arrow[dd, "R(s,r)"'] & \\
& \mathcal{S} \\
E \arrow[ur, "\phi_s"'] &
\end{tikzcd}
\qquad \phi_r = \phi_s \circ R(s,r)$$

Figure 1 –

II) Application aux équations différentielles linéaires non homogènes : méthode de variation de la constante. Nous avons dit que pour résoudre l'équation différentielle linéaire

$$y'(t) = A(t)(y(t)) + b(t)$$

il suffit de résoudre l'équation homogène et de connaître une solution particulière de l'équation. Nous allons voir que la connaissance de l'ensemble des solutions de l'équation homogène, donc de la résolvante de cette équation permet d'obtenir une solution particulière. Nous cherchons une solution particulière y de la forme

$$y(t) = R(t,s)(z(t))$$

et nous devons chercher $z(t)$.

(a) Montrer que y est solution si et seulement si $z'(t) = R(s,t)(b(t))$.

(b) En déduire que

$$z(t) = z(s) + \int_s^t R(s,\theta)((b(\theta))d\theta.$$

(c) Montrer que la fonction y définie par la relation ci-dessous est l'unique solution de l'équation non homogène satisfaisant $y(t_0) = y_0$,

$$y(t) = R(t,t_0)(y_0) + \int_{t_0}^t R(t,\theta)(b(\theta))d\theta.$$

60. On note $\mathbf{GL}_n(\mathbb{R})$ l'ensemble des matrices réelles inversibles d'ordre n. Si $X : t \in]a,b[\to GL_n(\mathbb{R})$ est dérivable, on définit $\mathcal{Q}(X)$ sur $]a,b[$ par

$$\mathcal{Q}(X)(t) = X'(t)X^{-1}(t), \quad t \in]a,b[.$$

I) Soit $A \in M_n(\mathbb{R})$. Expliciter la fonction $\mathcal{Q}(X)$ lorsque

(a) $X(t) = e^{tA}, \quad t \in \mathbb{R}$.

(b) $X(t) = e^{\ln(t-a)A}, \quad t \in]a, +\infty[$. Quelle est l'équation différentielle satisfaite par X ?

II) Soient X_i, $i = 1, 2$, deux fonctions dérivables de $]a,b[$ dans $\mathbf{GL}_n(\mathbb{R})$. Montrer que

$$\mathcal{Q}(X_1X_2) = \mathcal{Q}(X_1) + X_1\mathcal{Q}(X_2)X_1^{-1}.$$

III) Soit X une fonction dérivable de $]a,b[$ dans $\mathbb{R}$. Montrer que

$$\mathcal{Q}(X^{-1}) = -\left(\mathcal{Q}(X^T)\right)^T.$$

On justifiera soigneusement chacune des étapes du calcul. On rappelle que la fonction $\Delta : M \in \mathbf{GL}_n(\mathbb{R}) \to M^{-1}$ est différentiable en tout point M de $\mathbf{GL}_n(\mathbb{R})$ et $d\Delta(M)(H) = -M^{-1}HM^{-1}$.

61 (Une équation différentielle d'ordre 2, [1]). On considère une équation différentielle de la forme $y'' + 2py' + (p^2 + p')y = 0$ où p est une fonction dérivable sur $\mathbb{R}$. On note L l'opérateur définie sur $C^\infty(\mathbb{R})$ par $L(y) = y'' + 2py' + (p^2 + p')y$. Montrer que L se factorise sous la forme $L = l \circ l$. Utiliser cette observation pour résoudre l'équation lorsque $p(x) = \tanh x$.

62 (Le théorème de Floquet). On établit une propriété sur les équations différentielles linéaires dont les coefficients sont des fonctions périodiques.

On pourrait penser que les solutions d'une équation différentielle linéaire à coefficients 2π-périodique sont elles-mêmes 2π-périodique, comme c'est le cas par exemple pour l'équation $y' + y\cos x = 0$ dont les solutions sont $y = c\exp(-\sin x)$ mais la considération d'un exemple à peine moins simple comme $y' + y\cos^2 x = 0$ montre que la conjecture est erronée puisque les solutions, non périodiques, sont données par

$$y = c\,\exp\left(-\frac{\sin(2\,x)}{4} - \frac{x}{2}\right).$$

Le théorème de Floquet que nous établissons dans cet exercice explicite le défaut de périodicité des solutions des équations différentielles linéaires à coefficients périodiques.

On considère une équation différentielle linéaire

$$y' = A(t)y$$

d'inconnue $y : \mathbb{R} \to \mathbb{C}^n$ et où la fonction $A : \ t \in \mathbb{R} \to A(t) \in M_n(\mathbb{C})$ est continue 2π-périodique. On notera que y prend ses valeurs dans le plan complexe.

I) On se propose de montrer le théorème suivant.

Théorème 5. *Notons $(y_1,\dots,y_n)$ un base de l'espace vectoriel des solutions de l'équation différentielle ci-dessus. On note $S(t)$ la matrice carrée d'ordre n dont les colonnes sont les coordonnées des $y_i(t)$ dans la base canonique de $\mathbb{C}^n$. Il existe un matrice $B \in M_n(\mathbb{C})$ et une fonction $F : t \in \mathbb{R} \to M_n(\mathbb{C})$ continument dérivable et 2π-périodique telle que $S(t) = F(t)\cdot e^{tB}$, $t \in \mathbb{R}$.* ■

Le théorème de Floquet indique donc qu'à un facteur exponentiel près, les solutions d'une équation différentielle linéaire à coefficients périodiques sont encore périodiques. On s'assurera que cet énoncé est compatible avec les exemples simples considérés plus haut.
On rappelle que l'exponentielle de matrice

$$\exp : M \in M_n(\mathbb{C}) \to e^M = \sum_{j=0}^{\infty} \frac{1}{j!} M^j,$$

définit un difféomorphisme d'un voisinage de la matrice nulle sur un voisinage de la matrice identité *. Nous utiliserons le fait que l'ensemble image de $\exp$ est formé de toutes les matrices inversibles.

II) Comment se transforme la matrice $S(t)$ lorsqu'on remplace $(y_1,\dots,y_n)$ par une autre base de $\mathcal{S}$.

III) On considère la matrice $S(t+2\pi)$. Montrer qu'il existe $C \in M_n(\mathbb{C})$ telle que $S(t+2\pi) = S(t)\cdot C$ puis qu'il existe $B \in M_n(\mathbb{C})$ tel que $e^{2\pi B} = C$.

IV) Vérifier que la matrice $F(t)$ définie par $S(t)\cdot e^{-tB}$ satisfait la conclusion du théorème de Floquet.

La question demeure de savoir quand on peut affirmer que le système $y' = Ay$ possède une solution 2π-périodique. On apporte une solution à ce problème dans ce qui suit.

V) La matrice C considérée ci-dessus dépend de la base de solution $(y_1,\dots,y_n)$ de laquelle on part. Étudier comment se transforme C lorsqu'on utilise une autre base et en déduire que l'ensemble des valeurs propres de C ne dépend pas d'une choix de cette base.

Definition 3.1. Les valeurs propres ci-dessus sont appelées les **multiplicateurs** de l'équation.

VI) Démontrer le résultat suivant.

Théorème 6. *Le système $y' = Ay$ considéré ci-dessus admet une solution satisfaisant la relation $y(t+2\pi) = \mu y(t)$ si et seulement μ est un multiplicateur du système. En particulier, le système admet une solution 2π périodique, si et seulement si l'unité est un multiplicateur du système.* ■

VII) On note $S(\mu)$ l'ensemble des solutions du système différentiel satisfaisant $y(t+2\pi) = \mu y(t)$ est un sous-espace vectoriel de $\mathcal{S}$. Que peut-on dire de sa dimension ?

VIII) Nous revenons au cas simple des équations de la forme $y'(t) - g(t)y(t) = 0$ avec g 2π-périodique de $\mathbb{R}$ dans $\mathbb{R}$. Trouver une condition nécessaire et suffisante sur g assurant que les solutions sont 2π-périodiques.

*. Cette propriété est simple mais non élémentaire. Elle résulte d'une application du théorème d'inversion locale à la fonction $\exp$, le théorème s'appliquant puisque $d\exp(0)$ est l'identité de $M_n(\mathbb{C})$ dans $M_n(\mathbb{C})$

63 (Croissance des solutions d'une équation d'ordre 2, [1]). On considère l'équation différentielle linéaire d'ordre deux

$$y'' + (1+\theta(t))y = 0$$

où θ est continue sur $\mathbb{R}^+ = [0,\infty[$ avec $\int_0^\infty |\theta(t)|dt < \infty$.

(i) Justifiez l'existence de solutions définies sur $\mathbb{R}^+$.

On se propose de démontrer que toute solution f est bornée (sur $\mathbb{R}^+$).

(ii) On définit la fonction g sur $\mathbb{R}^+$ par

$$g(t) = f(t) + \int_0^t \theta(u)f(u)\sin(t-u)du.$$

Montrer que

$$g''(t) + g(t) = 0.$$

(iii) En déduire qu'il existe un réel A tel que

$$|f(t)| \leq A + \int_0^t |\theta(u)| \cdot |f(u)|\, du.$$

(iv) En déduire en étudiant la fonction

$$F(t) = \left\{A + \int_0^t |\theta(u)| \cdot |f(u)|\, du\right\} \cdot \exp\left(-\int_0^t |\theta(u)|du\right).$$

que la fonction f est bornée.

64 (Système de Lyapunov). On considère une équation différentielle linéaire à coefficients constants $y' = Ay$ où $A \in M_n(\mathbb{R})$ et $y : \mathbb{R} \to \mathbb{R}^n$. On fait l'hypothèse suivante. Il existe une matrice X définie positive telle que

$$A^T X + XA$$

est définie négative où A^T désigne la transposée de A. Rappelons qu'une matrice M est définie positive si elle la matrice d'un produit scalaire, autrement dit $\langle Mx \,,\, x\rangle > 0$ pour tout $x \in \mathbb{R}^n$ non nul. La définition de "définie négative" est similaire.

(i) Montrer qu'il existe $\epsilon > 0$ tel que $A^T X + XA + \epsilon X$ soit encore définie négative.

(ii) Montrer en considérant la dérivée de la fonction $t \to \langle Xy(t) \,,\, y(t)\rangle$ où y est une solution de $y' = Ay$ que

$$\langle Xy(t) \,,\, y(t)\rangle \leq \langle Xy(0) \,,\, y(0)\rangle e^{-\epsilon t}, \quad t \geq 0.$$

(iii) Déduire que

$$\|y(t)\| \leq \|y(0)\| \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}\, e^{-\frac{\epsilon t}{2}} \quad t \geq 0,$$

où $\lambda_{\max}$ est la plus grande, $\lambda_{\min}$ la plus petite des valeurs propres de X.

(iv) Trouver une hypothèse de même qui conduirait à une minoration plutôt qu'à une majoration de $\|y(t)\|$.

65 (Résolution par séries entières). On considère une équation différentielle de la forme

$$y''(z) + p(z)y'(z) + q(z)y(z).$$

Il s'agit ici d'une equation différentielle complexe mais le lecteur peut supposer que z est réel et que p, q sont des fonctions réelles et remplacer $\mathbb{C}$ par $\mathbb{R}$ dans ce qui suit. On dit que $a \in \mathbb{C}$ est un **point régulier** de l'équation si $p(z)$ et $q(z)$ sont développables en série entière au voisinage de a.

(i) Montrer la transformation

$$y(z) = w(z) \exp\left\{-\frac{1}{2}\int_0^z p(\xi)d\xi\right\}$$

ramène l'équation à la forme

$$w''(z) + J(z)w(z) = 0, \quad \text{avec} \quad J(z) = q(z) - \frac{1}{2}p'(z) - \frac{1}{4}p^2(z),$$

et que a est encore un point régulier de la nouvelle équation.

(ii) Montrer qu'il existe deux solutions linéairement indépendantes de l'équation développables en séries entières autour de a avec un rayon de convergence égal à celui de J en a. Sans perte de généralité, on pourra supposer que $a = 0$.

(iii) Résoudre l'équation $(1 - z^2)y'' - 2zy' + \frac{3}{4}z = 0$.

***

5

Appendice. Espaces vectoriels normés

§ 1. Définitions

1.1 Normes

Soit E un espace vectoriel sur $\mathbb{R}$. Une application N de $E \times E \to \mathbb{R}^+$ est une norme si elle vérifie les trois propriétés suivantes

(i) $N(x) = 0 \implies x = 0$, $x \in E$,

(ii) $N(\lambda x) = |\lambda| N(x)$, $x \in E$, $\lambda \in \mathbb{R}$,

(iii) $N(x + y) \leq N(x) + N(y)$, x, $y \in E$ (inégalité triangulaire).

Des deux dernières propriétés, nous déduisons

(D) $|N(x) - N(y)| \leq N(x - y)$, x, $y \in E$.

Les normes les plus classiques sont rappelées au paragraphe 2.1. Une des familles de normes les plus importantes est obtenue par l'emploi de produits scalaires, c'est-à-dire de formes bilinéaires symétriques définies positives sur E. En effet, si $(x, y) \in E \times E \to B(x, y)$ est une telle forme alors l'application N_B définie par $N_B(x) = (B(x, x))^{1/2}$ est une norme sur E. Nous dirons que c'est la norme induite par le produit scalaire B. Dans ce cas, l'inégalité triangulaire (iii) se démontre à l'aide de l'inégalité de Cauchy pour les produits scalaires.

1.2 Topologie

Les normes sont les outils fondamentaux pour étudier les notions topologiques sur un espace vectoriel. A chaque norme est associée une distance d_N définie par la relation $d_N(x, y) = N(x - y)$, $x, y \in E$, qui permet de construire une topologie sur

E. La boule ouverte de centre $a \in E$ et de rayon $r \in \mathbb{R}^+$ par rapport à la norme N est l'ensemble

$$B_N(a,r) = \{x \in E : N(x-a) < r\}. \tag{1.1}$$

Rappelons qu'un ensemble $V \subset E$ est un voisinage de a s'il contient une boule $B_N(a,r)$ pour un $r > 0$, $\Omega \subset E$ est ouvert s'il est un voisinage de chacun de ses points, autrement dit, chaque fois qu'il contient un élément a il contient aussi une boule $B_N(a,r)$ pour un rayon $r > 0$. Un ensemble est fermé si son complémentaire est ouvert. L'espace E muni de cette topologie -= qui est donc définie à partir des boules $B_N(a,r)$ et, par conséquent, à partir de N - est notée (E,N).

Soit F un second espace vectoriel et M une norme sur F. Une fonction f d'un ouvert Ω de (E,N) à valeurs dans F est dite continue en $a \in \Omega$ si pour tout $\epsilon > 0$, il existe $\eta > 0$ tel que $x \in B_N(a,\eta) \subset \Omega$ entraîne $f(x) \in B_M(f(a),\epsilon)$. Nous pouvons plus généralement définir $\lim_{x\to a} g(x)$ pour une fonction g définie sur un voisinage de a dans E à valeurs dans F *.

1.3 Equivalences des normes

Deux normes distinctes N_1 et N_2 sur E ne définissent pas nécessairement deux topologies différentes : lorsque les normes sont équivalentes, c'est-à-dire lorqu'il existe deux constantes c et C pour lesquelles

$$c\,N_1(x) \leq N_2(x) \leq C\,N_1(x), \quad x \in E, \tag{1.2}$$

un ensemble est un voisinage de a pour la première norme si et seulement si il l'est pour la seconde de sorte que toutes les notions topologiques coïncident.

Rappelons le théorème fondamental sur les normes d'un espace vectoriel de dimension finie.

Théorème 1. *Sur un espace vectoriel de dimension finie toutes les normes sont équivalentes.* ■

Le choix d'une norme sur un espace vectoriel de dimension finie n'a par conséquente aucune influence sur la validité d'un résultat de nature topologique. Cependant, dans certain cas, il peut arriver qu'une norme soit plus commode qu'une autre pour mener les calculs nécessaires à une démonstration. Notons aussi que dans certaines parties des mathématiques où il est nécessaire d'avoir des inégalités très précises, comme par exemple en analyse numérique matricielle, certaines normes sont privilégiées par rapport à d'autres.

---

*. Il suffit même, puisque la valeur en a de g n'intervient pas dans la définition de la limite, que g soit défini dans un voisinage de a privé de a.

1.4 Cas des espaces complexes

Tout ce qui a été dit s'étend de manière immédiate au cas où l'espace vectoriel E est un espace vectoriel complexe (les scalaires sont dans $\mathbb{C}$). Dans le deuxième axiome de la définition d'une norme, $|\lambda|$ doit être interprété comme le module du nombre complexe λ. Sauf cas exceptionnel, nous ne considérerons pas d'espace vectoriel normé sur $\mathbb{C}$ dans ce cours cependant l'ensemble des résultats des chapitres 1 et 2 restent valables dans ce cadre. Traditionnellement, pour désigner une norme, une notation de la forme $\|x\|$ est généralement employée à la place de $N(x)$. C'est surtout le cas lorsque une seule norme est employée dans le contexte. Nous nous conformerons à cette tradition.

§ 2. Exemples

2.1 Espaces fondamentaux

La table suivante présente quelques exemples fondamentaux d'espaces vectoriels normés. D'autres seront donnés plus bas après que nous aurons rappelé les résultats sur les applications linéaires continues.

$\mathbb{R}^n$	Les normes usuelles sont $\|x\| = \max_{i=1}^n |x_i|$, $\|x\|_1 = \sum_{i=1}^n |x_i|$ et $\|x\|_2 = (\sum_{i=1,\ldots,n} |x_i|^2)^{1/2}$. La dernière norme est induite du produit scalaire canonique sur $\mathbb{R}^n$
$M_n(\mathbb{R})$	L'espace des matrices $n \times n$ à coefficients réels. Les normes les plus intéressantes sur $M_n(\mathbb{R})$ sont les normes sous-multiplicatives qui vérifient $\|A\,B\| \leq \|A\|\,\|B\|$, $A, B \in M_n(\mathbb{R})$. Les exemples les plus importants sont rappelés au paragraphe 4.4.
$\mathsf{C}[a,b]$	L'espace des fonctions réelles continues sur l'intervalle fermé borné $[a,b]$. Il est muni de la norme Sup défini par $\|f\| = \max_{x\in[a,b]} |f(x)|$.
ℓ^2	L'espace des suite réelles $x = (x_n)$ à carrés sommables c'est-à-dire telles que la série $\sum_{n=0}^\infty x_n^2$ soit convergente. La norme est alors simplement $\|x\| = (\sum_{n=0}^\infty x_n^2)^{1/2}$ qui est induite du produit scalaire $(x\,,\,y) = \sum_{n=0}^\infty x_n\,y_n$.
$L^2(\Omega)$	L'espace des fonctions de carré (Lebesgue) intégrables sur l'ouvert Ω, plus précisément, l'espace de ces fonctions modulo la relation d'équivalence $f \sim g$ pour laquelle deux fonctions sont en relation si le sous-ensemble de Ω sur lequel elles diffèrent est de mesure de Lebesgue nulle. La norme est celle induite du produit scalaire $(f,g) = \int_\Omega fg dm$.

2.2 Produit cartésien d'espaces vectoriels normés

Signalons encore que si (E, N) et (F, M) sont deux espaces vectoriels normés alors leur produit cartésien $E \times F$ est lui-même normé par l'application

$$\|(x, y)\| = \max(N(x), M(y)).$$

Nous pourrions aussi prendre $\|(x, y)\| = N(x) + M(y)$ ou encore, par exemple, $\|(x, y)\| = (N^2(x) + M^2(y))^{1/2}$ dont nous vérifions immédiatement qu'elles sont équivalentes à la première. Chaque fois que nous considérons un produit d'espaces vectoriels normés, si nous ne signalons pas explicitement le contraire, ce produit sera muni de la première des normes ci-dessus, celle construite comme le maximum des normes de chaque composante. Nous pouvons naturellement étendre cette construction au cas d'un produit cartésien $E_1 \times E_2 \times \ldots E_k$ de k espaces vectoriels normés.

§ 3. Espaces de Banach

3.1 Définition

Un espace vectoriel normé qui est complet, c'est-à-dire pour lequel toute suite de Cauchy est une suite convergente, s'appelle un espace de Banach. Pour tous les développements profonds de l'analyse la propriété de complétude est essentielle et les espaces vectoriels normés qui ne sont pas complets jouent un rôle plutôt marginal en mathématiques. D'ailleurs, il est connu que tout espace vectoriel normé peut être plongé dans un espace de Banach. Lorsque la norme d'un espace de Banach est induite d'un produit scolaire, nous parlons d'espace de Hilbert. Les exemples d'espaces vectoriels normés du paragraphe 2.1 sont tous des espaces de Banach. Nous aurons besoin d'un espace Banach un peu plus général que $\mathsf{C}[a, b]$, l'espace de Banach $\mathsf{C}([a, b], X)$ formé des applications continues de $[a, b]$ dans X où X est un sous-ensemble fermé d'un espace de Banach F. Les espaces ℓ^2 et $L^2(\Omega)$ sont des espaces de Hilbert.

3.2 Le théorème du point fixe

Une application f d'un ensemble X dans lui-même, X étant un sous-ensemble d'un espace vectoriel normé $(E, \|\cdot\|)$ est dite contractante — ou encore, est appelée une contraction — s'il existe un réel k, compris strictement entre 0 et 1, tel que

$$\|f(x) - f(y)\| \leq k\,\|x - y\|, \qquad x,\, y \in X. \tag{3.1}$$

Autrement dit, les applications contractantes de X dans X sont celles qui sont lipschitziennes avec un rapport strictement plus petit que 1.

Le théorème suivant est un des outils les plus puissants des mathématiques, notamment pour montrer l'existence et l'unicité (mais aussi construire) des solutions d'équations fonctionnelles.

Théorème 2 (du point fixe). *Une application contractante f d'un fermé X d'un espace de Banach $(E, \|\cdot\|)$ dans lui-même, i.e. $f(X) \subset X$, possède un et un seul point fixe s. Autrement dit, l'équation $f(x) = x$ admet une et une seule solution dans X. Cette solution est donnée par la limite de la suite (x_n) définie par récurrence par $x_0 = a$ et $x_{n+1} = f(x_n)$, $n \geq 0$, le point de départ a étant librement choisi dans X.* ■

Nous aurons l'occasion d'utiliser une version d'apparence légèrement plus forte.

Corollaire 5.1. *Le théorème du point fixe demeure vrai si l'hypothèse que f est contractante est remplacée par l'hypothèse qu'une des itérées de f est contractante, cela signifie qu'il existe un entier $k \geq 1$ tel que $f^{[k]} = f \circ f \circ \cdots \circ f$ (k fois) est contractante.* ■

Démonstration. En effet si $f(X) \subset X$ alors $f^{[k]}(X) \subset X$ et nous pouvons appliquer le théorème à la fonction $f^{[k]}$. Le théorème nous dit que l'équation $f^{[k]}(x) = x$ admet une unique solution s. Nous avons alors $f^{[k]}(f(s)) = f^{[k+1]}(s) = f(f^{[k]}(s)) = f(s)$ de sorte que $f(s) \in X$ est une autre solution de l'équation $f^{[k]}(x) = x$ (dans X). Cette solution étant unique, nous avons nécessairement $f(s) = s$. Ceci montre que l'équation $f(x) = x$ admet au moins une solution. Mais si nous avons $f(s') = s'$ pour un autre éléments s' de X alors nous aurions par applications répétées de f, $f^{[k]}(s') = s'$ de sorte que s' serait aussi une seconde solution de l'équation $f^{[k]}(x) = x$ ce qui impossible. Il nous reste à établir que l'unique solution s est la limite de la suite (x_n) définie par la relation de récurrence $x_{n+1} = f(x_n)$ et $x_0 = a$, a étant un élément quelconque de X. Le théorème du point fixe appliqué à $f^{[k]}$ nous dit que suite x_{nk} converge vers s. La suite définie par $y_{n+1} = f^{[k]}(y_n)$ et $y_0 = f(a)$ converge aussi vers s mais celle-ci n'est autre que (x_{nk+1}). En continuant ainsi nous obtenons que toutes les suites (x_{nk+r}) avec $r = 0, \ldots, k-1$ convergent vers s. Il n'est alors plus difficile de vérifier que (x_n) elle-même converge vers s. ■

§ 4. Applications linéaires continues

4.1 *Caractérisations des applications linéaires continues*

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés et u une application linéaire de E dans F. Un théorème fondamental dit que les trois assertions suivantes sont équivalentes

(i) u est continue en l'origine 0 (nous appelons "origine" le vecteur nul de E),

(ii) u est continue en tout point de E,

(iii) Il existe $K \in \mathbb{R}$ tel que $\|u(x)\|_F \leq K\,\|x\|_E$), $x \in E$.

Remarquons que dans le cas d'une fonction quelconque f, seules son vraies les implications (ii) $\Longrightarrow$ (i) et, lorsque $f(0) = 0$, (iii) $\Longrightarrow$ (i). La propriété (iii) implique, en tenant compte de l'additivité des applications linéaires, que les applications linéaires continues sont lipschitziennes sur E et uniformément continues sur E.

4.2 Norme d'une application linéaire continue

A cause de l'homogénéité des applications linéaires, $u(\lambda x) = \lambda u(x)$, il suffit d'établir l'inégalité dans (iii) pour l'ensemble des vecteurs x dont la norme n'excède pas 1. En fait, il n'est pas difficile de vérifier que la condition (iii) est satisfaite si et seulement si la fonction $\|u(x)\|_F/\|x\|_E$ est bornée supérieurement sur $\{x \in E : 0 < \|x\|_E \leq 1\}$, ou, ce qui revient au même sur $\{x \in E : \|x\|_E = 1\}$. Posant

$$\|u\| = \sup_{x \neq 0} \frac{\|u(x)\|_F}{\|x\|_E} = \sup_{0<\|x\|_E \leq 1} \frac{\|u(x)\|_F}{\|x\|_E} = \sup_{\|x\|_E = 1} \|u(x)\|_F, \tag{4.1}$$

l'application linéaire u est continue si et seulement si $\|u\| < \infty$ et, dans ce cas,

$$\|u(x)\|_F \leq \|u\| \, \|x\|_E, \quad x \in E. \tag{4.2}$$

L'espace vectoriel des applications linéaires continues de (E, N) dans (F, M) sera noté $\mathcal{L}(E; F)$ ou, lorsque $E = F$, $\mathcal{L}(E)$. Il s'agit d'un abus de notation car la continuité de u dépend un général (lorsque nous ne travaillons pas sur des espaces vectoriels de dimension finie) des normes de E et F de sorte qu'il faudrait écrire, en toute rigueur, $\mathcal{L}((E, N); (F, M))$. Il faudrait aussi écrire $\|\cdot\|_{E,F}$ voire $\|\cdot\|_{(E,N)(F,M)}$ au lieu de $\|\cdot\|$. Le théorème suivant confirme ce que l'emploi de la notation $\|\cdot\|$ a déjà dû rappeler au lecteur.

Théorème 3. *La fonction $u \to \|u\|$ est une norme sur $\mathcal{L}(E; F)$. Lorsque $(F, \|\cdot\|_F)$ est un espace de Banach, $(\mathcal{L}(E; F), \|\cdot\|)$ l'est aussi.* ■

Remarquons que dire que deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sur E sont équivalentes revient à dire que l'application $x \in (E, \|\cdot\|_1) \to x \in (E, \|\cdot\|_2)$ est une application linéaire continue ainsi que sa réciproque.

4.3 Composition

Soient E, F, G, trois espaces vectoriels normés, $u \in \mathcal{L}(E; F)$ et $v \in \mathcal{L}(F; G)$. Puisque la composée de deux fonctions continues est encore continue nous avons aussi $v \circ u \in \mathcal{L}(E; G)$. Nous déduisons alors immédiatement de la formule (4.1) que

$$\|v \circ u\| \leq \|v\| \cdot \|u\|, \tag{4.3}$$

ou encore, de manière plus précise,

$$\|v \circ u\|_{E,G} \leq \|v\|_{F,G} \cdot \|u\|_{E,F}. \tag{4.4}$$

En particulier, lorsque $u \in \mathcal{L}(E)$, nous avons $\|u^2\| \leq \|u\|^2$ et, plus généralement,

$$\|u^n\| \leq \|u\|^n, \quad n \in \mathbb{N}. \tag{4.5}$$

4.4 Cas de la dimension finie

Dans le cas de la dimension finie, la question de la continuité d'une application linéaire ne se pose pas :

Théorème 4. *Les applications linéaires sur un espace vectoriel de dimension finie (à valeurs dans n'importe quel espace vectoriel normé) sont toujours continues.* ■

Insistons cependant sur le fait que, même si les applications linéaires sur un espace vectoriel la dimension finie sont toujours continues, le calcul ou l'estimation de leurs normes revêt dans bien des questions une grande importance. Il est utile de rappeler à ce propos le lien entre les normes sur les espaces de matrices et les normes d'application linéaire. Nous savons qu'à toute matrice $M = (a_{ij}) \in M_n(\mathbb{R})$ est canoniquement attachée une application linéaire $\mathcal{M}$ de $\mathbb{R}^n$ dans $\mathbb{R}^n$, cette application linéaire est définie par la relation

$$\mathcal{M}(e_j) = \sum_{i=1}^{n} a_{ij} e_i, \quad j = 1, \ldots, n \tag{4.6}$$

où e_i désigne le i-ème vecteur de la base canonique de $\mathbb{R}^n$. Nous pouvons alors définir la norme de M comme celle de son application linéaire associée $\mathcal{M}$. Pour calculer la norme de cette application linéaire à l'aide de la formule qui a été donnée ci-dessus nous devons d'abord fixer la norme de $\mathbb{R}^n$ avec laquelle nous calculerons $\|\mathcal{M}\|$,

$$\|M\| = \|\mathcal{M}\| = \sup_{\|x\|=1} \|\mathcal{M}(x)\|. \tag{4.7}$$

Le tableau suivant donne les trois exemples les plus importants.

Norme sur $\mathbb{R}^n$	Norme sur $M_n(\mathbb{R})$
$\|x\| = \max_{1\le i\le n} \|x_i\|$	$\|M\| = \max_{1\le i\le n} \left(\sum_{j=1}^n \|a_{ij}\|\right)$
$\|x\|_1 = \sum_{i=1}^n \|x_i\|$	$\|M\|_1 = \max_{1\le j\le n} \left(\sum_{i=1}^n \|a_{ij}\|\right)$
$\|x\|_2 = \left(\sum_{i=1}^n x_i^2\right)^{1/2}$	$\|M\|_2 = \rho({}^t M\, M)$.

Dans la troisième formule $\rho({}^t M\, M)$ désigne le rayon spectral de la matrice ${}^t M\, M$; c'est la le maximum des valeurs absolues des valeurs propres de cette matrice. Les normes obtenues comme ci-dessus à partir de normes d'applications linéaires sont dites induites de la norme correspondante sur $\mathbb{R}^n$. Il existe bien d'autres normes utiles en analyse matricielle.

§ 5. Applications multilinéaires continues

5.1 Applications multilinéaires

Nous avons déjà évoqué en donnant la définition d'une norme la notion de forme bilinéaire, en supposant que le lecteur avait une certaine familiarité avec elle. Nous rappelons dans cette partie les propriétés fondamentales d'une famille un peu plus générale d'applications, les applications multilinéaires. Les applications bilinéaires en resteront les exemples les plus d'importants.

Soient E_i, $i = 1, \dots, k$, et F, $k + 1$ espaces vectoriels (sur $\mathbb{R}$). Une application $u : E_1 \times E_2 \times \cdots \times E_k \to F$ est dite multilinéaire, ou, de manière plus précise, k-linéaire, si elle est linéaire par rapport à chacune de ses k variables. Cela signifie que la relation

$$\begin{aligned} &u(x_1, \dots, x_{i-1}, \lambda x + \mu y, x_{i+1}, \dots, x_k) = \\ &\qquad \lambda u(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_k) + \mu u(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_k) \end{aligned} \tag{5.1}$$

doit être satisfaite pour tous les $i = 1, \dots, k$, $\lambda, \mu \in \mathbb{R}$, $x, y \in E_i$, $x_j \in E_j$, $j \neq i$. L'ensemble de ces applications k-linéaires – que nous pouvons munir des lois habituelles d'addition des fonctions et de multiplication par un scalaire réel – forme un espace vectoriel noté $L(E_1, E_2, \dots, E_k; F)$. Lorsque $F = \mathbb{R}$ et $E_1 = E_2 = \cdots = E_k = E$, nous parlons simplement de forme k-linéaire sur E.

Théorème 5.

$$L(E_1, E_2, \dots, E_k; F) \sim L(E_1, L(E_2, \dots, E_k; F)). \tag{5.2}$$

■

Dans cet énoncé, le symbole $\sim$ signifie qu'il existe une application linéaire bijective (un isomorphisme linéaire) entre les deux espaces vectoriels. Cet isomorphisme est très facile à construire. Il suffit de poser

$$\Phi : L(E_1, L(E_2, \dots, E_k; F)) \to L(E_1, E_2, \dots, E_k; F) \tag{5.3}$$

avec $\Phi(\Gamma)$, $\Gamma \in L(E_1, L(E_2, \dots, E_k; F))$, définie par

$$\Phi(\Gamma)(x_1, x_2, \dots, x_k) = \Gamma(x_1)(x_2, \dots, x_k). \tag{5.4}$$

Il est facile de vérifier que cette application prend bien ses valeurs dans l'espace $L(E_1, E_2, \dots, E_k; F)$ et qu'elle est linéaire et bijective.

5.2 Continuité des applications multilinéaires

Lorsque les espaces E_i, $i = 1, \ldots, k$, sont des espaces vectoriels normés, $E_1 \times \cdots \times E_k$ est muni de la norme d'un espace vectoriel produit (voir 2.2) et il est légitime d'étudier la continuité de $u \in L(E_1, E_2, \ldots, E_k; F)$. Le sous-ensemble des applications multilinéaires continues forme alors un sous-espace vectoriel de $L(E_1, E_2, \ldots, E_k; F)$ que nous noterons

$$\mathcal{L}(E_1, E_2, \ldots, E_k; F).$$

La continuité des applications multilinéaires se traite d'une manière tout à fait similaire à celles des applications linéaires. Nous nous limiterons à indiquer le

Théorème 6. *Pour que $u \in \mathcal{L}(E_1, E_2, \ldots, E_k; F)$ il faut et il suffit que le nombre $\|u\|$ défini par la relation ci-dessous soit fini,*

$$\|u\| = \max\{\|u(x_1, \ldots, x_k)\|_F \ : \ \|x_i\|_{E_i} = 1, \ 1 \leq i \leq k\}. \tag{5.5}$$

nous avons alors l'inégalité

$$\|u(x_1, \ldots, x_k)\|_F \leq \|u\| \, \|x_1\|_{E_1} \, \|x_2\|_{E_2} \cdots \|x_k\|_{E_k}, \quad x_i \in E_i, \ i = 1, \ldots, k \tag{5.6}$$

et l'application $\|\cdot\|$ est une norme sur $\mathcal{L}(E_1, E_2, \ldots, E_k; F)$. En outre,

$$\mathcal{L}(E_1, E_2, \ldots, E_k; F) \sim \mathcal{L}(E_1, \mathcal{L}(E_2, \ldots, E_k; F)), \tag{5.7}$$

où le symbole $\sim$ signifie maintenant qu'il existe une application linéaire bijective entre les deux espaces qui est continue et dont la réciproque est aussi continue. ■

L'application linéaire bijective dont il est question dans cet énoncé est naturellement la même, à savoir l'application Φ de (5.4) que celle qui intervient dans le théorème 5 et, en utilisant la même notation pour deux normes différentes, $\|\Phi(\Gamma)\| = \|\Gamma\|$.

Théorème 7. *Lorsque tous les espaces E_i sont de dimension finie, toutes les applications multilinéaires sont continues, autrement dit,*

$$\mathcal{L}(E_1, E_2, \ldots, E_k; F) = L(E_1, E_2, \ldots, E_k; F). \tag{5.8}$$

■

Index

Bibliographie

[1] D. Azé, G. Constans, and Hiriart-Urruty J.-B. *Calcul différentiel et équations différentielles.* Dunod, Paris, 2002.

[2] H. Cartan. *Calcul différentiel.* Hermann, Paris, 1967.

[3] J. Dieudonné. *Calcul infinitésimal.* Hermann, Paris, 1980.

[4] W. Rudin. *Principles of Mathematical Analysis.* McGraw-Hill International Editions, Singapore, 1976.

[5] L. Schwartz. *Analyse mathématique. I, II.* Hermann, Paris, 1967. Cours de l'école polytechnique.