Calcul différentiel

On s'intéresse à des applications définies d'un ouvert de E, evn de dimension finie, dans F idem.
Le plus souvent, on considèrera les ev standards $\mathbb{R}^n, \mathbb{R}^p$ avec $n, p \leqslant 3$.

1 Continuité, dérivabilité

Il n'est pas évident de définir ce que l'on souhaite entendre par « continu, dérivable », dans le cas d'une application à plusieurs variables. Le point difficile est que l'on souhaite exprimer des propriétés vraies dans un **voisinage** (càd une boule ouverte) d'un point de U, alors que les méthodes naturelles consistent à se limiter à une direction — ou une coordonnée.

1.1 Continuité

DÉFINITION 1. *L'application f est continue en a ssi $f(x) \to f(a)$ quand $x \to a$.*

Il s'agit donc de la définition usuelle de la continuité dans un evn.

PROPOSITION.
Soient $(f_1, \ldots, f_p)$ les composantes *de f, c'est à dire que dans une base appropriée on a $f = f_1.e_1 + \ldots + f_p.e_p$.*
Alors f est continue $\iff$ toutes les f_i le sont.

En revanche, il ne sert de rien de considérer coordonnées par coordonnées **dans l'espace de départ** :

L'application $(x, y) \mapsto \begin{cases} 0 & \text{si } x = y = 0 \\ \dfrac{x^2 y}{x^4 + y^2} & \text{sinon} \end{cases}$ est continue par rapport à x et à y, mais elle est discontinue en l'origine.

> La continuité c'est quand on tend vers x de toutes les façons possibles ! Rappelons le critère séquentiel : f est continue en x si **pour toute suite** (x_n) tendant vers x on a la suite des images $(f(x_n))_n$ qui converge.

Néanmoins, on a tous les théorèmes usuels sur la continuité dans les evn :

PROPOSITION. *Tout « cocktail » d'applications continues (somme, produit, composée) est continu.*

ce qui permet d'assurer la continuité de toute combinaison de fonctions usuelles là où elle est définie.

1.2 Différentiabilité

La généralisation naturelle de la dérivabilité passe par le développement limité à l'ordre 1 :

1.2.1 Application linéaire tangente

> **DÉFINITION 2.** *On dit que f est différentiable en $a \in U$ ssi il existe une application **linéaire***
>
> $$d f_a \in \mathcal{L}(E, F) \quad \textit{telle que} \quad f(x) = f(a) + d f_a(x - a) + o(x - a)$$
>
> *pour x voisin de a. De façon équivalente :*
>
> $$f(a + h) = f(a) + d f_a(h) + o(h).$$
>
> $d f_a$ *s'appelle la* différentielle de f en a, *ou encore* l'application linéaire tangente.

Rappelons que $o(h)$ signifie $\|h\| \times \varepsilon(h)$ où $\varepsilon(h)$ est une fonction qui tend vers 0 avec h. Par exemple, si $h = (a, b)$ on a

$$ab = O(a^2 + b^2) = O(\|h\|^2 = o(\|h\|) \quad \text{(pour toute norme)}$$

Quand une application est différentiable, elle est donc localement « proche » d'une application affine (son graphe ressemblera donc à une variété affine, etc).

Exemple : On considère l'application $A \mapsto A^{-1}$ (dans le groupe des matrices inversibles, $\mathcal{G}\ell_n(\mathbb{R})$. Elle est continue car ses composantes sont des fractions rationnelles en les termes de A. Est-elle différentiable ?
On peut essayer de calculer

$$(A+H)^{-1} - A^{-1} = (A(I + A^{-1}H))^{-1} - A^{-1} = (I + A^{-1}H)^{-1}A^{-1} - A^{-1}$$

Or on peut faire le cas particulier du DL plus simple :

$$(I+U)^{-1} = I - U + o(U) \quad \text{quand } U \to 0.$$

Démonstration.

$$(I+U) \times \big((I+U)^{-1} - (I-U)\big) = U^2 \Rightarrow (I+U)^{-1} - (I-U) = (I+U)^{-1}U^2$$

et ceci est un $o(U)$ (et même un $O(\|U\|^2)$) car $(I+U)^{-1} \sim I$ quand $U \to 0$. ♦

Avec ce petit lemme, il reste (en posant $U = A^{-1}H$) :

$$(A+H)^{-1} - A^{-1} = (I - A^{-1}H + o(A^{-1}H))A^{-1} - A^{-1} = -A^{-1}HA^{-1} + o(H)$$

et on a bien notre application linéaire tangente (linéaire en la variable H) qui est $H \mapsto -A^{-1}HA^{-1}$.

EXERCICE 1. *Montrer à la main que l'application*

$$(x, y) \mapsto (x^2 + y^2, 3xy)$$

est différentiable au point (a, b).

PROPOSITION. *Toute application différentiable en* a *est continue en* a.

PROPOSITION. *Toute application linéaire est différentiable, et égale en tout point à sa différentielle.*

DÉFINITION 3. *Si* $d\,f_a$ *existe pour tout* $a \in \Omega$ *alors on dit que* f *est différentiable (tout court).*
Sa différentielle est l'application $d\,f : a \mapsto d\,f_a$, *qui va de* Ω *dans* $\mathcal{L}(E, F)$ *(!).*

Cette notion s'avère insuffisante pour les applications que nous envisageons, tout comme on ne s'est pas contenté des applications dérivables (de $\mathbb{R}$ dans $\mathbb{R}$), on est allé définir la classe $\mathcal{C}^1$.
Essayons donc de généraliser la notion de dérivabilité dans une autre direction :

1.2.2 Dérivée dans une direction

Soit $a \in U$; notons que si u est un vecteur quelconque (non nul), le segment $[a - \varepsilon u, a + \varepsilon u]$ est inclus dans U pour $\varepsilon > 0$ assez petit.
On peut donc définir une application vectorielle d'une variable réelle par

$$\phi_u(t) = f(a + tu)$$

pour t voisin de 0. Ce qui suffit à dériver ϕ_u :

DÉFINITION 4. f *admet en* a *une dérivée* ***selon le vecteur*** u *ssi* ϕ_u *est dérivable en 0 ; on pose alors*

$$D_u f(a) = \phi'_u(0)$$

L'idée est à la fois simple et profonde : on considère la **restriction** de f à un segment (passant par a) et on dérive cette restriction. On espère acquérir suffisamment de renseignements sur f en opérant ainsi dans toutes les directions possibles. (On peut aussi dériver la restriction de f à une courbe de l'espace, comme on le verra plus tard). Nous verrons que cela est à la fois trop optimiste et trop pessimiste :

— D'une part il suffit de dériver dans n directions indépendantes.
— D'autre part la dérivabilité dans ces (toutes les) directions ne suffira pas, nous exigerons que ces dérivées partielles soient continues.

Néanmoins il y a un premier lien entre dérivées directionnelles et différentielle :

THÉORÈME 1. *Si f est différentiable en a, alors elle est dérivable dans toute direction. De plus, on aura alors pour tout vecteur u*

$$D_u f(a) = d\,f_a(u)$$

Démonstration aisée :
$$\frac{f(a+tu)-f(a)}{t} = \frac{d\,f_a(tu)+o(tu)}{t} \to d\,f_a(u)$$
par linéarité de $d\,f_a$. L'idée est simple, si on sait faire un DL dans toute direction, *a fortiori* on sait le faire dans une direction fixée ! Examinons ce qui en résulte dans une base de E, soit $(e_1 \dots e_n)$:

DÉFINITION 5. *La dérivée $d\,f_a(e_j)$ dans la direction du vecteur de base e_j est la $j^{\text{ième}}$ dérivée partielle. C'est une fonction de E dans F, comme f. On la note $D_j f(a)$ ou $\dfrac{\partial f}{\partial x_j}(a)$ ou $\partial_j f$.*

REMARQUE 1. *Il s'agit de la dérivée de l'application* partielle

$$t \mapsto f(a_1, \dots a_{j-1}, a_j + \mathbf{t}, a_{j+1} \dots a_n) :$$

on fait bouger une seule variable à la fois.

REMARQUE 2. *La notation traditionnelle avec ∂ est ambigüe : que signifie $\dfrac{\partial f(u+v, u-v)}{\partial x_2}$?... la notation x_2 n'a pas d'autre vertu que de signaler le* NUMÉRO *de la variable de dérivation considérée. En ce sens la notation D_j ou ∂_j est plus pure.*

Par linéarité de $d\,f_a$, on déduit du théorème précédent la formule

PROPOSITION. *Si $h = (h_1, \dots h_n)$ on a $D_h f(a) = d\,f_a(h) = \sum_{j=1}^{n} h_j D_j f(a)$*

On a donc une formule de Taylor à l'ordre 1 :

$$f(a+h) = f(a) + d_a(h) + o(h) = f(a) + \sum_{j=1}^{n} h_j D_j f(a) + o(h)$$

EXERCICE 2. *Quel rapport y a-t-il entre $D_u f(a)$ et $D_{\lambda u} f(a)$?*

Le gros malheur est que la réciproque de ce théorème est fausse :

EXERCICE 3. *On définit* $f : (x,y) \mapsto \begin{cases} \dfrac{x^2 y}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } x = y = 0 \end{cases}$.

Montrer que f admet en l'origine des dérivées dans toutes les directions, sans être différentiable [hint : ses dérivées selon (Ox) et (Oy) étant nulles, une éventuelle différentielle serait nulle].

Faisons le point : on sait facilement calculer les dérivées partielles, mais leur existence ne suffit pas pour avoir la différentielle, qui seule permet de faire des DL c'est à dire de l'Analyse. Mais calculer une application linéaire tangente, c'est difficile ! Que faire ? Corser les hypothèses :

1.2.3 Classe $\mathcal{C}^1$

DÉFINITION 6. *f est de classe $\mathcal{C}^1$ sur U ssi pour tout vecteur $u \in E$, l'application $a \mapsto D_u f(a)$ est continue sur U — autrement dit, si toutes les dérivées partielles [dans n'importe quelle base] sont des applications continues.*

Ce n'est pas le cas dans l'exemple précédent.

THÉORÈME FONDAMENTAL.
Si il existe une base de E dans laquelle toutes les dérivées partielles sont continues, i.e. si f est de classe $\mathcal{C}^1$, alors f est différentiable.
On dit alors que f est **continûment différentiable.**

Démonstration. Supposons qu'il existe une base dans laquelle les dérivées partielles de f sont continues, et pour simplifier les notations considérons le cas d'une application à deux variables, à valeurs réelles : on suppose donc que $D_1 f$ et $D_2 f$ sont continues et on veut montrer que f est différentiable. Il en résultera que les dérivées directionnelles sont continues, puisque pour tout vecteur $u = (u_1, u_2)$ on a $D_u f = u_1 D_1 f + u_2 D_2 f$.
On écrit donc

$$f(x+h)-f(x) = f(x_1+h_1, x_2+h_2)-f(x_1, x_2) = \big(f(x_1+h_1, x_2+h_2)-f(x_1, x_2+h_2)\big)+\big(f(x_1, x_2+h_2)-f(x_1, x_2)\big).$$

Considérons la quantité $f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2 + h_2)$.
Elle est de la forme $g(h_1) - g(0)$ où $g(t) = f(x_1 + \mathbf{t}, x_2 + h_2)$ est de classe $\mathcal{C}^1$ (à une variable), par hypothèse.
On a donc par DL en 0 de la fonction g

$$f(x_1+h_1, x_2+h_2)-f(x_1, x_2+h_2) = g(h_1)-g(0) = h_1 g'(0)+o(h_1) = h_1 D_1 f(x_1, x_2+h_2)+o(h_1) = h_1 D_1 f(x_1, x_2)+o(h)$$

en utilisant la **continuité** de $D_2 f$ pour la dernière égalité : en effet $D_1 f(x_1, x_2 + h_2) \to D_1 f(x_1, x_2)$ quand $h_2 \to 0$.
De même et plus facilement, il vient

$$f(x_1, x_2 + h_2) - f(x_1, x_2) = h_2 D_2 f(x_1, x_2) + o(h_2)$$

En recollant, on a bien trouvé

$$f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2 + h_2) = \big(h_1 D_1 f + h_2 D_2 f\big)(x_1, x_2) + o(h)$$

donc f est différentiable et ite missa est. ♦

Cela signifie exactement que l'application $d\,f : a \mapsto d\,f_a$ est continue (combinaison linéaire des $D_j f(a)$).
Nous ne considérerons plus que des fonctions de ce type, au moins, pour éviter les pathologies. Concrètement, pour établir qu'une fonction à plusieurs variables est de classe $\mathcal{C}^1$, on utilisera les théorèmes ≪ cocktails ≫ ; à défaut, on calculera les dérivées partielles **et on vérifiera qu'elles sont continues**.

EXERCICE 4. *Montrer* **sans calculs** *que l'application $A \mapsto A^{-1}$ définie dans $\mathcal{G}\ell_n(\mathbb{R})$ est de classe $\mathcal{C}^1$.*

EXERCICE 5. *(délicat) Montrer que la fonction $f : (x, y) \mapsto \dfrac{e^{-1/(xy)^2}}{e^{-1/x^4} + e^{-1/y^4}}$ est discontinue en l'origine. Pourtant elle admet (en posant $f(0, 0) = 0$) des dérivées partielles à tout ordre (qui sont nulles).*

1.3 Combinaisons de différentielles

Il est clair que sommes, combinaisons linéaires, produits (scalaires, vectoriels, etc...) respectent aussi bien la différentiabilité que la classe $\mathcal{C}^1$. Ce qui permet d'affirmer que des applications compliquées, mais élémentaires, sont $\mathcal{C}^1$ (et même plus, cf. infra).

EXERCICE 6. *Si f, g sont différentiables de $\Omega \to \mathbb{R}$, quelle est au juste la différentielle en a de $f \times g$? Idem si l'espace d'arrivée est $\mathcal{M}_n(\mathbb{R})$.*

Considérons maintenant le problème de la composition :

LEMME 1. *Si f est différentiable en a et g est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et*

$$d(g \circ f)_a = d\,g_{f(a)} \circ d\,f_a$$

C'est assez évident intuitivement : infiniment près de a (resp. $f(a)$), f (resp. g) est affine ; donc $g \circ f$ aussi !

Démonstration. On écrit $f(a+h) = f(a) + \mathrm{d}\, f_a(h) + o(h) = f(a) + k$.
Comme $\mathrm{d}\, f_a$ est linéaire en dimension finie, elle est lipschitzienne et donc $\mathrm{d}\, f_a(h) = O(h)$.
De même, $g(f(a+h)) = g(b+k) = g(b) + \mathrm{d}\, g_b(k) + o(k)$ en posant $b = f(a)$. Par linéarité, il vient

$$\mathrm{d}\, g_b(k) = \mathrm{d}\, g_b\big(\mathrm{d}\, f_a(h) + o(h)\big) = \mathrm{d}\, g_b \circ \mathrm{d}\, f_a(h) + \mathrm{d}\, g_b(o(h)) = \mathrm{d}\, g_b \circ \mathrm{d}\, f_a(h) + o(h)$$

en utilisant cette fois la lipschitzianéité de $\mathrm{d}\, g_b$. On conclut en remarquant que $o(k) = o(O(h)) = o(h)$. ♦

L'expression de $\mathrm{d}(g \circ f)$ permet de calculer les dérivées partielles de $g \circ f$ comme sommes et produits de celles de f et de g. Donc :

COROLLAIRE 1. *Si f et g sont de classe $\mathcal{C}^1$, il en est de même de $g \circ f$.*

Explorons plus en détail l'expression analytique de ces dérivées partielles;

1.4 Matrice jacobienne

Tout d'abord, rappelons que f est une combinaison linéaire de ses **composantes** f_i :

$$f = \sum_{i=1}^{p} f_i e_i$$

En conséquence, la différentiabilité, etc, de f revient à celle des f_i et l'on a

$$\mathrm{d}\, f_a = \sum_{i=1}^{p} \mathrm{d}\, f_{i\,a} e_i \qquad D_j f(a) = \sum_{i=1}^{p} D_j f_i(a) e_i \qquad (\text{ i.e. } \frac{\partial f}{\partial x_j} = \sum_{i=1}^{p} \frac{\partial f_i}{\partial x_j} e_i).$$

Les coordonnées de $D_j f$ sont donc les $D_j f_i$. On peut noter, symboliquement,

$$\mathrm{d}\, f_a = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \mathrm{d}\, x_j = \sum_{j=1}^{n} \sum_{i=1}^{p} D_j f_i(a)\, \mathrm{d}\, x_j e_i$$

qui a l'avantage d'avoir encore un sens quand les x_j sont des fonctions d'autres variables (cf. différentielle de fonctions composées infra). Physiquement $\mathrm{d}\, x_j$ évoque une petite variation de x_j, mathématiquement vous voyez que c'est la forme linéaire $h \mapsto h_j$!

REMARQUE 3. *Cette formule donne sens à l'expression « continûment différentiable » : on voit que $\mathrm{d}\, f$ est continue (comme application de $\Omega \to \mathcal{L}(E, F)$ qui à a associe $\mathrm{d}\, f_a$) ssi les $D_j f$ le sont.*

Prenons un exemple : $(r, \theta) \mapsto (x = r\cos\theta, y = r\sin\theta)$. Les dérivées partielles de $x = r\cos\theta$ et $y = r\sin\theta$ sont données par

$$\begin{aligned} dx &= \cos\theta\, dr - r\sin\theta\, d\theta \\ dy &= \sin\theta\, dr + r\cos\theta\, d\theta \end{aligned} \quad \text{ou plus rigoureusement} \quad \begin{pmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & +r\cos\theta \end{pmatrix}.$$

On voit surgir la matrice de terme général $D_j f_i(a)$:

1.4.1 Matrice jacobienne

On reconnaît la matrice de l' application linéaire $\mathrm{d}\, f(a)$.

DÉFINITION 7. *La* matrice jacobienne *de f au point a est la matrice de terme général $\left[\dfrac{\partial f_i}{\partial x_j}\right]$. C'est la matrice de $\mathrm{d}\, f_a$! On a donc*

$$\mathrm{d}\, f_a(h) = \left[\frac{\partial f_i}{\partial x_j}\right]_{i,j} .h = \begin{pmatrix} \sum_{j=1}^{n} D_j f_1.h_j \\ \vdots \\ \sum_{j=1}^{n} D_j f_i.h_j \\ \vdots \end{pmatrix} = \sum_{i=1}^{p} \big(\sum_{j=1}^{n} D_j f_i.h_j\big) e_i$$

Ce qui permet de caractériser aisément la différentielle d'une composée :

PROPOSITION. *La matrice jacobienne de* $g \circ f$ *est le produit des matrices jacobiennes de* g *et de* f.

EXERCICE 7. *En prenant* $g(u, v) = (x, y)$, *justifier (et apprendre par cœur !) la formule de la chaîne :*

$$(\heartsuit) \qquad \frac{\partial f \circ g(u,v)}{\partial u} = \frac{\partial f(x,y)}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} \qquad \text{où } (x,y) = g(u,v)$$

et plus généralement

$$\frac{\partial f(x_1, \dots, x_n)}{\partial u} = \sum \frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial u}$$

quand les x_j *sont des fonctions d'autres variables* $u, v \dots$.

EXERCICE 8. *Montrer que si* f *et* g *sont réciproques l'une de l'autre et différentiables, la différentielle de* f^{-1} *en* $f(a)$ *est l'application linéaire inverse de la différentielle de* f *en* a *: la matrice jacobienne de* f^{-1} *est l'inverse de celle de* f *[en particulier elle est carrée !]*
NB : dans ce cas on parle de difféomorphisme (HP).

EXERCICE 9. *Matrice jacobienne, et son inverse, des changements de variables en polaires, en sphériques, et en cylindriques.*

1.4.2 Composée avec une application vectorielle

Il s'agit de dériver une application qui court sur une courbe.
Différentions $f \circ \phi$, où $\phi : I \subset \mathbb{R} \to U$. On a donc $\phi(t) = (x_1(t), x_2(t), \dots, x_n(t))$.
Commençons par clarifier la différentielle d'une application bêtement dérivable de $\mathbb{R}$ dans $\mathbb{R}$:

LEMME 2. *La différentielle de* $\phi : I \to \mathbb{R}$ *en* a *est l'application linéaire qui multiplie son argument (réel) par* $\phi'(a)$ *:*

$$d\,\phi_a(h) = h.\phi'(a) \quad \text{ou encore} \quad \phi'(a) = d\,\phi_a(1).$$

Démonstration. Cela se voit en écrivant le DL de deux façons :

$$\Phi(a+h) - \Phi(a) = h\Phi'(a) + o(h) = d\,\Phi_a(h) + o(h). \qquad \blacklozenge$$

L'application linéaire tangente, quand on a une seule variable, consiste donc à multiplier la variable (réelle) par le vecteur dérivé.
Donc l'application linéaire tangente à $f \circ \Phi$, appliquée à un réel h, le multiplie par $\Phi'(a)$ puis applique $d\,f_{\phi(a)}$. Mais par ailleurs, elle multiplie le même h par la dérivée de $f \circ \Phi$. Par identification :

PROPOSITION. *La dérivée de* $f \circ \phi$ *en* a *est l'image de* $\phi'(a)$ *par* $d\,f_{\phi(a)}$ *:*

$$(f \circ \phi)'(a) = d\,f_{\phi(a)}(\phi'(a)) = \nabla f(\phi(a)).\phi'(a)$$

(cette dernière notation fera sens quand on aura parlé de gradient, cf. paragraphe infra)
C'est bien ce que l'on pouvait espérer de plus simple.
Cette proposition est importante du point de vue géométrique : les propriétés de l'arc paramétré par ϕ seront conservées *mutatis mutandis* par f. En particulier, pour un point régulier $\phi(a)$ où la tangente est dirigée par $\phi'(a)$, l'arc image admettra aussi une tangente en $f(\phi(a))$, laquelle sera dirigée par $d\,f_{\Phi(a)}(\phi'(a))$.

$\boxed{\text{En résumé : } f \text{ transporte les points, } d\,f \text{ transporte les tangentes.}}$

EXERCICE 10. *On considère encore une inversion :* $x \mapsto x/\|x\|^2$. *Calculer sa différentielle en* a, *montrer que c'est la composée d'une homothétie et d'une réflexion orthogonale ; en déduire qu'elle conserve les angles. Et l'orientation ? Voyez le dessin, Simba conserve sa belle mâchoire carrée. . . !*

EXERCICE 11. *On considère une application* f *qui vérifie :* $\forall x \in E\ \forall t \in \mathbb{R} \quad f(tx) = tf(x)$. *En dérivant par rapport à* t, *montrer que* $f = d\,f_0$ *et donc que* f *est linéaire.*

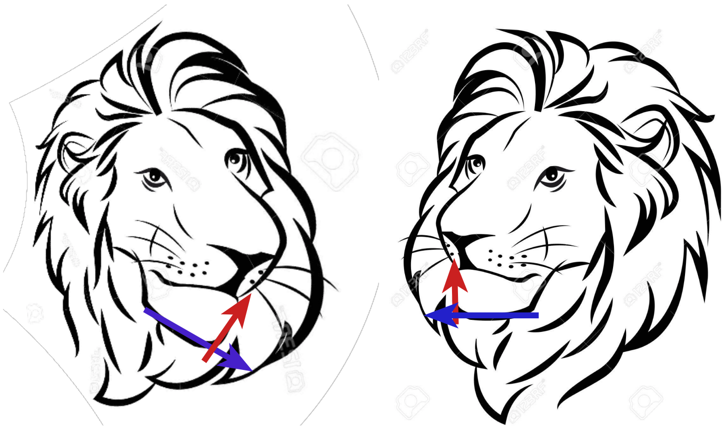

FIGURE 1 – L'inversion conserve les angles, pas les proportions

1.5 Le cas numérique

On prend ici $p = 1$, c'est à dire qu'on regarde une seule composante d'une application de $\mathbb{R}^n \to \mathbb{R}^p$.

1.5.1 Gradient

La matrice jacobienne d'une application de $U \subset \mathbb{R}^n$ dans $\mathbb{R}^1 = \mathbb{R}$ se réduit à un vecteur ligne, la différentielle en un point est une forme linéaire.
Munissons l'espace de départ d'une structure euclidienne (jusqu'à la fin du chapitre), alors toute forme linéaire s'exprime de façon unique par un produit scalaire :

LEMME 3. *L'application qui transforme un vecteur $a \in E$ en la forme linéaire $\phi_a : x \mapsto \langle\, a \,|\, x \,\rangle$ est un isomorphisme de E dans son dual E^*.*

Démonstration. La bilinéarité du produit scalaire montre qu'on a bien affaire à une application linéaire de E dans E^*. Ces deux espaces ont même dimension, or $\phi_a = 0_{E^*} \Rightarrow \phi_a(a) = \|a\|^2 = 0 \Rightarrow a = 0_E$. Le morphisme $a \mapsto \Phi_a$ est injectif, donc bijectif. Ce qui nous sert est que pour toute forme $\phi \in E^*$ il existe un $a \in E$ tel que $\phi = \phi_A$, i.e. la surjectivité. ♦

Ceci permet de définir le vecteur gradient indépendamment d'une base.

DÉFINITION 8. *Le gradient en a de l'application f, de classe $\mathcal{C}^1$ de U dans $\mathbb{R}$, est défini par le vecteur tel que*

$$\forall a \in U \qquad \mathrm{d}f_a = \langle \nabla f(a) \,|\, .\rangle \quad \textit{i.e.} \quad \forall h \in E,\ \mathrm{d}f_a(h) = \langle\, \nabla f(a) \,|\, h \,\rangle.$$

PROPOSITION. *Dans toute BON, on a $\nabla f = (D_1 f, D_2 f, \dots, D_n f)$.*

Cela vient de la formule déjà vue dans le cas général.[1] On peut noter tout cela

$$\mathrm{d}f(a)(h) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} h_j = (D_1 f, D_2 f, \dots, D_n f).(h_1, \dots, h_n).$$

EXERCICE 12. *Calculer les gradients des applications $x \mapsto \|x\|, \|x\|^2, 1/\|x\|$ là où ils existent.*

1. Noter que ∇f a une expression différente dans une base non orthonormale ! On évitera de se poser le problème...

REMARQUE 4. *On a pour tout vecteur* $\mathrm{d}f_a(u) = \langle \nabla f(a) \mid u \rangle$ *et en particulier si* $\|u\| = 1$ *on a par Cauchy-Schwarz* $\mathrm{d}f_a(u) \leqslant \|\nabla f(a)\|$ *et le cas d'égalité est connu. On a donc une interprétation géométrique du gradient :*

Le gradient [si il est non nul] donne la direction de u qui maximise $D_u(f)$.

On peut dire que le gradient va « dans le sens qui augmente le plus f *» : en effet pour* u *positivement colinéaire au gradient on a*

$$f(a + \delta u) - f(a) \sim \delta \|\nabla f(a)\| \quad \text{ce qui est optimal.}$$

C'est l'origine des méthodes dites « de gradient » pour rechercher un maximum (ou un minimum) : à chaque instant on cherche la direction du gradient, et on avance d'un (petit) pas dans cette direction ; si on a diminué f*, on diminue le pas.*

EXERCICE 13. *Implémenter cela à l'aide d'une boucle (tant que* `pas`> 0.0001 *et* $\|\nabla\| > 0.001\ldots$*) et tester sur une fonction de votre choix comme* $x^4 + 3x^2y^2 - 4xy + 9y^2$ *(on cherche le minimum).*

1.5.2 Algèbre des applications numériques de classe $\mathcal{C}^1$

Soit $\mathcal{C}^1(U)$ l'ensemble des applications de classe $\mathcal{C}^1$ de U dans $\mathbb{R}$. Les applications partielles sont $\mathcal{C}^1$ de $\mathbb{R}$ dans $\mathbb{R}$, et donc

PROPOSITION. *Le produit de deux éléments de* $\mathcal{C}^1(U)$ *est encore dans* $\mathcal{C}^1(U)$.

En conséquence, $\mathcal{C}^1(U)$ est une algèbre.
La formule explicite est utile :

$$\nabla(f \times g) = \nabla f \times g + f \times \nabla g.$$

Elle se généraliserait à d'autres sortes de produits.

1.5.3 Caractérisation des fonctions constantes

Nous allons encore exploiter l'idée des dérivées directionnelles (*ie* de restreindre f à un segment bien placé) pour généraliser l'inégalité des AF.

LEMME 4. *Soit* U *un ouvert* **convexe** *et* $f \in \mathcal{C}^1(U)$. *Si* $\|\nabla f\|$ *reste borné sur* U *par la constante* M, *on aura pour tout couple* $(x, y) \in U^2$

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

Pourquoi convexe ? parce que pour démontrer cette inégalité, on va raisonner sur le **segment** $[x, y]$ — qui reste bien inclus dans U du moment que celui-ci est convexe.

Démonstration. On pose pour $t \in [0, 1] \quad g(t) = f(x + t(y - x))$.
Alors g est dérivable, et par la formule démontrée précédemment on a

$$g'(t) =< \nabla f_{x+t(y-x)} \mid y - x >$$

Par CAUCHY-SCHWARZ, on en déduit que $|g'| \leqslant M.\|x - y\|$, d'où l'on tire le résultat voulu en appliquant l'inégalité des AF vectorielle à g entre 0 et 1. ♦

COROLLAIRE FONDAMENTAL. *Sur un ouvert* **convexe***, les applications constantes sont exactement les applications différentiables dont la différentielle est en tout point l'application nulle.*

Ceci reste vrai pour des applications de $\mathbb{R}^n$ dans $\mathbb{R}^p$, en se ramenant aux composantes. Nous admettrons que cette propriété est encore vraie sur tout **ouvertf connexe par arcs** (démontré dans l'exercice 16).

EXERCICE 14. *Etudier l'application* $(x, y) \mapsto \arctan x + \arctan y - \arctan \dfrac{x + y}{1 - xy}$ *sur* $\mathbb{R}^2$ *privé de* $\{xy = 1\}$.

1.5.4 Existence d'un extremum local

Le même principe peut resservir à étudier une condition nécessaire d'extrémalité : si f atteint en a un maximum (par exemple), alors toutes les restrictions de f à des segments passant par a y atteignent aussi un maximum : tous les chemins montent, qui vont au sommet de la montagne (proverbe chinois). Leurs dérivées, qui sont des dérivées directionnelles, ne peuvent donc que s'annuler en a et

PROPOSITION. *Si* f *atteint un extremum en* a*, alors* $\nabla f(a) = 0$.

Un tel point s'appelle un **point critique**, ils sont importants dans l'étude des systèmes différentiels (champs de vecteurs...). Bien entendu, la condition n'est pas suffisante (comme en dimension 1). Ceci dit on a parfois un seul point critique et on sait pour d'autres raisons (compacité...) qu'il y a un maximum (par exemple) : alors on sait que ce maximum est au point critique ! Géométriquement, cela signifie que le plan tangent à la surface $z = f(x, y)$ est horizontal. La question est alors la position de la surface par rapport à ce plan : dessus, dessous, à travers ?...

Exemple : **Problème de Fermat**.
On cherche dans un plan euclidien le minimum de la fonction $M \mapsto f(M) = AM + BM + CM$.

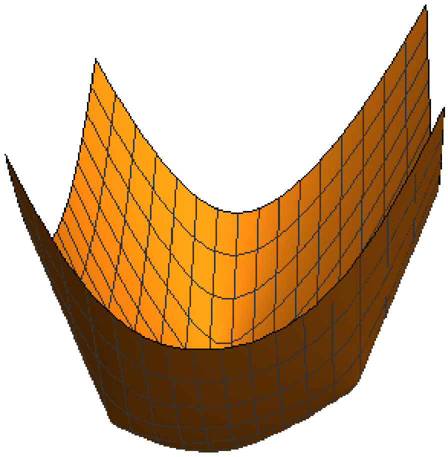

FIGURE 2 – Le graphe de f montre qu'elle est différentiable sauf en A, B, C

• par compacité f admet un minimum m sur le disque fermé $\mathcal{D}$ (et donc compact) de centre A et de rayon $f(A) = BA + CA$. Mais si $M \notin \mathcal{D}$ c'est que $f(M) > AM > f(A) \geqslant m$ et donc $f(M)$ n'est pas minimal. Le minimum sur $\mathcal{D}$ est donc un minimum absolu.

• Le gradient de $M \mapsto AM$ est le vecteur unitaire $\dfrac{\overrightarrow{AM}}{AM} = e_A$. *Sauf si on a un minimum en* A, B *ou* C *(tester)*, le minimum est un point critique où l'on a $e_A + e_B + e_C = 0$.

• la somme de trois vecteurs unitaires est nulle ssi ils forment des angles mutuels de $2\pi/3$.

• On a donc (par exemple l'angle $\widehat{AMB}$ qui vaut $2\pi/3$, ce qui signifie que M appartient à un arc de cercle (circonscrit au triangle équilatéral construit sur AB). Finalement

$$\boxed{\text{le minimum est atteint en l'intersection des trois cercles circonscrits aux triangles équilatéraux } ABC', BCA', CAB'.}$$

1.5.5 Calcul de flux

Considérons une application f de Ω dans $\mathbb{R}$ (un champ scalaire), et un arc dessiné dans l'espace γ. Si on imagine un point matériel qui se déplace le long de l'arc, la partie du champ qui travaille durant le déplacement est celle qui est alignée avec ce déplacement, ce qui suggère de considérer le produit scalaire $\nabla f(\gamma(t)).\gamma'(t)$. Mais ceci est la dérivée de l'application composée $f \circ \gamma$ qui va d'un intervalle de $\mathbb{R}$ dans $\mathbb{R}$. On en déduit

PROPOSITION. *Soit* γ *un arc paramétré* $\gamma : t \mapsto \big(x(t), y(t)\big) \in \Omega, t \in [0, 1]$.
On pose $a = \gamma(0), b = \gamma(1)$. *Alors si* f *est une application de classe* $\mathcal{C}^1$ *de* Ω *dans* $\mathbb{R}$*, on a*

$$f(b) - f(a) = \int_0^1 d\, f_{\gamma(t)}(\gamma'(t))\, dt = \int_0^1 \nabla f(\gamma(t)).\gamma'(t)\, dt.$$

Remarquablement, ce flux ("travail") ne dépend *que des points de départ et d'arrivée* et pas du chemin parcouru !

Considérons maintenant une courbe plane d'équation $F(x,y) = C^{te}$ (le même raisonnement conviendrait avec une surface dans $\mathbb{R}^3$, etc). L'application $F \circ \gamma$ susdite serait donc constante (si on imagine qu'une partie au moins de la courbe peut être paramétrée), et on en déduit un corollaire géométrique simple (généralisé plus bas) :

COROLLAIRE 3. *Le gradient de F dirige la normale à la courbe équipotentielle $F(x,y) = C^{te}$.*

Ceci permet d'écrire très facilement, par exemple, la normale à une conique : au point (x_0, y_0) la normale à la courbe d'équation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ est $\frac{xx_0}{a^2} + \frac{yy_0}{b^2} = 1$ (en effet $\nabla f = \left(\frac{2x_0}{a^2}, \frac{2y_0}{b^2}\right)$). **Rappel** : le vecteur $N = (a,b,c)$ est normal au plan (P) ssi ce dernier a pour équation $ax + by + cz = C^{te}$ (et similairement pour définir un droite du plan par un vecteur normal). On obtient la constante en exprimant que le plan (ou la droite) passe par un point particulier.

EXERCICE 15. *Calculer le gradient de $M \mapsto MF + MF'$; en déduire une propriété de la normale à une ellipse de foyers F, F'.*

EXERCICE 16. *Si Ω est un ouvert connexe par arcs* **de classe** $\mathcal{C}^1$ *[on peut montrer que tout ouvert c.p.a. vérifie cette propriété plus forte], et si $\mathrm{d}\,f$ est nulle sur Ω, alors f est constante.*

1.5.6 Vecteurs tangents à une partie de E

On veut définir la notion de vecteur tangent à une partie de E. Un tel vecteur doit être tangent à une courbe, car on ne dispose pas d'autre notion de tangente ! D'où l'idée simple : **un vecteur tangent à une partie est un vecteur tangent à une courbe tracée dans la partie.**

DÉFINITION 9. *Soit $X \subset E$, $x \in X$ alors $v \in E$ est tangent à X s'il existe un arc γ défini sur un intervalle $]-\varepsilon, \varepsilon[$ tel que*

$$\forall t \in]-\varepsilon, \varepsilon[\quad \gamma(t) \in X \qquad \gamma(0) = x \qquad \gamma'(0) = v.$$

EXERCICE 17. *Montrer que l'ensemble des vecteurs tangents en x à X est stable par homothétie (ce n'est pas toujours un sev ceci dit).*

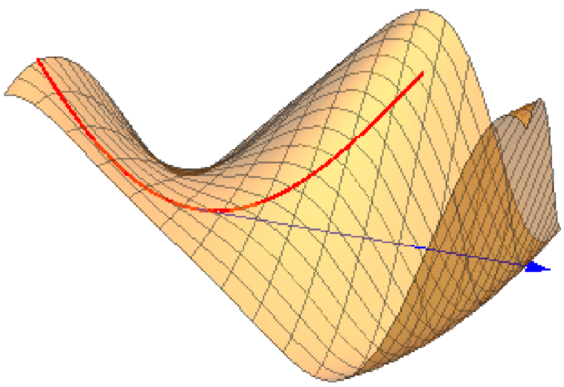

FIGURE 3 – Exemple de vecteur tangent (bleu) à l'arc (rouge) tracé sur X

Exemple : Supposons que X soit un arc paramétré : alors les vecteurs tangents en $x \in X$ forment la tangente !

Exemple : Soit maintenant $X = SO_n(\mathbb{R})$, le groupe des rotations, et $x = I_n$. Si on a un arc paramétré comme ci-dessus, il vérifie

$${}^t\gamma(t)\gamma(t) = I_n$$

et donc par dérivation en $t = 0$ il vient

$${}^t\gamma'(0) + \gamma'(0) = 0$$

c'est à dire que tout vecteur tangent est une matrice antisymétrique. La réciproque est vraie mais plus difficile (Oral Mines-Ponts) ; et l'espace tangent est $\mathcal{A}_n(\mathbb{R})$.
Considérons maintenant une surface donnée comme le graphe d'une application $\mathcal{C}^1$ $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^3$ i.e. $z = f(x, y)$. Toute surface n'est pas de cette forme, mais quitte à se restreindre à une partie d'une surface donnée cela couvre beaucoup de cas.
Pour un arc tracé sur cette surface et passant par le point $(x_0, y_0, z_0) = f(x_0, y_0)$ on aura $\gamma(t) = (x, y, f(x, y))$ et donc

$$\gamma'(0) = \frac{d}{dt}(x, y, f(x, y)) = \left(\dot{x}, \dot{y}, D_1 f\dot{x} + D_2 f\dot{y}\right) = \dot{x}\begin{pmatrix}1\\0\\D_1 f\end{pmatrix} + \dot{y}\begin{pmatrix}0\\1\\D_2 f\end{pmatrix}$$

en prenant les dérivées $\dot{x}, \dot{y}$ en $t = 0$.
Réciproquement il est facile de tracer un arc dont la dérivée a cette forme :
en posant $\gamma(t) = (x_0 + \alpha t, y_0 + \beta t, f(x_0 + \alpha t, y_0 + \beta t))$ qui est bien une courbe tracée sur la surface, on trouve $\gamma'(0) = (\alpha, \beta, \alpha D_1 f + \beta D_2 f)$. En conséquence

PROPOSITION. *L'ensemble des vecteurs tangents à cette surface est le plan Vect* $\left(\begin{pmatrix}1\\0\\D_1 f\end{pmatrix}, \begin{pmatrix}0\\1\\D_2 f\end{pmatrix}\right)$ *(au point* (x_0, y_0, z_0)*).*

NB : ces deux vecteurs ont un sens, ce sont les vecteurs tangents aux deux applications partielles

$$x \mapsto f(x, y_0) \qquad \text{et} \qquad y \mapsto f(x_0, y).$$

Aucun force au monde ne va nous empêcher de définir alors le plan **affine** passant par le point et dirigé par ce plan vectoriel. Il est paramétré, avec les notations précédentes, par

$$(x, y, z) = (x_0, y_0, z_0) + (\alpha, \beta, \alpha D_1 f + \beta D_2 f) = (x_0 + \alpha, y_0 + \beta, z_0 + (\alpha D_1 f + \beta D_2 f)) \quad \alpha, \beta \in \mathbb{R}$$

d'où – éliminant les constantes arbitraires α, β entre les expressions de x, y et z – la formule et définition :

DÉFINITION 10. *Le plan tangent à la surface* $z = f(x, y)$ *au point* (x_0, y_0, z_0) *est le plan d'équation*

$$z = z_0 + (x - x_0)D_1 f(x_0, y_0) + (y - y_0)D_2 f(x_0, y_0) \qquad \textit{i.e.} \qquad z - z_0 = \nabla f.(x - x_0, y - y_0)$$

C'est facile à retenir : on réduit f à son DL à l'ordre 1.
Calcul alternatif : on pose $F(x, y, z) = z - f(x, y)$ et on calcule son gradient, qui donne la normale et donc l'équation du plan tangent.

Exemple : Cherchons le plan tangent en un point non équatorial de la sphère $x^2 + y^2 + z^2 = R^2$: spdg mettons qu'on aie $0 < z = \sqrt{R^2 - x^2 - y^2} = f(x, y)$, alors au point (x_0, y_0, z_0) il vient $\nabla f = \left(\frac{-2x_0}{z_0}, \frac{-2y_0}{z_0}\right)$ d'où une équation du plan tangent :

$$z - z_0 = -(x - x_0)\frac{x_0}{z_0} - (y - y_0)\frac{y_0}{z_0} \qquad \Longleftrightarrow \qquad x_0 x + y_0 y + z_0 z = R^2$$

(le rayon vecteur dirige la normale à la sphère).
Pour généraliser ce résultat, considérons maintenant $X = f^{-1}(a)$ c'est à dire que X est une ligne/surface de niveau (une équipotentielle) du champ scalaire $f : \Omega \to \mathbb{R}$. Que peut-on dire d'un vecteur tangent à X ? Un arc γ tracé sur X vérifie forcément $f \circ \gamma = a$. Si tout le monde est dérivable, on en déduit par dérivation de fonctions composées $\langle \nabla f \,|\, \gamma' \rangle = 0$.
On a démontré

THÉORÈME 3. *Si* f *est une fonction différentiable d'un ouvert* Ω *de* E *(euclidien) à valeurs dans* $\mathbb{R}$*, si* $X = \{f = a\}$ *est une ligne ou surface de niveau de* f*, alors le gradient de* f *en* $x \in X$ *est orthogonal à tout vecteur tangent à* X*.*

En résumé le gradient est normal aux lignes de niveau (vu en électrostatique !). On retrouve immédiatement la normale à une sphère ainsi.

1.6 Dérivées ultérieures : classe $\mathcal{C}^k$

1.6.1 Définition

On observe que $D_j f = \partial_j f = \dfrac{\partial f}{\partial x_j}$ est encore une application de U dans $\mathbb{R}^p$: il se peut qu'elle admette aussi des dérivées partielles. On pose

$$D_{k,j} f = \frac{\partial^2 f}{\partial x_k \partial x_j} = D_k(D_j f) = \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right)$$

et ainsi de suite.

DÉFINITION 11. f *est de classe* $\mathcal{C}^k$ *ssi toutes ses dérivées partielles d'ordre* k *existent et sont continues.*

EXERCICE 18. *Combien y en a-t-il, a priori ?*

1.6.2 Propriétés

De même que pour la classe $\mathcal{C}^1$, on peut combiner des fonctions de classe $\mathcal{C}^k$ et rester dans la même classe : en effet, les dérivées d'ordre k de sommes ou produits de fonctions s'expriment à l'aide des dérivées d'ordre au plus k.

COROLLAIRE 4. *L'ensemble des fonctions de classe* $\mathcal{C}^k$ *de* Ω *dans* $\mathbb{R}$ *est une algèbre.*

Notons la suite d'inclusions : $\mathcal{C}^k(\Omega) \subset \ldots \subset \mathcal{C}^1(\Omega)$. Il est parfois très utile de faire apparaître les opérateurs différentiels comme des applications linéaires :

PROPOSITION. *Pour tout vecteur* u*, la dérivée directionnelle* D_u *définit une application linéaire* $f \mapsto D_u f$ *de* $\mathcal{C}^k(\Omega)$ *dans* $\mathcal{C}^{k-1}(\Omega)$*. On a dans toute base de* E *où* $u = (u_1 \ldots u_n)$

$$D_u = \sum_{j=1}^{n} u_j D_j$$

Résoudre des équations aux dérivées partielles revient donc souvent à chercher des noyaux (ou des espaces propres) de tels opérateurs. Pour pouvoir utiliser des propriétés similaires au lemme des noyaux, on a besoin d'une propriété de commutativité :

THÉORÈME DE SCHWARZ. *Soit* f *de classe* $\mathcal{C}^2$ *sur* U *; alors pour tout couple* (i, j) *on a*

$$D_{i,j} f = D_{j,i} f = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Démonstration. Prenons $(x_0, y_0) \in U \subset \mathbb{R}^2$ spdg et posons $g(y) = f(x_0 + t, y) - f(x_0, y)$ puis

$$F(t) = g(y_0 + t) - g(y_0) = f(x_0 + t, y_0 + t) - f(x_0, y_0 + t) + f(x_0, y_0) - f(x_0 + t, y_0)$$

Il vient en dérivant par rapport à y

$$g'(y) = \partial_2 f(x_0 + t, y) - \partial_2 f(x_0, y)$$

Par le T.A.F. il existe $\theta \in]0, 1[$ tel que

$$F(t) = g(y_0 + t) - g(y_0) = t g'(y_0 + \theta t) = t\big(\partial_2 f(x_0 + t, y_0 + \theta t) - \partial_2 f(x_0, y_0 + \theta t)\big) = t\big(\psi(x_0 + t) - \psi(x_0)\big)$$

en posant $\psi(x) = \partial_2 f(x, y_0 + \theta t)$. Il est irrésistible d'appliquer encore le T.A.F. à ψ :

$$F(t) = t\big(\psi(x_0 + t) - \psi(x_0)\big) = t^2 \psi'(x_0 + \eta t) = t^2 \partial_1 \partial_2 f(x_0 + \eta t, y_0 + \theta t) \quad (\eta \in]0, 1[)$$

Par continuité des dérivées partielles secondes (c'est pourquoi on en a besoin !) on a donc

$$\partial_1 \partial_2 f(x_0, y_0) = \lim_{t \to 0} \frac{F(t)}{t^2}$$

En travaillant non plus sur g mais sur $h : x \mapsto f(x, y_0 + t) - f(x, y_0)$ on trouvera symétriquement

$$\partial_2 \partial_1 f(x_0, y_0) = \lim_{t \to 0} \frac{F(t)}{t^2}$$

Par unicité de la limite, le théorème de Schwarz est démontré. ♦

On en déduit (par une récurrence triviale) que les opérateurs de dérivation **commutent** sur les espaces $\mathcal{C}^k(U)$ — pour $k \geqslant 2$.

On peut alors interpréter la proposition ci-dessous en 1.6.3 – les solutions de $\dfrac{\partial^2 f}{\partial x \partial y} = 0$ sont les $f(x, y) = a(x) + b(y)$ – comme une application du théorème des noyaux :

$$\operatorname{Ker}(D_1 \circ D_2) = \operatorname{Ker} D_1 \oplus \operatorname{Ker} D_2!$$

EXERCICE 19. *Combien une fonction de $\mathcal{C}^k(U)$ a-t-elle, finalement, de dérivées d'ordre k ?*

Exemple : Un contre-exemple au Thm de Schwarz : les dérivées croisées de $f : (x, y) \mapsto \dfrac{xy(x^2 - y^2)}{x^2 + y^2}, f(0, 0) = 0$ existent mais sont distinctes car discontinues.

- f est $\mathcal{C}^0$: en effet $|f(x, y)| \leqslant \dfrac{1}{2}(x^2 + y^2) = \dfrac{1}{2} r^2 \to 0$.
- f est $\mathcal{C}^1$: pas de problème hors (0,0) pour calculer $D_1 f(x, y) = \dfrac{x^4 y + 4x^2 y^3 - y^5}{(x^2 + y^2)^2}$ et comme $|x|, |y| \leqslant r$ on a $|D_1 f(x, y)| \leqslant 6r \to 0$ à l'origine ; or en $(0, 0)$ on a bien une dérivée partielle, en revenant à la définition :

$$D_1 f(0, 0) = \lim_{x \to 0} \frac{f(x, 0) - f(0, 0)}{x - 0} = \lim_{x \to 0} \frac{0}{x} = 0$$

De même pour $D_2 f$ qui existe et est continue sur $\mathbb{R}^2$.
- calcul des dérivées croisées en revenant à la définition :

$$D_{2,1} f(0, 0) = \lim_{y \to 0} \frac{D_1 f(0, y) - D_1 f(0, 0)}{y - 0} = \lim_{y \to 0} \frac{\dfrac{-y^5}{(0^2 + y^2)^2} - 0}{y - 0} = -1$$

et par antisymétrie de f, $D_{1,2} f(0, 0) = +1 \neq D_{2,1} f(0, 0)$.
On pourrait vérifier (très pénible à la main...) que ces dérivées croisées ne sont pas continues quoique bien définies. En effet à côté de l'origine il vient

$$D_{1,2} f(x, y) = \frac{(x^2 - y^2)(x^4 + 10x^2 y^2 + y^4)}{(x^2 + y^2)^3} = \cos(2\theta)(2 - \cos(4\theta))$$

fonction purement orthoradiale qui n'a pas de limite en (0,0).

1.6.3 EDP

On peut être amené à rechercher les fonctions dont les dérivées partielles vérifient une certaine relation.

Exemple : L' équation des ondes (P) $\quad c^2 \dfrac{\partial^2 f}{\partial x^2} = \dfrac{\partial^2 f}{\partial t^2}$ ou l'équation de la chaleur $\dfrac{\partial^2 f}{\partial x^2} = k \dfrac{\partial f}{\partial t}$.

Une telle équation s'appelle évidemment **équation aux dérivées partielles**, ou EDP pour les intimes. Le prix ABEL 2019 a été attribué à Karen Uhlenbeck, qui a contribué des progrès considérables à la théorie de ces équations.
Le principe le plus simple de résolution d'une EDP est le suivant :

PROPOSITION. *Les solutions de* $\dfrac{\partial f(x, y, z \dots)}{\partial x} = 0$ *sont des fonctions* ***arbitraires*** *de* $y, z, \dots$.

Exemple : L'équation $\dfrac{\partial^2 f}{\partial x \partial y} = 0$ se ramène à $\dfrac{\partial f}{\partial y} = C^{te}$ *par rapport à* x, *i.e.* $\dfrac{\partial f}{\partial y} = \beta(y)$; elle a donc pour solution générale

$$f(x,y) = a(x) + b(y) \qquad \text{où } a \text{ et } b \text{ sont dérivables, avec } b' = \beta$$

(voire même $\mathcal{C}^2$, si l'on recherche des solutions de cette classe).
On en déduit les solutions de (P) par le changement de variables **et donc aussi de fonction inconnue**[2]

$$u = x + ct, v = x - ct \qquad g(u,v) = f(x,y).$$

EXERCICE 20. *Le faire (on calculera* $\dfrac{\partial^2 g}{\partial u \partial v}$*).*

Exemple : $x\partial_1 f + y\partial_2 f = x^2 + y^2$ incite (surtout par son second membre) à passer en coordonnées polaires, i.e. à poser $x = r\cos\theta, y = r\sin\theta$ et forcément $f(x,y) = f(r\cos\theta, r\sin\theta) = g(r,\theta)$. On constate avec une joie sans mélange que

$$D_1 g = \frac{d\,g}{d\,r} = \frac{d\,f(r\cos\theta, r\sin\theta)}{d\,r} = \cos\theta D_1 f + \sin\theta D_2 f$$

c'est à dire que l'équation initiale se réécrit

$$D_1 g = r^2 \quad \text{d'où} \quad g(r,\theta) = \frac{r^2}{2} + h(\theta)$$

où h est une fonction quelconque ($\mathcal{C}^1$ quand même, si l'on veut que g le soit).
De manière similaire, l'équation

$$y\,D_1 f - x\,D_2 f = y/x \qquad (x > 0)$$

se résout bien par le même changement de variable. Cette fois c'est

$$D_2 g = \frac{d\,f(r\cos\theta, r\sin\theta)}{d\,\theta} = -r\sin\theta D_1 f + r\cos\theta D_2 f = -yD_1 f + xD_2 f$$

et l'équation devient $D_2 g = -\tan\theta$ d'où $g = \ln(\cos\theta) + h(r)$ i.e. $f(x,y) = \ln\left(\dfrac{x}{\sqrt{x^2+y^2}}\right) + h(\sqrt{x^2+y^2})$.

EXERCICE 21. *Autre exemple : on cherche les applications* f *de classe* $\mathcal{C}^1$ *qui vérifient*

$$\frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} + 3(x-y)f = 0$$

Montrer d'abord que $\phi : (x,y) \mapsto (u,v) = (x+y, xy)$ *est une bijection de* $\Omega = \{x > y\}$ *sur un ouvert que l'on précisera. Ensuite en posant* $f(x,y) = g(u,v)$ *(càd* $g = f \circ \phi^{-1}$*), montrer que notre équation se traduit par* $D_1 g - 3g = 0$*. Résoudre cette dernière équation, en déduire une expression de* f*.*

2. Attention ! En physique f est une grandeur physique (enthalpie, etc) et pas une fonction au sens mathématique, on a donc $H(p,v) = H(v,T)$, etc. Mathématiquement c'est un non-sens : il faut changer de fonction car techniquement on effectue un changement de variable i.e. une composition de fonctions.