Géométrie différentielle

Marc SAGE

8 avril 2009

Table des matières

La différentielle d'une application f en un point a sera notée indifféremment

$$d_a f = df_a = df(a).$$

La dérivée partielle d'une application f selon la i-ième coordonnée sera notée indifféremment

$$\partial_i f = \frac{\partial f}{\partial x_i}$$

Rappelons les liens entre extrema locaux et hessiennes. Soit H_a la hessienne d'une application f de classe C^2 au voisinage d'un point a, à valeurs réelles. En écrivant

$$f(a+h) = f(a) + d_a f(h) + H_a(h) + o\left(h^2\right),$$

on montre aisément les implications suivantes :

Si f admet un minimum local en a, alors H_a est positive ;
Si H_a est *définie* positive, alors f admet un minimum local en a.

On rappelle également que le *laplacien* d'une telle fonction est défini par

$$\Delta f := \sum \partial_{ii}^2 f = \sum \frac{\partial^2 f}{\partial x_i^2} = \operatorname{tr} H$$

et qu'une fonction est dite *harmonique* lorsque son laplacien est nul.

1 Mise en jambe

Soit f une application de $\mathbb{R}^n$ dans $\mathbb{R}$ différentiable en 0 telle que

$$\forall x \neq 0,\ \forall \lambda > 0,\ f(\lambda x) = \lambda f(x).$$

Montrer que f est linéaire.

Solution proposée.
Il s'agit de montrer que f vaut sa différentielle en 0.
f étant continue en 0 car différentiable en 0, on obtient $f(0) = 0$ en faisant tendre λ vers 0 dans la relation $f(\lambda x) = \lambda f(x)$.
Par ailleurs, la différentiabilité de f permet d'écrire

$$\forall a \in \mathbb{R}^n,\ f(a) = \underbrace{f(0)}_{=0} + d_0 f(a) + \varepsilon(a)\|a\| \text{ où } \lim_0 \varepsilon = 0.$$

En réinjectant dans les relations données, on obtient (à $a \in \mathbb{R}^n$ et $\lambda \neq 0$ fixés) :

$$\begin{aligned} f(\lambda a) &= \lambda f(a) \\ d_0 f(\lambda a) + \varepsilon(\lambda a)\|\lambda a\| &= \lambda d_0 f(a) + \lambda \varepsilon(a)\|a\| \\ \lambda \varepsilon(\lambda a)\|a\| &= \lambda \varepsilon(a)\|a\| \\ \varepsilon(\lambda a) &= \varepsilon(a). \end{aligned}$$

Le terme $\varepsilon(a)$ de droite ne dépend plus de λ : en prenant la limite quand $\lambda \longrightarrow 0$, on trouve $\varepsilon(a) = 0$, et ceci pour tout a, d'où $f = d_0 f$, *CQFD*.

Remarque. Si l'on avait utilisé la notation $o(a)$, on n'aurait probablement pas pu raisonner aussi fininement. On voit ici les limites de cette notation très souple : $o(x)$ dénote à la fois la fonction devant laquelle le o est négligeable et la variable de cette fonction. Le lecteur pourra faire le lien avec l'exercice 5 de la feuille sur la dérivation.

2 Sur les difféomorphismes

1. *Montrer que la racine carrée est un C^∞-difféomorphisme des matrices réelles symétriques définies positives.*
2. *Montrer que l'application suivante est un C^∞-difféomorphisme global sur les polynômes réels scindés simples unitaires de degré n.*

$$\varphi : \left\{ \begin{array}{ccc} \left\{ \overrightarrow{\lambda} \in \mathbb{R}^n \ ; \ \lambda_1 < \cdots < \lambda_n \right\} & \longrightarrow & \mathbb{R}_n [X] \\ \overrightarrow{\lambda} & \longmapsto & \prod (X - \lambda_i) \end{array} \right. .$$

Solution proposée.

1. Regardons plutôt l'inverse de la racine : l'élévation au carré. C'est injectif par unicité de la racine carré, C^∞ car polynomial. Si l'on montre que sa différentielle est injective sur S_n^{++}, les conditions d'un C^∞-difféomorphisme seront réunies.

   La différentielle du carré en une matrice A vaut $(A\cdot)+(\cdot A)$. Lorsque $A \in S_n^{++}$, on peut la diagonaliser, mettons $A = P\Lambda P^{-1}$ où Λ est diagonale. Un élément PMP^{-1} du noyau de $(A\cdot) + (\cdot A)$ vérifiera donc $MD + DM = 0$, *i. e.* $M = -D^{-1}MD$. Prenant les coefficients, il vient $m_{i,j} = -\frac{\lambda_j}{\lambda_i} m_{i,j}$ pour tout i, j ; le quotient $-\frac{\lambda_j}{\lambda_i}$ étant strictement négatif, on doit avoir $m_{i,j} = 0$, d'où $M = 0$, *CQFD*.
2. φ est clairement injective. Il s'agit de montrer que sa différentielle en tout $\overrightarrow{\lambda}$ est également injective. Soit donc $\overrightarrow{x}$ dans le noyau d'un $d_\lambda \varphi$, ce qui s'écrit

$$\sum \partial_i \varphi (\lambda) \ x_i = 0 \iff \sum_i x_i \prod_{j \neq i} (X - \lambda_j) = 0.$$

   Faire $X = \lambda_k$ donne $x_k \prod_{j \neq k} (\lambda_k - \lambda_j) = 0$, d'où $x_k = 0$ car les λ_i sont tous distincts, *CQFD*.

Remarque. On vient de montrer que les racines d'un polynôme réel unitaire scindé simple variaient de manière C^∞ en fonction de leur coefficients.

Il vient également que les polynômes unitaires[1] scindés simples de $\mathbb{R}_n [X]$ forment un ouvert de ce dernier, ce qui peut se retrouver à l'aide du résultant[2]

3 Différentiabilité du déterminant

1. *Différentier le déterminant dans $M_n(\mathbb{R})$. Comment s'exprime son gradient ? Interprêter.*
2. *Calculer la hessienne du déterminant dans $M_n(\mathbb{R})$. Qu'observe-t-on ?*

Solution proposée.

1. Notons déjà que det est C^∞ car polynomial en ses coordonnées. Nous présentons trois méthodes.
   (a) Revenons à la définition de la différentielle. Fixons une matrice $A \in M_n$. Pour $M \in M_n$, on étudie l'accroissement $\det (A + M) - \det A$. Pour développer le premier terme, on explicite les colonnes de A et M :

$$\det (A + M) = \det (A_1 + M_1 \mid \cdots \mid A_n + M_n) .$$

   Lorsque l'on développe, les termes contenant exactement k colonnes de M seront des polynômes de degré k en les coefficients de M ; comme tous les termes de degré ≥ 2 sont des $O\left(H^2\right)$, donc des $o(H)$, on ne garde que les termes pour $k = 0$ et 1 :

$$\det (A + M) = \det (A_1 \mid \cdots \mid A_n) + \sum_{j=1}^{n} \det (A_1 \mid \cdots \mid A_{j-1} \mid M_j \mid A_{j+1} \mid \cdots \mid A_n) + o(M) .$$

[1] hypothèse dont on peut aisément se passer
[2] *cf.* seconde feuille sur les polynômes

Le premier terme est $\det A$. En développant le j-ième terme de la somme selon la colonne M_j, on voit qu'il vaut $\sum_{i=1}^{n} m_{i,j} \left[\operatorname{Com} A\right]_{i,j}$. On obtient par conséquent

$$\det\left(A+M\right)=\det A+\sum_{i,j}\left[\operatorname{Com} A\right]_{i,j} m_{i,j}+o\left(M\right),$$

d'où la valeur de la différentielle $d_A \det$ prise en une matrice M :

$$\left[d_A \det\right]\left(M\right)=\sum_{i,j}\left[\operatorname{Com} A\right]_{i,j}\left[M\right]_{i,j}=\left\langle \operatorname{Com} A \mid M\right\rangle.$$

Pour obtenir le gradient, on revient à la définition : la différentielle en A est le produit scalaire contre un vecteur, appelé gradient. On en déduit immédiatement

$$\overrightarrow{\nabla} \det=\operatorname{Com}.$$

On obtient ainsi une interprétation géométrique de la comatrice : c'est la direction vers laquelle le déterminant augmente le plus.

(b) On sait que la donnée d'une base (e_i) permet de décomposer la différentielle d'une application scalaire f comme somme des accroissements $\partial_i f$ selon la base dx_i. On s'intéresse donc à la dérivée partielle de det selon les matrices élémentaire $E_{i,j}$ (qui forment une base de M_n) en une matrice donnée A.

On développe $\det\left(A+tE_{i,j}\right)$ par linéarité par rapport à la j-ième colonne, ce qui donne $\det A+t\det A_j$ où A_j désigne la matrice A où l'on a retiré la j-ième colonne pour y mettre une colonne avec que des 0 et un 1 à la i-ième place. C'est une incitation à dévélopper $\det A_j$ selon la j-ième colonne : il sort le cofacteur de place (i,j). Finalement, on obtient

$$\left[\partial_{i,j}\det\right]\left(A\right)=\left[\operatorname{Com} A\right]_{i,j}.$$

On conclut comme dans la première méthode.

(c) Proposons une troisième méthode adaptée aux calculs de différentielles sur des groupes.

On commence par calculer la différentielle en 1 (l'identité). Le calcul direct est ici aisé en faisant intervenir le polynôme caractéristique χ_H :

$$\det\left(1+H\right)=\chi_H\left(-1\right)=1+\operatorname{tr} H+O\left(H^2\right).$$

On en déduit la différentielle en toute matrice A inversible :

$$\begin{aligned}\det\left(A+H\right)&=\left|A\right|\det\left(1+HA^{-1}\right)\\&=\left|A\right|\left(1+\operatorname{tr}\left(A^{-1}H\right)+o\left(H\right)\right)\\&=\det A+\operatorname{tr}\left(\left|A\right|A^{-1}H\right)+o\left(H\right).\end{aligned}$$

Or, en se souvenant que $A^{-1}=\frac{{}^t\operatorname{Com} A}{|A|}$, on retrouve la différentielle

$$\begin{aligned}d_A\det\left(H\right)&=\operatorname{tr}\left(\left|A\right|A^{-1}H\right)\\&=\operatorname{tr}\left({}^t\operatorname{Com} A\times H\right)\\&=\left\langle \operatorname{Com} A \mid H\right\rangle.\end{aligned}$$

On propage enfin le résultat par densité de GL_n. L'application det étant C^1, l'application $d\det$ est continue. De même, Com est continue car polynomiale. L'identité $d_A\det=\left\langle \operatorname{Com} A \mid \cdot\right\rangle$, valable pour $A\in GL_n$, se propage donc par continuité à tout $A\in M_n$, *CQFD.*

2. Le calcul de la hessienne en une matrice A revient au calcul des dérivées partielles $\left[\partial_{i,j}\partial_{k,l}\det\right]\left(A\right)$. Puisque $\left[\partial_{k,l}\det\right]\left(A\right)$ est le déterminant de la matrice $A_{k,l}$ obtenue en retirant de A ses k-ième ligne et l-ième colonne, dériver $\det A_{k,l}$ selon $E_{i,j}$ donne d'après le calcul précédent $\left[\operatorname{Com} A_{k,l}\right]_{i,j}$, autrement dit (à un signe près) le déterminant de la matrice A privée des lignes i et k ainsi que les colonnes j et l. Pour $i=k$ ou $j=l$, cela donne 0. En particulier, les termes diagonaux sont nuls, donc la trace aussi, ce qui montre que le déterminant est harmonique.

4 Différentiabilité de l'inverse

Différentier l'inverse dans $GL_n(\mathbb{K})$.

Solution proposée.
Notons Inv l'application « inverse ». Elle est définie sur un ouvert de $M_n(\mathbb{K})$ comme image réciproque de l'ouvert $\mathbb{K}^*$ par l'application continue det. Elle y est même C^∞ en vertu de la formule

$$\operatorname{Inv} A = \frac{1}{\det A}\,{}^t\operatorname{Com} A \text{ (tout est polynomial).}$$

1. S'inpirant de la troisième méthode de l'exercice précédente, on calcule d'abord la différentielle en l'identité puis on la « transportera » en un point quelconque de GL_n.

   Pour obtenir $d_{I_n}\operatorname{Inv}$, on calcule les dérivées partielles selon les matrices de bases $E_{i,j}$.

   Pour $i=j$, on a

$$\frac{\operatorname{Inv}(I_n+tE_{i,i})-\operatorname{Inv}(I_n)}{t}=\frac{\left(\frac{1}{1+t}-1\right)E_{i,i}}{t}=\frac{-1}{1+t}E_{i,i}\longrightarrow -E_{i,i},$$

   tandis que pour $i\neq j$ on obtient

$$\frac{\operatorname{Inv}(I_n+tE_{i,j})-\operatorname{Inv}(I_n)}{t}=\frac{(I_n-tE_{i,j})-I_n}{t}\longrightarrow -E_{i,j},$$

   formule valable pour $i=j$ d'après la ligne précédente. On trouve donc

$$d_{I_n}\operatorname{Inv}(M)=\sum m_{i,j}(-E_{i,j})=-M,$$

   d'où le résultat[3]

$$d_{I_n}\operatorname{Inv}=-\operatorname*{Id}_{GL_n}.$$

   Pour « transporter » cette différentielle en une matrice quelconque de GL_n, on utilise la structure de groupe de GL_n. Étant donnée un matrice H de perturbation, on peut écrire

$$\begin{aligned}\operatorname{Inv}(A+H) &= \left[A\left(I_n+A^{-1}H\right)\right]^{-1}=\operatorname{Inv}\left(I_n+A^{-1}H\right)A^{-1}=\left(I_n-A^{-1}H+o(H)\right)A^{-1}\\ &= \operatorname{Inv}A-A^{-1}HA^{-1}+o(H),\end{aligned}$$

   d'où le résultat[4]

$$d_A\operatorname{Inv}=-A^{-1}\cdot A^{-1}.$$

2. On différentie l'égalité $1_{M_n}=\operatorname{Id}_{GL_n}\times\operatorname{Inv}$ valable sur GL_n en une matrice A :

$$\begin{aligned}0 &= d_A\operatorname{Id}\times\operatorname{Inv}(A)+\operatorname{Id}(A)\times d_A\operatorname{Inv}\\ &= \operatorname{Id}\times A^{-1}+A\times d_A\operatorname{Inv}\text{, d'où}\\ d_A\operatorname{Inv} &= -A^{-1}\times\operatorname{Id}\times A^{-1},\ \textit{CQFD.}\end{aligned}$$

   Noter l'importance ici de raisonner avec les *fonctions* et non avec les *valeurs* des fonctions : on confondrait aisément le point où l'on différentie (qui doit rester dans l'ouvert de $\mathbb{R}^d$ où l'on peut différentier) avec l'argument de la différentielle (qui peut varier dans tout l'ev $\mathbb{R}^d$). Même si les physiciens font cela souvent et si c'est efficace, revenir aux définitions permet de bien tout contrôler afin d'éviter les bétises.

Remarque. On peut facilement prolonger ce résultat au groupe linéaire de n'importe quel espace de Banach. Évidemment, on ne regardera que les isomorphismes *continus*, leur réciproque est alors automatiquement continue par le théorème de Banach[5].

---

[3] lequel ne fait que conforter l'intuition du développement

$$\frac{1}{1+x}=1-x+O(x^2)$$

[4] lequel ne fait décidément que conforter l'intuition du développement

$$\frac{1}{a+x}=\frac{1}{a}\frac{1}{1+\frac{x}{a}}=\frac{1}{a}\left(1-\frac{x}{a}+O(x^2)\right)=\frac{1}{a}-\frac{x}{a^2}+O(x^2)$$

[5] *cf.* seconde feuille sur les espaces de Banach

5 Théorème de Rolle multidimensionnel

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ continue sur un compact K, différentiable sur son intérieur $\mathring{K}$ et constante sur son bord ∂K.

Montrer que df s'annule à l'intérieur de $\mathring{K}$.

Solution proposée.

Lorsque $n = 1$ et $K = [a, b]$, on retombe sur l'énoncé unidimensionnel du théorème de Rolle. Calquons donc la démonstration pour généraliser à la dimension n.

K étant compact, f y atteint ses extrema, mettons $\inf_K f =: \alpha \leq \beta := \sup_K f$. Si $\alpha = \beta$, f est constante sur tout K, d'où $df = 0$ en tout point de $\mathring{K}$. Dans le cas contraire, α ou β est atteint en dehors du bord ∂K (car f y est constante), et en ce point intérieur à K la différentielle s'annule.

6 Principe du maximum

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe C^2 et Ω un ouvert borné de $\mathbb{R}^n$.

1. *Montrer que si $\Delta f > 0$ sur Ω, alors*
$$f_{|\Omega} < \max_{\partial\Omega} f.$$

2. *En déduire que, si f est harmonique sur Ω, alors*
$$\min_{\partial\Omega} f \leq f_{|\overline{\Omega}} \leq \max_{\partial\Omega} f.$$

3. *Montrer que, si f est harmonique et est constante sur $\partial\Omega$, alors f est constante sur Ω.*

Solution proposée.

Remarquons tout d'abord que $\partial\Omega = \overline{\Omega} \Big\backslash \mathring{\Omega}$ est un fermé borné, donc compact[6], donc f (qui est continue) y atteint bien ses infimum et supremum.

1. Soit par l'absurde un $a \in \Omega$ tel que $f(a) \geq \max f_{|\partial\Omega}$. Par compacité du fermé borné $\overline{\Omega}$, f atteint son supremum sur $\overline{\Omega}$ en un point μ ; si $f(\mu) > f(a)$, alors $\mu \in \Omega$, sinon on peut prendre $\mu = a$; dans tous les cas, on a un maximum en $\mu \in \Omega$, donc la hessienne en μ est négative, donc de trace $\Delta f \leq 0$, ce qui est contraire aux hypothèses.

2. On se ramène à la question précédente en perturbant f. Pour un $\varepsilon > 0$, $f_\varepsilon := f + \varepsilon \sum x_i^2$ vérifie $\Delta f_\varepsilon = \Delta f + 2n\varepsilon > 0$, d'où pour tout point $a \in \Omega$
$$\begin{aligned} f_\varepsilon(a) &< \max_{\partial\Omega} f_\varepsilon \leq \max_{\partial\Omega} f + \max_{\partial\Omega}\left(\varepsilon \sum x_i^2\right) \\ f(a) + \varepsilon \sum a_i^2 &< \max_{\partial\Omega} f + \varepsilon M^2 \text{ où } M \text{ majore } \Omega \\ \text{et } f(a) &\leq \max_{\partial\Omega} f \text{ lorsque } \varepsilon \longrightarrow 0. \end{aligned}$$
Pour avoir le min, on applique à $-f$.

3. Puisque $f_{|\partial\Omega}$ est constante, on a l'égalité $\max_{\partial\Omega} f = \min_{\partial\Omega} f$, d'où $f_{|\Omega}$ constante coincée entre ces deux valeurs.

[6] Le lecteur aura remarqué la correspondance $\left\{ \begin{array}{l} \overline{\Omega} \longleftrightarrow K \\ \Omega \longleftrightarrow \mathring{K} \end{array} \right.$ avec les notations de l'exercice précédent.

7 Conditions d'holomorphie de Cauchy-Riemann

Soit $f : \mathbb{C} \longrightarrow \mathbb{C}$ une application vue comme couple de fonctions de deux variables réelles

$$(x,y) \longmapsto \begin{pmatrix} X(x,y) \\ Y(x,y) \end{pmatrix}.$$

Montrer que f est dérivable[7] *en un complexe c ssi f est différentiable en c et vérifie*

$$\begin{cases} [\partial_1 X](c) = [\partial_2 Y](c) \\ [\partial_2 X](c) = -[\partial_1 Y](c) \end{cases}.$$

Solution proposée.
Pour $h := x + iy$ proche de zéro, on peut écrire

$$f(c+h) - f(c) = [X(a+x, b+y) - X(a,b)] + i[Y(a+x,b+y) - Y(a,b)].$$

1. Supposons f dérivable en c. Alors le terme de droite s'écrit $hf'(c) + o(h)$, soit encore (avec des notations évidentes)

$$(x+iy)(\lambda + i\mu) + o(h) = (x\lambda - y\mu) + i(x\mu + y\lambda) + o(h).$$

Prenant la partie réelle, il vient

$$X(a+x, b+y) - X(a,b) = (x\lambda - y\mu) + \underbrace{\operatorname{Re} o(h)}_{=o(h)}.$$

Faire $y = 0$ montre que $\partial_1 X(c)$ existe[8] et vaut λ, tandis que faire $x = 0$ donne $\partial_2 X(c) = -\mu$.

Prenant à présent la partie imaginaire, on obtient

$$Y(a+x, b+y) - Y(a,b) = x\mu + y\lambda + o(h).$$

On dérive de même $\partial_1 Y(c) = \mu$ et $\partial_2 Y(c) = \lambda$, d'où les conditions cherchées.

2. Supposons X et Y différentiables en c et vérifiant les égalités de l'énoncés, mettons $\begin{cases} \partial_1 X(c) = \lambda = \partial_2 Y(c) \\ \partial_1 Y(c) = \mu = -\partial_2 X(c) \end{cases}$.
On peut alors écrire

$$\begin{aligned} f(c+h) - f(c) &= [X(a+x,b+y) - X(a,b)] + i[Y(a+x,b+y) - Y(a,b)] \\ &= [\lambda x - \mu y + o(h)] + i[\mu x + \lambda y + o(h)] \\ &= x(\lambda + i\mu) + y(-\mu + i\lambda) + o(h) \\ &= (x+iy)(\lambda + i\mu) + o(h), \end{aligned}$$

d'où le résultat escompté avec $f'(c) = \lambda + i\mu$.

Remarque. La dérivation complexe est bien plus contraignante que la dérivation réelle. On peut en effet montrer qu'une fonction complexe dérivable (au sens complexe) sur un ouvert de $\mathbb{C}$ est non seulement de classe C^2, mais aussi C^∞ et même développable en série entière sur son domaine de définition. Une telle fonction est dite *holomorphe*[9] ; il y aurait beaucoup à en dire[10].

---

[7] Il s'agit de la dérivation au sens complexe : $f'(c)$ est défini par la limite $\frac{f(c+h)-f(c)}{h}$ lorsque le *complexe* h tend vers 0 dans $\mathbb{C}^*$.

[8] Bien noter que le $o(h)$ devient un $o(\alpha)$ lorsque $\beta = 0$.

[9] du grec *holos* signifiant « entier », pour évoquer le développement en série **entière**

[10] On se reportera aux exercices sur les séries entières pour quelques résultats supplémentaires qui n'ont nullement la prétention d'introduire au sujet.

8 Quelques résultats sur les fonctions holomorphes

1. *Soit f une fonction holomorphe. Montrer que, si f est constante sur le bord $\partial\Omega$ d'un ouvert borné du plan complexe, alors f est constante sur cet ouvert.*
2. *En déduire que deux fonctions holomorphe qui coïncident sur le bord d'un ouvert (borné) sont égales sur cet ouvert.*

Solution proposée.

1. Soit $f = X + iY$ comme dans l'énoncé. Au vu de l'exercice précédent et du lemme de Schwarz, la fonction X (qui est de classe C^2 d'après la remarque de l'exercice précédent) est harmonique :

$$\Delta X = \partial_{11}^2 X + \partial_{22}^2 X = \partial_1 (\partial_2 Y) + \partial_2 (-\partial_1 Y) = 0.$$

   On peut donc appliquer le résultat de l'avant-dernier exercice : X est constante sur Ω. On fait évidemment la même chose pour Y, d'où le résultat.
2. On prend la différence : nulle au bord, donc nulle partout !

Remarque. Nous venons de voir que les fonctions holomorphes sont harmoniques. On peut montrer réciproquement qu'une fonction harmonique sur $\mathbb{R}^2$ est localement la partie réelle d'une fonction holomorphe. Ceci explique les propriétés communes à des deux classes de fonctions.

9 Lemme d'Hadamard

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe C^∞. On notera abusivement x_i l'application $x \mapsto x_i$.

1. *Si $f(0) = 0$, montrer que*

$$f = \sum_i x_i f_i$$

   où les fontions f_i sont C^∞.
2. *Si de plus $d_0 f = 0$, montrer que*

$$f = \sum_{i,j} x_i x_j f_{i,j}$$

   où les fontions $f_{i,j}$ sont C^∞.

Solution proposée.

1. On applique Taylor-Lagrange avec reste intégral à la fonction $F : t \mapsto f(tx)$:

$$f(x) = F(1) = F(0) + \int_0^1 F'(t)\, dt = \int_0^1 \sum \partial_i f(tx)\, x_i dt = \sum x_i \int_0^1 \partial_i f(tx)\, dt.$$

   En posant $f_i(x) = \int_0^1 \partial_i f(tx)\, dt$, qui est C^∞ car l'intégrande l'est, on a gagné.
2. Si de plus $d_0 f = 0$, les dérivées partielles de f en 0 sont nulles, ce qui s'écrit

$$0 = \partial_k \left(\sum x_i f_i \right)(0) = f_k(0) + 0.$$

   On peut donc appliquer ce qui précède aux f_i :

$$f(x) = \sum_i x_i f_i(x) = \sum_i x_i \sum_j x_j f_{i,j}(x) = \sum_{i,j} x_i x_j f_{i,j}(x)$$

   avec les $f_{i,j}$ qui sont C^∞

10 Dérivations de germes C^∞

Soit G l'ev des fonctions définies sur un voisinage[11] de 0 de $\mathbb{R}^n$, à valeurs réelles et de classe C^∞.

On appelle *dérivation* sur G toute forme linéaire δ vérifiant pour tout $f,g\in G$

$$\delta(fg) = f(0)\,\delta(g) + g(0)\,\delta(f).$$

1. *Montrer que dériver (en 0) selon un vecteur donné est une dérivation.*
2. *Montrer la réciproque. (On pourra utiliser le lemme d'Hadamard.)*

Solution proposée.

1. Soit v le vecteur considéré et $\delta : f \mapsto f'_v(0) = [d_0 f](v)$. On vérifie que, pour $f,g\in G$:

$$\begin{aligned}\delta(fg) &= [d_0(fg)](v) = [d_0 f\cdot g(0) + f(0)\cdot d_0 g](v)\\ &= g(0)[d_0 f](v) + f(0)[d_0 g](v)\\ &= g(0)\,\delta(f) + f(0)\,\delta(f),\ \textit{CQFD.}\end{aligned}$$

2. Faire $f = g \equiv 1$ donne $\delta(1) = 2\delta(1)$, d'où $\delta(1) = 0$.
Suivons l'énoncé : d'après Hadamard, f s'écrit

$$f = f(0) + \sum x_i f_i$$

où x_i dénote abusivement l'application $x\mapsto x_i$ et où $f_i\in G$. Appliquant δ et utilisant ses propriétés, il vient :

$$\delta(f) = f(0)\,\delta(1) + \sum 0 + f_i(0)\,\delta(x_i) = \sum \delta(x_i)\frac{\partial f}{\partial x_i}(0).$$

En posant $v := (\delta(x_1),...,\delta(x_n))$, on obtient

$$\delta(f) = \sum \frac{\partial f}{\partial x_i}(0)\,v_i = d_0 f(v) = f'_v(0),\ \textit{CQFD.}$$

11 Isométries et différentielles

Soit $E := \mathbb{R}^n$ muni de sa structure d'espace euclidien et $f : E \longrightarrow E$ de classe C^1 telle que la différentielle en tout point soit une isométrie de E. On veut montrer que f est une isométrie affine.

1. *Étant donné un point $a\in E$, montrer qu'il y un ouvert U_a où*

$$\|f(x) - f(y)\| = \|x-y\| \text{ pour tous } x,y\in U_a.$$

2. *Montrer que*

$$\forall h,k\in E,\ \langle d_x f(h) \mid d_y f(k)\rangle = \langle h \mid k\rangle \text{ pour tous } x,y\in U_a.$$

En déduire que df est constante sur U_a.

3. *Conclure.*

Solution proposée.

[11] On parle de *germe* en 0, d'où la notation.

1. En norme triple, $|||d_a f||| = \sup_{x \neq 0} \frac{\|d_a f(x)\|}{\|x\|} = 1$ car $d_a f$ est une isométrie, donc l'inégalité des accroissements finis s'applique :

$$\|f(x) - f(y)\| \leq \|x - y\| \text{ pour tous } x, y \in E.$$

On voudrait pouvoir retourner cette inégalité pour avoir égalité, ce qui sera possible si l'on arrive à inverser f. Avanti.

Soit $a \in E$. $d_a f$ est une isométrie, donc injective, donc f est un difféomorphisme local en a, disons $f : U_a \longrightarrow V_a$ difféo. Quitte à restreindre U_a, on peut supposer V_a connexe afin de lui appliquer l'inégalité des accroissements finis. Les différentielles de f^{-1} sont les inverses de celles de f, donc sont encore des isométries, donc de norme 1 ; pour $f(x)$ et $f(y)$ dans V_a, on aura donc

$$\left\|f^{-1}(f(x)) - f^{-1}(f(y))\right\| \leq \|f(x) - f(y)\|,$$

d'où l'égalité souhaitée.

2. On différentie l'égalité $\|f(x) - f(y)\|^2 = \|x - y\|^2$ en x selon un vecteur h, puis en y selon un vecteur k :

$$\begin{aligned} 2\langle d_x(h) \mid f(x) - f(y)\rangle &= 2\langle h \mid x - y\rangle \\ \langle d_x(h) \mid -d_y f(k)\rangle &= \langle h \mid -k\rangle \\ \langle d_x(h) \mid d_y f(k)\rangle &= \langle h \mid k\rangle. \end{aligned}$$

Cela permet de comparer les isométries $d_x f$ et $d_y f$ en appliquant en un vecteur h :

$$\begin{aligned} \|d_x f(h) - d_y f(h)\|^2 &= \|d_x f(h)\|^2 - 2\langle d_x f(h) \mid d_y f(h)\rangle + \|d_y f(h)\|^2 \\ &= \|h\|^2 - 2\langle h \mid h\rangle + \|h\|^2 \\ &= 0. \end{aligned}$$

3. La question précédente montre que l'ensemble $\{a \in E \ ; \ d_a f = d_0 f\}$ est ouvert. Comme de plus f est C^1, ce dernier est également fermé (comme image réciproque du singleton $\{d_0 f\}$ par l'application continue df). Par connexité de E, on doit avoir $d_a f = d_0 f$ pour tout $a \in E$, donc la différentielle de $f - d_0 f$ est nulle sur tout E, donc $f - d_0 f$ est une constante. Finalement :

$$f = f(0) + d_0 f, \ \textit{CQFD}.$$

12 Du local au global

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ un C^1-difféomorphisme local tel que $\left\{ \begin{array}{c} f^{-1}(\{0\}) = \{0\} \\ \lim_\infty f = \infty \end{array} \right.$. Le problème est de montrer que f est un C^1-difféomorphisme *global*.

1. *Montrer que la préimage par f de tout compact est un compact.*
2. *Montrer que f est surjective.*

   On fixe un $y \in \mathbb{R}^n$.
3. *Montrer que y n'a qu'un nombre fini d'antécédents que l'on notera $x_1, ..., x_p$.*
4. *Pour tout $\varepsilon > 0$, montrer qu'il existe un voisinage V de y dont la préimage par f est recouverte par les p boules $x_i + \varepsilon\mathbb{B}$*
5. *Montrer que $a \mapsto \operatorname{Card} f^{-1}(\{a\})$ est localement constante.*
6. *Conclure.*

Solution proposée.

1. Soit K un compact de $\mathbb{R}^n$. Sa préimage est fermé car f est continue. Il reste à montrer qu'elle est bornée. Or, K est borné par un certain M. L'hypothèse $\lim_\infty f = \infty$ permet par ailleurs d'affirmer l'existence d'un $A > 0$ tel que $\|x\| > A \implies \|f(x)\| > M \implies x \notin K$, de sorte que $f^{-1}(K)$ est borné par A.
2. Montrons que $\operatorname{Im} f$ est ouvert et fermé dans $\mathbb{R}^n$. Comme ce dernier est connexe, nécessairement $\operatorname{Im} f$ vaudra $\mathbb{R}^n$ (ou le vide, mais cela est trivialement à exclure). $d_a f$ étant inversible pour tout $a \in \mathbb{R}^n$, f est ouverte, donc $f(\mathbb{R}^n)$ est ouvert. Montrons que ce dernier est également fermé. Soit $b = \lim f(a_n)$. La réunion des a_n et de b forment un compact, donc sa préimage est un compact contenant tous les a_n, d'où une sous-suite convergente $a_{\varphi(n)} \longrightarrow a$. La continuité de f permet de conclure $b = \lim f\left(a_{\varphi(n)}\right) = f(a) \in \operatorname{Im} f$.
3. Il est déjà clair que les antécédents de y sont tous isolés : en effet, si x est l'un deux, f est un C^1-difféomorphisme local en x, donc injective sur un voisinage de x. Or, la préimage du compact $\{y\}$ est compacte d'après le premier point, donc ne peut être infinie, sinon on en extrairait une suite convergente vers un point d'accumulation de $f^{-1}(\{y\})$.
4. Supposons par l'absurde que la préimage $f^{-1}\left(y + \frac{1}{n}\mathbb{B}\right)$ de toute boule centrée en y n'est pas incluse dans $\bigcup_{f(a)=y} a + \varepsilon\mathbb{B}$. On a donc un a_n dans $f^{-1}\left(y + \frac{1}{n}\overline{\mathbb{B}}\right)$, *i. e.* tel que $\|f(a_n) - y\| < \frac{1}{n}$, qui se trouve à distance au moins ε de tous les antécédents de y. Quitte à extraire du compact $f^{-1}\left(y + \overline{\mathbb{B}}\right)$, on peut supposer (a_n) convergente vers un $a \in \mathbb{R}^n$, lequel doit vérifier $f(a) = y$ (par continuité de f) ; le point a est donc un antécédent de f. Mais puisque $a_n \longrightarrow a$, la suite a_n tombe dans $a + \varepsilon\mathbb{B}$ à partir d'un certain rang, ce qui est impossible.
5. f étant un difféomorphisme local en les x_i, il y a un voisinage V_i de chaque x_i où f réalise un difféomorphisme. Quitte à réduire ε et les V_i, on peut supposer les $V_i = x_i + \varepsilon\mathbb{B}$ disjoints. Alors tout b dans $\bigcap f(V_i)$ admet exactement p antécédents dans $\bigcup V_i$ (considérer les $f_{V_i}^{-1}$). Pour éviter d'en trouver d'autres, il faudrait que tous les antécédents de b tombent dans $\bigcup V_i$, ce qui sera réalisé en remplaçant $\bigcap f(V_i)$ par son intersection avec $\bigcup V_i$ (qui est bien ouvert comme intersercction finie d'ouverts). Sur ce voisinage de y, le nombre d'antécédent est constant, *CQFD*.
6. Nous venons de montrer que les parties
$$U_k := \left\{a \in \mathbb{R}^n \ ; \ \operatorname{Card} f^{-1}(\{a\}) = k\right\}$$
sont ouvertes. Comme elles partitionnent le connexe $\mathbb{R}^n$, nécessairement elles sont toutes vides sauf une. Mais U_1 contient 0 par hypothèse, ce qui montre $U_1 = \mathbb{R}^n$, d'où l'injectivité de f et la conclusion.

Remarque. L'aspect différentiel intervient ici de très loin. Le bon énoncé serait plutôt « Un homéomorphisme local $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ propre[12] est un homéomorphisme global ». Sous certaines hypothèses de connexité (vérifiées pour $\mathbb{R}^n$), on peut montrer qu'un homéomorphisme local est global ssi il est propre[13].

13 Théorème des extrema liés

Soit à trouver les extrema d'une fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ sur une ligne de niveau du plan, par exemple sur l'ellipse d'équation $\frac{x^2}{4} + y^2 = 1$.

L'idée est de faire varier les lignes de niveau de la fonction f jusqu'à trouver un point de contact avec le lieu où l'on cherche les extrema. Puisque le gradient est toujours orthogonal aux lignes de niveau, en ce point de tangence des deux lignes de niveau, les gradients de h et φ seront colinéaires.

Prenons par exemple $h : (a,b) \mapsto ab$. On cherche les points P de notre ellipse tels qu'il y a réel λ vérifiant $\overrightarrow{\nabla}_P h = \lambda \overrightarrow{\nabla}_P e$ avec $e : (a,b) \mapsto \frac{a^2}{4} + b^2$. On trouve $\left\{\begin{array}{l} b = \lambda\frac{a}{2} \\ a = \lambda 2b \end{array}\right.$, d'où $\lambda^2 = 1$ et $a = \pm 2b$, ce qui donne (en réinjectant dans l'équation de l'ellipse) $b = \pm\frac{1}{\sqrt{2}}$ puis $a = \pm\sqrt{2}$, d'où les extrema $h(a,b) = \pm 1$.

Pour nous conforter dans notre intuition, retrouvons le maximum de h directement par une inégalité $2uv \leq u^2 + v^2$: pour a et b de même signe, on aura
$$h(a,b) = ab = 2\frac{a}{2}b \leq \left(\frac{a}{2}\right)^2 + b^2 = 1$$
avec $=$ ssi $\frac{a}{2} = b$.[3]

[12] Une application est dit *propre* lorsque la préimage de tout compact est compacte.
[13] Le lecteur intéressé pourra consulter *A Note on Proper Maps* de Chung-Wu Ho, Proc. Amer. Math. Soc. 51 (1975), 237-241.

De la même manière, si l'on cherche les extrema d'une application $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ sur la courbe définie par l'intersection de deux surfaces de niveau, les trois gradients seront orthogonaux à la courbe, donc coplanaires. Si le lieu est décrit par deux équations $u = v = 0$, on pourra donc trouver deux réels λ et μ tels que $\overrightarrow{\nabla} f = \lambda \overrightarrow{\nabla} u + \mu \overrightarrow{\nabla} v$, sous réserve que $\overrightarrow{\nabla} u$ et $\overrightarrow{\nabla} v$ ne soient pas liés[14].

Nous nous proposons de montrer proprement notre intuition en termes différentiels, le pont entre gradients et différentielles que nous emprunterons étant l'équivalence

$$\overrightarrow{\nabla}_a f \in \operatorname{Vect}\left\{\overrightarrow{\nabla} u, \overrightarrow{\nabla} v, \overrightarrow{\nabla} w\right\} \iff d_a f \in \operatorname{Vect}\left\{d_a u, d_b v, d_a w\right\}.$$

Soit f une fonction réelle définie sur $\mathbb{R}^n$ de classe C^1. On cherche les extrema de f sur le lieu

$$\Gamma := \left\{a \in \mathbb{R}^n \ ; \ g^1(a) = \cdots = g^r(a) = 0\right\}$$

où les g^i sont aussi des fonctions réelles sur $\mathbb{R}^n$ de classe C^1.

Soit a un tel extremum. Si les $d_a g^i$ sont libres, on veut montrer qu'ils engendrent $d_a f$.

1. *Montrer que la jacobienne de f au point a est de rang r. En déduire un « cassage » de $\mathbb{R}^{n-r} \times \mathbb{R}^r$ selon deux coordonnées.*
2. *Montrer que la fonction*

$$G := \left(g^1, ..., g^r\right) : \left\{ \begin{array}{ccc} \mathbb{R}^{n-r} \times \mathbb{R}^r & \longrightarrow & \mathbb{R}^r \\ z & \longmapsto & \left(g^1(z), ..., g^r(z)\right) \end{array} \right.$$

vérifie les hypothèses du théorème des fonctions implicites en $a =: (\alpha, \beta)$.

On peut donc paramétrer Γ par $\gamma(t) := (t, \varphi(t))$ où φ est de classe C^1 vérifiant $\varphi(\alpha) = \beta$.

3. *Montrer les inclusions*

$$\bigcap \operatorname{Ker} d_a g^j = \operatorname{Im} d_\alpha \gamma \subset \operatorname{Ker} d_a f$$

et conclure.

Solution proposée.

1. Rappelons que la matrice jacobienne $\left(\partial_i g^j(a)\right)$ des g^j au point a n'est autre que la matrice des $d_a g^i$ dans la base des dx_i. Par liberté des $d_a g^i$, le rang de cette matrice est r, donc on peut en extraire r lignes libres. Les indices des lignes $i_1, ..., i_r$ ainsi sélectionnées fournissent un cassage de $\mathbb{R}^n \simeq \mathbb{R}^{n-r} \times \mathbb{R}^r$. On suppose par commodité que $i_k = k$ pour tout $k = 1, ..., r$.
2. On a déjà $G(\alpha, \beta) = 0$. De plus, étant donné un vecteur $\varepsilon \in \mathbb{R}^r$, on regarde $G(\alpha, \beta + \varepsilon) - G(\alpha, \beta)$. Pour cela, on regarde les $g^k(\alpha, \beta + \varepsilon) - g^k(\alpha, \beta)$. Les fonctions g^k étant C^1, les fonctions $g^k(\alpha, \cdot)$ sont aussi C^1 et on a

$$\forall 1 \leq k \leq r, \ g^k(\alpha, \beta + \varepsilon) = \sum_{i=1}^{r} \partial_i g^1(a) \varepsilon_i + o(\varepsilon),$$

de sorte que

$$G(\alpha, \beta + \varepsilon) - G(\alpha, \beta) = \sum_{i=1}^{r} \left(\partial_i g^1(a), ..., \partial_i g^r(a)\right) \varepsilon_i + o(\varepsilon).$$

G est donc différentiable en a selon sa seconde coordonnée, et la matrice de sa différentielle est la jacobienne en a qui est inversible.

3. Puisque $\gamma(t) \in \Gamma$, on a $g^j \circ \gamma = 0$ pour tout j, d'où, en différentiant en α,

$$0 = d_{\gamma(\alpha)} g^j \circ d_\alpha \gamma = d_a g^j \circ d_\alpha \gamma,$$

ce qui montre l'inclusion

$$\operatorname{Im} d_\alpha \gamma \subset \bigcap \operatorname{Ker} d_a g^j.$$

[14] La liberté de $\overrightarrow{\nabla} u$ et $\overrightarrow{\nabla} v$ implique l'unicité des réels λ et μ qui apparaissent : ces derniers sont appelés *multiplicateurs de Lagrange*.

On a par ailleurs l'égalité des dimensions :

$$\left\{\begin{array}{c} \operatorname{rg} d_\alpha \gamma = \operatorname{rg}\left(\operatorname{Mat} d_\alpha \gamma\right) = \operatorname{rg}\begin{pmatrix} I_{n-r} \\ * \end{pmatrix} = n-r \\ \dim \bigcap \operatorname{Ker} d_a g^j = \operatorname{corg} d_a g^j = n-r \end{array}\right. .$$

D'autre part, $f \circ \gamma$ étant extrémal en α, sa différentielle en α est nulle, *i. e.* $\operatorname{Im} d_\alpha \gamma \subset \operatorname{Ker} d_a f$ comme ci-dessous, d'où l'égalité

$$\bigcap \operatorname{Ker} d_a g^j \subset \operatorname{Ker} d_a f.$$

Le résultat souhaité se déduit par dualité, en utilisant que $(\operatorname{Ker} \varphi)^\perp = \mathbb{R}\varphi$ pour toute forme linéaire φ :

$$\begin{array}{rcl} \bigcap \operatorname{Ker} d_a g^j & \subset & \operatorname{Ker} d_a f \\ \Longleftrightarrow (\operatorname{Ker} d_a f)^\perp & \subset & \left(\bigcap \operatorname{Ker} d_a g^j\right)^\perp \\ \Longleftrightarrow \mathbb{R} d_a f & \subset & \sum \left(\operatorname{Ker} d_a g^j\right)^\perp \\ \Longleftrightarrow \mathbb{R} d_a f & \subset & \sum \mathbb{R} d_a g^j, \text{ \textit{CQFD.}} \end{array}$$

Remarque 1. Une interprétation géométrique de cet énoncé est la suivante.
Voyons Γ comme une surface. Au voisinage de a, les g^i s'approximent bien par leurs applications tangentes $d_a g^i$, de sorte que les équations $g^i = 0$ définissant Γ permettent d'approximer ce dernier (autour de a) par le « plan tangent » $\bigcap \operatorname{Ker} d_a g^i$.
Traçons maintenant un chemin c sur Γ passant par $c(0) = a$. La dérivée en 0 de c est un vecteur tangent u à Γ en a. Comme $f \circ c$ est extrémal en 0, sa dérivée doit s'annuler, d'où $u \in \operatorname{Ker} d_a f$. Or, en faisant varier c, on peut obtenir tous les vecteurs u du plan tangent, d'où l'inclusion

$$\bigcap \operatorname{Ker} d_a g^j \subset \operatorname{Ker} d_a f.$$

Remarque 2. Ce théorème est sans doute puissant, mais il faut d'une part prendre garde aux hypothèses qui apparaissent lorque f est définie sur autre chose qu'un ouvert (l'extremum doit être *intérieur* au domaine de définition, ce qui oblige à regarder le bord à part) et d'autre part **connaître ses inégalités classiques**. Il est en effet intolérable qu'un taupin invoque les extrema liés pour maximiser le produit de réels positifs dont la somme est fixée (lorsqu'une inégalité arithmético-géométrique trivialise la chose). L'impression produite est celle du même taupin qui dérive la fonction $\theta \mapsto \frac{1-r^2}{r^2-2r\cos\theta+1}$ pour en étudier les variations (en r !) : arrêtons le massacre.

14 Théorème de la boule chevelue par le lemme de Milnor

On note $\mathbb{S}_n$ la sphère de dimension $n \geq 1$ vue dans $\mathbb{R}^{n+1}$. On s'intéresse aux champs tangents à $\mathbb{S}_n$, *i. e.* aux applications $f : \mathbb{S}_n \longrightarrow \mathbb{R}^{n+1}$ telles que $f(a) \perp a$ pour tout point $a \in \mathbb{S}_n$.
Le but est de montrer qu'un tel champ f, supposé C^1, s'annule forcément pour n pair.

1. *Montrer par l'absurde qu'on peut supposer f définie sur $\mathbb{R}^{n+1}$, conservant la norme, tout en restant tangent et C^1 sur l'adhérence de l'anneau*

$$A := \left\{ a \in \mathbb{R}^{n+1} \; ; \; \frac{1}{2} < \|a\| < \frac{3}{2} \right\}.$$

On perturbe ensuite l'identité par f : pour λ « petit » réel, on pose

$$f_\lambda = \operatorname{Id} + \lambda f.$$

2. *Montrer que, pour λ assez petit, f_λ est un difféomorphisme local sur A, puis global (**lemme de Milnor***

3. *Montrer que, pour de tels λ, l'image $f_\lambda(A)$ vaut*

$$A' := f_\lambda(A) = \sqrt{1+\lambda^2}A.$$

On pourra invoquer des arguments de connexité pour un sens.

4. *Conclure en prenant les volumes.*

Solution proposée.

1. Pour définir $f(a)$ où a se balade dans A, on se ramène naturellement à la sphère unité en considérant $f\left(\frac{a}{\|a\|}\right)$. Ensuite, pour atterrir dans la sphère unité, il suffit de normaliser, ce qui est possible puisque l'on a supposé par l'absurde le champ f sans zéros. Enfin, on homogénéise pour avoir la condition sur la norme. Finalement, on remplace $f(a)$ par

$$\widetilde{f}(a) = \|a\| \frac{f\left(\frac{a}{\|a\|}\right)}{\left\|f\left(\frac{a}{\|a\|}\right)\right\|}.$$

qui reste bien C^1 sur l'anneau A. On vérifie que $\widetilde{f}$ reste tangent :

$$\left\langle \widetilde{f}(a) \mid a \right\rangle = \left\langle \|a\| \frac{f\left(\frac{a}{\|a\|}\right)}{\left\|f\left(\frac{a}{\|a\|}\right)\right\|} \mid a \right\rangle = (\cdot)\left\langle f\left(\frac{a}{\|a\|}\right) \mid \frac{a}{\|a\|} \right\rangle = 0.$$

2. Soit $a \in A$: la différentielle de f_λ s'obtient aisément par

$$d_a f_\lambda = \mathrm{Id} + \lambda d_a f.$$

On veut qu'elle soit injective, *i. e.* que

$$\det(\mathrm{Id} + \lambda d_a f) \neq 0$$

Or, f étant C^1 sur $\overline{A}$, *df* est continue sur $\overline{A}$ donc bornée, donc le déterminant ci-dessus reste non nul (et même > 0) pour λ assez proche de 0.

Pour inverser globalement, il faut vérifier l'injectivité :

$$f_\lambda(a) = f_\lambda(b) \implies a - b = \lambda(f(b) - f(a)).$$

f étant continue sur le compact $\overline{A}$, elle y est bornée, donc le terme de droite est un $O(\lambda)$, d'où $a = b$ en faisant tendre λ vers 0.

3. Pour un point $a \in A$, on a

$$\begin{aligned} \|f_\lambda(a)\|^2 &= \|a + \lambda f(a)\|^2 = \|a\|^2 + 2\lambda \langle a \mid * \rangle + \lambda^2 \|f(a)\|^2 \\ &= \|a\|^2 + 0 + \lambda^2 \|a\|^2 \text{ car } f \text{ tangent et conserve la norme} \\ &= \left(1+\lambda^2\right)\|a\|^2, \end{aligned}$$

d'où l'inclusion $A' \subset \sqrt{1+\lambda^2}A$.

Pour l'inclusion réciproque, on observe que, puique f_λ est un difféomorphisme, A' est un ouvert de $\sqrt{1+\lambda^2}A$; ce dernier étant connexe, il suffit de montrer que A' est fermé **dans** $\sqrt{1+\lambda^2}A$.

Soit donc $f(a_n)$ une suite de A' convergent vers un $l \in \sqrt{1+\lambda^2}A$. Quitte à extraire du compact $\overline{A}$, on peut supposer $a_n \longrightarrow a \in \overline{A}$. On obtient $f(a) = l$, avec

$$\|a\| = \|f(a)\| = \|l\| \in \left]\frac{1}{2}, \frac{3}{2}\right[, \ i.\ e.\ a \in A.$$

Ceci montre que $l \in f(A) = A'$, *CQFD.*

4. Le volume de l'anneau A peut s'obtenir par une intégrale

$$\mathrm{vol}(A) = \int_A 1.$$

Or, un changement de variable via le difféomorphisme f_λ montre que

$$\operatorname{vol}\left(\sqrt{1+\lambda^2}A\right) = \operatorname{vol} f(A) = \int_{f(A)} 1 = \int_A |\det df_\lambda|\,,$$

d'où pour $\lambda > 0$ assez petit

$$\left(1+\lambda^2\right)^{\frac{n+1}{2}} \operatorname{vol}(A) = \int_{a\in A} \det\left(\operatorname{Id} + \lambda d_a f\right).$$

Le terme de droite est un polynôme en λ, tandis que celui de gauche ne l'est pas pour n pair ; ploum. C'est-y pas joli ?

Remarque. Pour $n = 2p-1$ impair, on peut toujours exiber un champ tangent sans zéros $f : \mathbb{S}_n \longrightarrow \mathbb{R}^{n+1} = \mathbb{R}^{2p}$ en faisant p copies du champ tangent au cercle :

$$f(x_1, ..., x_{2p}) = (x_2, -x_1, x_4, -x_3, ..., x_{2p}, -x_{2p-1}).$$

15 Méthode de Newton généralisée

On se donne un C^1-difféomorphisme $f : \overline{B} \longrightarrow \mathbb{R}^n$ où B désigne la boule ouverte de centre un point $a_0 \in \mathbb{R}^n$ et de rayon un réel $r > 0$. On définit[15] une suite (a_n) par la relation de récurrence

$$a_{n+1} := a_n - df_{a_n}^{-1}(f(a_n))$$

et on note α le réel $\|a_1 - a_0\|$. On suppose que a n'est pas un zéro de f, ce qui revient à dire, par injectivité de df_a, que $\alpha > 0$.

On suppose que, sur B, f est $2k$-lipschitzienne pour un $k > 0$ et df est bornée par un $M > 0$.
On fait de plus l'hypothèse $\alpha \leq \frac{1}{s+Mk}$ avec $rs = 1$.

1. *Montrer le lemme suivant :*

$$\forall a, b \in B,\ \|f(a) - f(b) - df_b(a-b)\| \leq k\|a-b\|^2.$$

2. *En déduire que tous les a_n sont dans B et que*

$$\|a_{n+1} - a_n\| \leq \alpha h^{2^n - 1}$$

pour un h à déterminer.

3. *Montrer que la suite a_n converge vers un zéro de f. Caractériser la convergence.*

Solution proposée.

1. Le segment $[a, b]$ restant dans B par convexité de ce dernier, on peut considérer l'application

$$\varphi(t) := f(b + t(a-b)) \text{ pour } 0 \leq t \leq 1.$$

Sa dérivée est donnée par

$$\varphi'(t) = df_{b+t(b-a)}(a-b).$$

[15] C'est l'analogie avec la méthode de Newton pour les polynômes pour approcher rapidement les zéros (simples). On part d'une abscisse a, on suit la tangente au point $(a, P(a))$ qui coupe l'axe des abscisses en un point b, puis on recommence. Cela définit une suite par la formule

$$a_{n+1} := a_n - \frac{P'(a_n)}{P(a_n)}.$$

On peut montrer que cette suite, sous de bonnes hypothèses, converge quadratiquement vers une racine de P.

La quantité qui nous intéresse s'écrit donc

$$f(a) - f(b) - df_b(a-b) = \varphi(1) - \varphi(0) - \varphi'(0).$$

Pour utiliser l'hypothèse lipschitzienne sur df, qui est reliée à φ', on fait ne fait apparaître que du φ' :

$$\|\varphi(1) - \varphi(0) - \varphi'(0)\| = \left\| \int_0^1 [\varphi'(t) - \varphi'(0)]\, dt \right\| \leq \int_0^1 \|\varphi'(t) - \varphi'(0)\|\, dt.$$

Par ailleurs, l'intégrande se majore par

$$\begin{aligned}
\|\varphi'(t) - \varphi'(0)\| &= \left\| df_{b+t(b-a)}(a-b) - df_b(a-b) \right\| \\
&\leq |||df(b+t(b-a)) - df(b)|||\, \|a-b\| \\
&\leq \gamma \|t(b-a)\|\, \|a-b\| \\
&= t\|a-b\|^2,
\end{aligned}$$

d'où le lemme en intégrant pour $0 \leq t \leq 1$.

2. Montrons par récurrence que $a_n \in B$ et $\|a_{n+1} - a_n\| < \alpha h^{2^n - 1}$.
Pour $n = 0$, c'est trivial.
Supposons le résulat acquis pour un $n - 1 \geq 0$. On cherche à majorer

$$\|a_{n+1} - a_n\| = \left\| -df_{a_n}^{-1}(f(a_n)) \right\| \leq M\|f(a_n)\|.$$

Afin d'utiliser le lemme, on peut faire apparaître la quantité nulle

$$0 = df_{a_{n-1}}(0) = df_{a_{n-1}}\left[a_n - a_{n-1} + df_{a_{n-1}}^{-1}(f(a_{n-1}))\right] = f(a_{n-1}) + df_{a_{n-1}}(a_n - a_{n-1}).$$

On obtient

$$\|f(a_n)\| = \left\|f(a_n) - df_{a_{n-1}}(0)\right\| = \left\|f(a_n) - f(a_{n-1}) - df_{a_{n-1}}(a_n - a_{n-1})\right\| \leq k\|a_n - a_{n-1}\|^2,$$

d'où

$$\|a_{n+1} - a_n\| \leq M\|f(a_n)\| \leq Mk\|a_n - a_{n-1}\|^2 \leq Mk\left(\alpha h^{2^{n-1}-1}\right)^2 = \frac{\alpha Mk}{h}\alpha h^{2^n - 1}.$$

La majoration souhaitée en découle pour $h := \alpha Mk \leq \frac{1}{s+Mk}Mk < 1$.

Il reste à montrer que $a_n \in B$, *i. e.* que $\|a_n - a_0\| < r$. Il suffit d'utiliser les majorations précédemment montrées

$$\|a_n - a_0\| \leq \sum_{i=1}^n \|a_i - a_{i-1}\| \leq \sum_{i=1}^n \alpha h^{2^{i+1}-1} < \alpha \sum_{i \geq 0} h^i = \frac{\alpha}{1-h} = \frac{1}{\frac{1}{\alpha} - Mk} \leq \frac{1}{(s+Mk) - Mk} = r.$$

3. La série $\sum \|a_{n+1} - a_n\|$ converge absolument, donc simplement, *i. e.* (a_n) converge vers un $a \in \overline{B}$. En prenant la limite dans la ligne ci-dessus, on trouve

$$\|a - a_0\| < \frac{\alpha}{1-h} \leq r,$$

ce qui montre que a reste dans B. Or, df^{-1} étant continue sur cette dernière, on peut prendre la limite dans la définition de la suite a_n, ce qui donne[16]

$$a = a - df_a^{-1}(f(a)) \implies df_a^{-1}(f(a)) \implies f(a) = 0.$$

[16] Pour être propre, il faudrait justifier pourquoi, lorsqu'on se donne une application continue $x \mapsto \varphi_x$ à valeurs dans $\mathcal{L}_c(E,F)$, on peut écrire $\lim_{x \to a} \varphi_x(x) = \varphi_a(a)$. Il suffit de découper comme pour la démonstration de la continuité du produit :

$$\begin{aligned}
\|\varphi_x(x) - \varphi_a(a)\| &\leq \|[\varphi_x - \varphi_a](x)\| + \|\varphi_a(x-a)\| \\
&\leq \underbrace{\|\varphi_x - \varphi_a\|}_{\longrightarrow 0}\underbrace{\|x\|}_{\text{borné}} + \underbrace{\|\varphi_x\|}_{\text{borné}}\underbrace{\|x-a\|}_{\longrightarrow 0}.
\end{aligned}$$

Quant à la convergence, on a majore

$$\begin{aligned}\|a_n - a\| &\leq \sum_{i\geq n} \|a_{i+1} - a_i\| = \alpha \sum_{i\geq n} h^{2^{i+1}-1} = \alpha h^{2^n-1} \sum_{i\geq n} h^{2^n\left(2^{i-n}-1\right)} \\ &\leq \alpha h^{2^n-1} \sum_{i\geq 0} \left(h^{2^n}\right)^i = \alpha h^{2^n-1} \frac{1}{1-h^{2^n}} = O\left(h^{2^n}\right).\end{aligned}$$

On obtient ainsi une convergence *quadratique*.

Remarque. Ce résultat est en pratique peu applicable, l'égalité $\alpha < \frac{1}{\frac{1}{r}+kM}$ étant difficilement vérifiable.