Chapitre 15

Calcul différentiel

Hugo Salou MPI⋆
Dernière mise à jour le 17 mars 2023

Exemple 1:
Dans le plan $\mathbb{R}^2$, soit $\mathscr{C}$ le cercle de centre $(0,0)$ et de rayon 1 :

$$(x,y) \in \mathscr{C} \iff \underbrace{x^2+y^2=1}_{\text{équation implicite}} \iff \overbrace{\exists t \in \mathbb{R}, \begin{cases} x = \cos t \\ y = \sin t. \end{cases}}^{\text{équations paramétrées}}$$

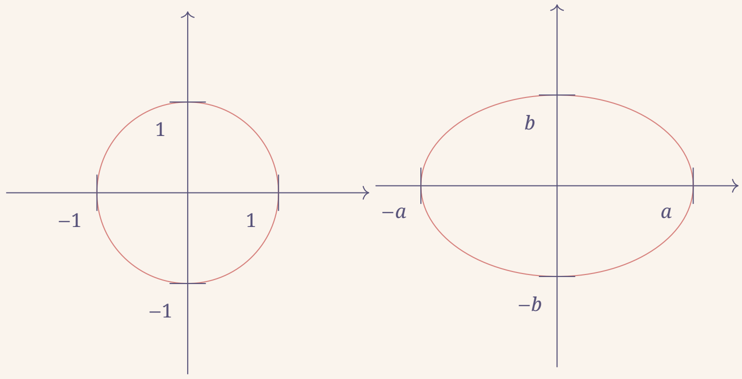
Plus généralement, soient deux réels $a>0$ et $b>0$. La courbe $\mathscr{E}$ d'équation implicite

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

est appelée *ellipse*. Elle est représentée sur la figure ci-dessous. C'est aussi une *courbe paramétrée* :

$$(x,y) \in \mathscr{E} \iff \exists t \in \mathbb{R}, \begin{cases} x = a\cos t \\ y = b\sin t. \end{cases}$$

Les deux courbes sont *bornées*. On peut passer de $\mathscr{C}$ à $\mathscr{E}$ au moyen des changements de variables $X = x/a$ et $Y = y/b$. On nomme a le « demi grand axe » et b le « demi petit axe » de l'ellipse $\mathscr{E}$.



$x^2+y^2=1$ $\qquad$ $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$

Figure 1 – Cercle $\mathscr{C}$ et ellipse $\mathscr{E}$

Exemple 2:
Soient deux réels $a>0$ et $b>0$. La courbe $\mathscr{H}$ d'équation implicite

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

est appelée *hyperbole*. Elle est représentée sur la figure ci-dessous. Cette hyperbole est la réunion de deux branches $\mathscr{H}_+ = \{(x,y) \in \mathscr{H} \mid x \geqslant 0\}$ et $\mathscr{H}_- = \{(x,y) \in \mathscr{H} \mid x \leqslant 0\}$. La courbe $\mathscr{H}_+$ est une courbe paramétrée

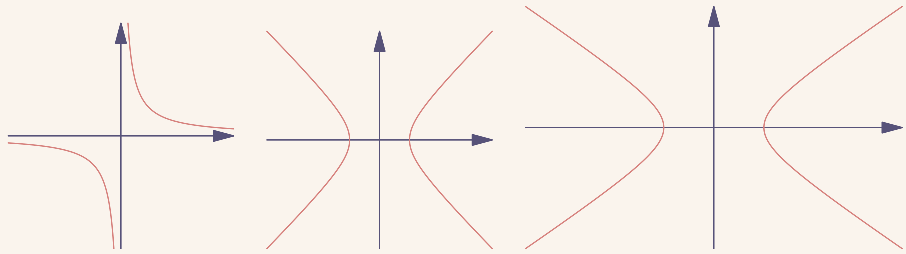
$$(x,y) \in \mathscr{H}_+ \iff \exists t \in \mathbb{R}, \begin{cases} x = a\,\mathrm{ch}\,t \\ y = b\,\mathrm{sh}\,t. \end{cases}$$

L'hyperbole $\mathscr{H}$ possède deux asymptotes $\frac{y}{b} = \frac{x}{a}$ et $\frac{y}{b} = -\frac{x}{a}$.

$$\begin{aligned} \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 &\iff X^2 - Y^2 = 1 && \text{où } \begin{cases} X = x/a \\ Y = y/b \end{cases} \\ &\iff U \cdot V = 1 && \text{où } \begin{cases} U = X+Y \\ V = X-Y \end{cases} \\ &\iff V = \frac{1}{U} \end{aligned}$$

L'hyperbole $\mathscr{H}$ a une *zone interdite* entre -1 et 1. En effet

$$X^2 - Y^2 = 1 \iff X^2 = 1 + Y^2 \implies X^2 \geqslant 1.$$

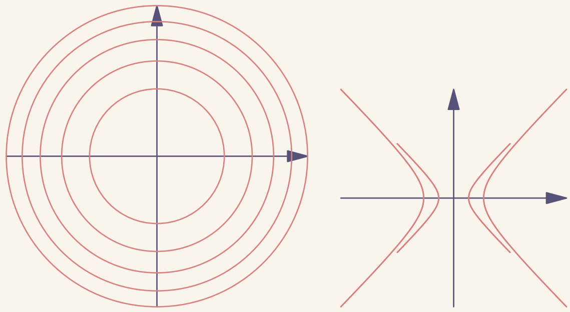


$V = \dfrac{1}{U}$ $\qquad$ $X^2 - Y^2 = 1$ $\qquad$ $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$

FIGURE 2 – Hyperboles

DÉFINITION 3:
Soit $D \subset \mathbb{R}^2$ une partie du plan et soit $f : D \longrightarrow \mathbb{R}$ une fonction définie sur D. Pour chaque réel K, la *courbe de niveau* K de la fonction f est l'ensemble C_K des points $(x, y) \in D$ tels que

$$f(x, y) = K.$$



$f(x, y) = x^2 + y^2$ $\qquad$ $f(x, y) = x^2 - y^2$

FIGURE 3 – Courbes de niveaux

Faire varier le niveau K sur la figure ci-dessus revient à faire varier la hauteur K du plan de la figure ci-dessous.

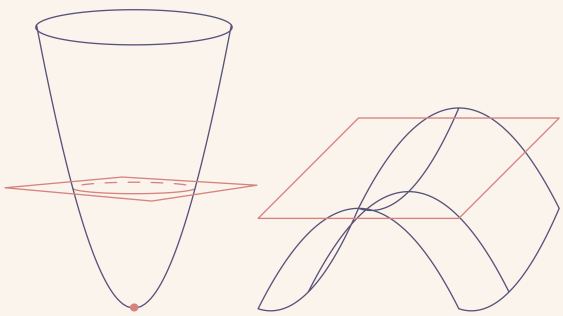


FIGURE 4 – L'intersection d'un paraboloïde et d'un plan (à gauche), d'une selle de cheval et d'un plan (à droite)

1 Dérivées partielles

2 Dérivées partielles

DÉFINITION 4:
Soit $D \subset \mathbb{R}^2$ un ouvert, et soit $f : D \longrightarrow \mathbb{R}$, $(x, y) \mapsto f(x, y)$ une fonction. Soit un point $(a, b) \in D$.

1. Si
$$\lim_{h \to 0} \frac{f(a+h, b) - f(a, b)}{h}$$
existe et est finie, alors ce nombre réel est noté $\partial_1 f(a, b)$ ou $\frac{\partial f}{\partial x}$ et est appelé la première *dérivée partielle* de f en (a, b).
2. Si
$$\lim_{h \to 0} \frac{f(a, b+h) - f(a, b)}{h}$$
existe et est finie, alors ce nombre réel est noté $\partial_2 f(a, b)$ ou $\frac{\partial f}{\partial y}$ et est appelé la seconde *dérivée partielle* de f en (a, b).
3. Pour $i \in [\![1, 2]\!]$, si $\partial_i f(a, b)$ existe pour tout $(a, b) \in D$, alors al fonction $\partial_i f : D \longrightarrow \mathbb{R}$ est appelée la i-ème *dérivée partielle* de f.
4. On dit que la fonction f est de classe $\mathscr{C}^1$ sur D si les deux dérivées partielles $\partial_1 f$ et $\partial_2 f$ existent et sont continues sur D.

EXERCICE 5 (Une fonction qui possède des dérivées partielles non continues):
Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ la fonction définie par
$$f(x, y) = x^2 \cdot \sin\left(\frac{y}{x}\right) \quad \text{si } x \neq 0 \quad \text{et} \quad f(0, y) = 0 \text{ sinon..}$$

1. *Montrer que f est continue sur $\mathbb{R}^2$.*
2. *Montrer que les dérivées partielles $\partial_1 f(x, y)$ et $\partial_2 f(x, y)$ existent si $x \neq 0$, et les calculer.*
3. *Montrer que les dérivées partielles $\partial_1 f(0, y)$ et $\partial_2 f(0, y)$ existent et les calculer.*
4. *Montrer que la fonction f n'est pas de classe $\mathscr{C}^1$.*

1. La fonction f est continue sur $\mathbb{R}^\star \times \mathbb{R}$ d'après les théorèmes généraux (par produit et par composition). Montrons que $f(0+h, y+k) \xrightarrow[(h,k)\to(0,0)]{} f(0, y)$, *i.e.* $|f(0+h, y+k) - f(0, y)| \longrightarrow 0$. On a
$$\begin{aligned} 0 \leqslant |f(0+h, y+k) - f(0, y)| &= \left|h^2 \sin\left(\frac{y+k}{h}\right) - 0\right| \\ &= h^2 \left|\sin\left(\frac{y+k}{h}\right)\right| \\ &\leqslant h^2 \leqslant \sqrt{h^2+k^2}^2 \leqslant \|(h,k)\|_2^2 \end{aligned}$$
D'après le théorème des gendarmes, f est continue en $(0, y)$, f est donc continue sur $\mathbb{R}^2$.
2. En tout point de $\mathbb{R}^2 \times \mathbb{R}$, $\partial_1 f$ et $\partial_2 f$ existent par composition et par produit. Et,
$$\begin{aligned} \partial_1 f(x, y) &= \frac{\partial f}{\partial x}(x, y) \\ &= 2x \sin\left(\frac{y}{x}\right) + x^2 \times \frac{-y}{x^2} \cos\left(\frac{y}{x}\right) \\ &= 2x \sin\left(\frac{y}{x}\right) - y \cos\left(\frac{y}{x}\right) \end{aligned}$$
$$\begin{aligned} \partial_2 f(x, y) &= \frac{\partial f}{\partial y}(x, y) \\ &= x^2 \cdot \frac{1}{x} \cdot \cos\left(\frac{y}{x}\right) \\ &= x \cos\left(\frac{y}{x}\right) \end{aligned}$$
3. Calculons
$$\frac{f(0+h, y) - f(0, y)}{h} = \frac{h^2 \sin\left(\frac{y}{h}\right) - 0}{h} = h \sin\left(\frac{y}{h}\right) \xrightarrow[h \to 0]{} 0$$

car $h \to 0$ et sin est borné. D'où, $\partial_1 f(0,y)$ existe et $\partial_1 f(0,y) = 0$. De plus, $\big[f(0,y+h) - f(0,y)\big]/k = (0-0)/k = 0 \xrightarrow[k\to 0]{} 0$. D'où $\partial_2 f(0,y)$ existe et $\partial_2 f(0,y) = 0$.

4. On a $\partial_1 f(x,1) \underset{x\to 0}{\not\longrightarrow} \partial_1 f(0,1)$ car $1 \times \cos\left(\frac{1}{x}\right)$ n'a pas de limite.

Théorème 6 (Formule de Taylor & Young à l'ordre 1):
Soit $D \subset \mathbb{R}^2$ un ouvert, et soit $f : D \to \mathbb{R}$ une fonction. Si f est de classe $\mathscr{C}^1$, alors, pour tout $(a,b) \in D$,

$$f(a+h, b+k) = f(a,b) + h\,\partial_1 f(a,b) + k\,\partial_2 f(a,b) + \underbrace{\|(h,k)\| \cdot \varepsilon(h,k)}_{o(h,k)}$$

où $\varepsilon(h,k) \xrightarrow[(h,k)\to(0,0)]{} 0$.

Rappel (Accroissements finis):
Si φ est une fonction continue sur un segment $[\alpha,\beta]$, et qu'elle est dérivable sur $]\alpha,\beta[$, alors il existe $c \in\ ]\alpha,\beta[$ tel que

$$\varphi(\beta) - \varphi(\alpha) = \varphi'(c) \cdot (\beta - \alpha).$$

Démonstration:
On a $f(a+h,b+k) - f(a,b) = \big[f(a+h,b+k) - f(a,b+k)\big] + \big[f(a,b+k) - f(a+k,b+k)\big]$. Soit $\varphi(\vec{x}) = f(x,b+k)$. Alors, d'après le théorème des accroissements finis, il existe $c \in\ ]a, a+h[$ tel que $\varphi(a+h) - \varphi(a) = \varphi'(c) \cdot \big((a+h) - a\big)$. D'où, $f(a+h,b+k) - f(a,b+k) = \partial_1 f(c,b+k) \times h$. De même, soit $\psi(y) = f(a,y)$; alors, il existe $d \in\ ]b, b+k[$ tel que $f(a,b+k) - f(a,b) = \partial_2 f(a,d) \times k$. Or, f est de classe $\mathscr{C}^1$, d'où $\partial_1 f(c,b+k) \to \partial_1 f(a,b)$ quand $(h,k) \to (0,0)$, car $(c,b+k) \to (a,b)$ et $\partial_1 f$ est continue. De même, $\partial_2 f(a,d) \to \partial_2 f(a,b)$ quand $(h,k) \to (0,0)$. Ainsi, $\partial_1 f(c,b+k) = \partial_1 f(a,b) + \varepsilon_1(h,k)$ où $\varepsilon_1(h,k) \to 0$ quand $(h,k) \to (0,0)$. De même, $\partial_2 f(a,d) = \partial_2 f(a,b) + \varepsilon_2(h,k)$ avec $\varepsilon(h,k) \to 0$ quand $(h,k) \to (0,0)$. On a donc $f(a+h,b+k) - f(a,b) = h\,\partial_1 f(a,b) + h\,\varepsilon_1(h,k) + k\,\partial_2 f(a,b) + k\,\varepsilon_2(h,k)$. On pose $R(h,k)$ le reste.

$$\begin{aligned} 0 \leqslant |R(h,k)| &\leqslant |h| \cdot |\varepsilon_1(h,k)| + |k| \cdot |\varepsilon_2(h,k)| \\ &\leqslant \sqrt{h^2+k^2} \times \underbrace{\big[|\varepsilon_1(h,k)| + |\varepsilon_2(h,k)|\big]}_{\varepsilon(h,k)} \\ &= \|(h,k)\| \cdot \varepsilon(h,k) = o(h,k) \end{aligned}$$

Remarque 7:

1. (fonctions de $\mathbb{R}^p$ vers $\mathbb{R}$) Soit $p \in \mathbb{N}^\star$. De même que les fonctions de $\mathbb{R}^2$ vers $\mathbb{R}$, on peut étudier une fonction de $\mathbb{R}^p$ vers $\mathbb{R}$, définir ses p dérivées partielles $\partial_1 f, \ldots, \partial_p f$ si elles existent, et dire que la fonction f est de classe $\mathscr{C}^1$ sur un ouvert $U \subset \mathbb{R}^p$ si ses p dérivées partielles sont continues.
   D'après la formule de Taylor & Young, si la fonction f est de classe $\mathscr{C}^1$ sur U, alors : en tout point $\vec{a} = (a_1, \ldots, a_p) \in U$,
   $$f(\vec{a} + \vec{h}) = f(\vec{a}) + h_1\,\partial_1 f(\vec{a}) + \cdots + h_p\,\partial_p f(\vec{a}) + o(\vec{h})$$
   où $\vec{h} = (h_1, \ldots, h_p)$.
2. (fonctions de $\mathbb{R}^p$ vers $\mathbb{R}^n$) Soient $p \in \mathbb{N}^\star$ et $n \in \mathbb{N}^\star$, un ouvert $U \subset \mathbb{R}^p$ et une fonction
   $$\begin{aligned} f : U &\longrightarrow \mathbb{R}^n \\ \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} &\longmapsto \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_n(\vec{x}) \end{pmatrix} \end{aligned}$$
   Soit $\vec{a} \in U$. Si chacune des n fonctions $f_i : U \to \mathbb{R}$ est de classe $\mathscr{C}^1$ sur U, alors
   $$\forall i \in [\![1,n]\!], \qquad f_i(\vec{a}+\vec{h}) = f_i(\vec{a}) + h_1\,\partial_1 f_i(\vec{a}) + \cdots + h_p\,\partial_p f_i(\vec{a}) + o(\vec{h}).$$
   Matriciellement, cette égalité s'écrit
   $$\underbrace{\begin{pmatrix} f_1(\vec{a}+\vec{h}) \\ \vdots \\ f_n(\vec{a}+\vec{h}) \end{pmatrix}}_{f(\vec{a}+\vec{h})} = \underbrace{\begin{pmatrix} f_1(\vec{a}) \\ \vdots \\ f_n(\vec{a}) \end{pmatrix}}_{f(\vec{a})} + \underbrace{\begin{array}{c} \begin{array}{ccc} \partial_1 f(\vec{a}) & \cdots & \partial_p f(\vec{a}) \\ \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} \partial_1 f_1(\vec{a}) & \cdots & \partial_p f_1(\vec{a}) \\ \vdots & \ddots & \vdots \\ \partial_1 f_n(\vec{a}) & \cdots & \partial_p f_n(\vec{a}) \end{pmatrix} \end{array}}_{\mathrm{J}_f(\vec{a})} \cdot \underbrace{\begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix}}_{\vec{h}} + \|\vec{h}\| \cdot \underbrace{\begin{pmatrix} \varepsilon_1(\vec{h}) \\ \vdots \\ \varepsilon_n(\vec{h}) \end{pmatrix}}_{\varepsilon(\vec{h})}.$$

Définition 8:
Soit f une fonction définie sur un ouvert $U \subset \mathbb{R}^p$ comme

$$\begin{aligned} f : U &\longrightarrow \mathbb{R}^n \\ \vec{x} = (x_1, \ldots, x_p) &\longmapsto \big(f_1(\vec{x}), \ldots, f_n(\vec{x})\big). \end{aligned}$$

Si chaque fonction f_i admet p dérivées partielles en $\vec{a}$, alors la matrice $\mathrm{J}_f(\vec{a}) \in \mathcal{M}_{n,p}(\mathbb{R})$ définie par $\big(\mathrm{J}_f(\vec{a})\big)_{i,j} = \frac{\partial f_i}{\partial x_j}(\vec{a})$, pour tout $i \in [\![1, n]\!]$ et $j \in [\![1, p]\!]$, est appelée *jacobienne* de f en $\vec{a}$.

$$\begin{array}{c} \\ \\ \nabla f_1 \longrightarrow \\ \vdots \\ \nabla f_n \longrightarrow \end{array} \begin{array}{c} \begin{array}{ccc} \partial_1 f(\vec{a}) & \cdots & \partial_p f(\vec{a}) \\ \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} \partial_1 f_1(\vec{a}) & \cdots & \partial_p f_1(\vec{a}) \\ \vdots & \ddots & \vdots \\ \partial_1 f_n(\vec{a}) & \cdots & \partial_p f_n(\vec{a}) \end{pmatrix} \end{array} = \mathrm{J}_f(\vec{a}).$$

Exemple 9:
La fonction

$$\begin{aligned} M : \mathbb{R}_\star^+ \times \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ (r, \varphi) &\longmapsto (x, y) = (r\cos\varphi, r\sin\varphi) \end{aligned}$$

change les coordonnées polaires (r, φ) en coordonnées cartésiennes (x, y). La fonction M possède des dérivées partielles et sa matrice jacobienne est

$$\mathrm{J}_M(r, \varphi) = \begin{pmatrix} \frac{\partial x}{\partial r}(r, \varphi) & \frac{\partial x}{\partial \varphi}(r, \varphi) \\ \frac{\partial y}{\partial r}(r, \varphi) & \frac{\partial y}{\partial \varphi}(r, \varphi) \end{pmatrix} = \begin{array}{c} \begin{array}{cc} \frac{\partial M}{\partial r} & \frac{\partial M}{\partial \varphi} \\ \downarrow & \downarrow \end{array} \\ \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} \end{array}.$$

Ses vecteurs colonnes sont représentés sur la figure ci-dessous. Le vecteur $\frac{\partial M}{\partial r}$ est tangent à la droite paramétrée par $r \mapsto M(r, \varphi)$; le vecteur $\frac{\partial M}{\partial \varphi}$ est tangent au cercle paramétré par $\varphi \mapsto M(r, \varphi)$.

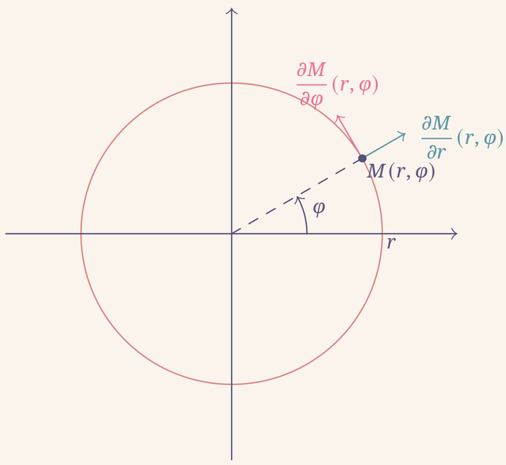


Figure 5 – Coordonnées polaires

3 La différentielle d'une fonction

Dans toute la suite, E et F sont des espaces vectoriels normés de dimensions finies $p = \dim E$ et $n = \dim F$. Si l'on choisit une base de E et une base de F, alors on pourra confondre $\vec{x} \in E$ et $(x_1, \ldots, x_p) \in \mathbb{R}^p$ d'une part, et $f(\vec{x}) \in F$ et $\big((f_1(\vec{x}), \ldots, f_n(\vec{x})\big) \in \mathbb{R}^n$ d'autre part.

Proposition – Définition 10:
Soient U un ouvert de E, un point $\vec{a} \in U$, et une fonction $f : U \to F$. On dit que f est *différentiable* en $\vec{a}$ s'il existe une application linéaire $\ell_{\vec{a}} : E \to F$ telle que

$$f(\vec{a}+\vec{h}) = f(\vec{a}) + \ell_{\vec{a}}(\vec{h}) + o(\vec{h}).$$

Si f est différentiable en $\vec{a}$, alors

1. l'application $\ell_{\vec{a}}$ est unique, on l'appelle *la différentielle* de f en $\vec{a}$, et on note

$$\begin{aligned} \ell_{\vec{a}} = \mathrm{d}f(\vec{a}) : E &\longrightarrow F \\ \vec{h} &\longmapsto \ell_{\vec{a}}(\vec{h}) = \mathrm{d}f(\vec{a}) \cdot \vec{h}\ ; \end{aligned}$$

2. les dérivées partielles de f en $\vec{a}$ existent et

$$\forall \vec{h} \in E, \qquad \mathrm{d}f(\vec{a}) \cdot \vec{h} = h_1\, \partial_1 f(\vec{a}) + \cdot + h_p\, \partial_p f(\vec{a}) = \sum_{i=1}^{p} h_i\, \partial_i f(\vec{a}).$$

Démonstration:

1ère preuve de l'unicité. Supposons qu'il existe L_1 et L_2 deux applications linéaires telles que, pour tout vecteur $\vec{h}$, $f(\vec{a}+\vec{h}) = f(\vec{a})+L_1(\vec{h})+o(\vec{h})$ et $f(\vec{a}+\vec{h}) = f(\vec{a})+L_2(\vec{h})+o(\vec{h})$. Montrons que $L_1 = L_2$. On soustrait les deux expressions de $f(\vec{a}+\vec{h})$ et on trouve $\vec{0} = \vec{0} + L_1(\vec{h}) - L_2(\vec{h}) + o(\vec{h})$. On fixe $\vec{h} \in E$, et soit $t \in \mathbb{R}^\star : L_1(t\vec{h}) - L_2(t\vec{h}) = o(t\vec{h}) = \|t\vec{h}\|\varepsilon(t\vec{h})$. Ainsi, par homogénéité de la norme et par linéarité de L_1 et L_2, on a $t\,L_1(\vec{h}) - t\,L_2(\vec{h}) = |t| \cdot \|\vec{h}\| \cdot \varepsilon(t\vec{h})$. Ainsi, $L_1(\vec{h}) - L_2(\vec{h}) = \|\vec{h}\| \cdot \hat{\varepsilon}(t\vec{h})$ où $\hat{\varepsilon}(t\vec{h}) = \pm\varepsilon(t\vec{h})$. On fait tendre t vers 0, et donc $L_1(\vec{h}) - L_2(\vec{h}) \to 0$. D'où, $L_1(\vec{h}) - L_2(\vec{h}) = \vec{0}$ pour tout vecteur $\vec{h} \in E$. On en déduit $L_1 = L_2$.

2nde preuve de l'unicité & formule. Soit $\ell_{\vec{a}}$ une application linéaire telle que $f(\vec{a}+\vec{h}) = f(\vec{a}) + \ell_{\vec{a}}(\vec{h}) + \|\vec{h}\|\,\varepsilon(\vec{h})$, pour tout vecteur $\vec{h} \in E$. En particulier, on pose $\vec{h} = t\,\vec{e}_i$, où $\vec{e}_i$ est l'un des vecteurs d'une base de E, et $t \in \mathbb{R}^\star$. Ainsi, $f(\vec{a}+t\vec{e}_i) = f(\vec{a}) + \ell_{\vec{a}}(t\vec{e}_i) + \|t\vec{e}_i\|\,\varepsilon(t\vec{e}_i)$. En développant, on trouve $f(a_1,\dots,a_{i-1},a_i+t,a_{i+1},\dots,a_p) - f(a_1,a_2,\dots,a_{i-1},a_i,a_{i+1},\dots,a_p) = t\ell_{\vec{a}}(\vec{e}_i) + |t|\,\|\vec{e}_i\|\,\varepsilon(t\vec{e}_i)$, par linéarité de $\ell_{\vec{a}}$ et par homogénéité de la norme. On divise par t et on trouve

$$\frac{f(a_1,\dots,a_{i-1},a_i+t,a_{i+1},\dots,a_p) - f(a_1,a_2,\dots,a_{i-1},a_i,a_{i+1},\dots,a_p)}{t} = \ell_{\vec{a}}(\vec{e}_i) + \|\vec{e}_i\|\,\hat{\varepsilon}(t\vec{e}_i) \xrightarrow[t\to 0]{} \ell_{\vec{a}}(\vec{e}_i).$$

D'où, $\partial_i f(\vec{a})$ existe et vaut $\ell_{\vec{a}}(\vec{e}_i)$.
Si $\vec{h}$ est quelconque, alors on pose $\vec{h} = h_1\vec{e}_1 + \cdots + h_p\vec{e}_p$. Par linéarité de $\ell_{\vec{a}}$, on a

$$\mathrm{d}f(\vec{a}) \cdot \vec{h} = \ell_{\vec{a}}(\vec{h}) = h_1\,\ell_{\vec{a}}(\vec{e}_1) + \cdots + h_p\,\ell_{\vec{a}}(\vec{e}_p) = h_1\,\partial_1 f(\vec{a}) + \cdots + h_p\,\partial_p f(\vec{a}).$$

Exemple 11:

1. La fonction

$$\begin{aligned} f : \mathcal{M}_n(\mathbb{R}) &\longrightarrow \mathcal{M}_n(\mathbb{R}) \\ A &\longmapsto A^2 \end{aligned}$$

est différentiable en chaque point $A \in \mathcal{M}_n(\mathbb{R})$ et $\mathrm{d}f(A) \cdot H = AH + HA$. En effet, $f(A+H) = (A+H)^2 = A^2 + AH + HA + H^2 = f(A) + [AH + HA] + H^2$. La formule entre crochets est linéaire en H, et dépend de A, on la note $\ell_A(H)$. Et, montrons que $H^2 = o(H)$ *i.e.* $H^2 = \|H\|\varepsilon(H)$, avec $\varepsilon(H) \to 0$ quand $H \to 0$. Montrons alors que $H^2/\|H\| \to 0$ quand $H \to 0$, *i.e.* montrons que $\|H^2\|/\|H\| \to 0$ quand $\|H\| \to 0$.[1] On choisit une norme sous-multiplicative. Alors, $\|H^2\| \leqslant \|H\| \times \|H\|$. D'où, $0 \leqslant \|H^2\|/\|H\| \leqslant \|H\|^2/\|H\| \leqslant \|H\|$. Par le théorème des gendarmes, on a bien $\|H^2\|/\|H\| \to 0$ quand $H \to 0$.

Rappel (comatrice):

1. L'enjeu est de choisir une norme adaptée à la question : toutes les normes sont équivalentes en dimension finie.

$$
\begin{aligned}
\det A &= \begin{array}{c|cccccc|}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{j} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} \\
\multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{\downarrow} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} \\
 & a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\
 & a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\
 & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
i \to & a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\
 & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
 & a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn}
\end{array} \\
&= (-1)^{1+j} a_{1j}\Delta_{1j} + (-1)^{2+j} a_{2j}\Delta_{2j} + \cdots + (-1)^{n+j} a_{n,j}\Delta_{n,j} \\
&= \sum_{i=1}^{n} (-1)^{i+j} a_{i,j}\Delta_{i,j},
\end{aligned}
$$

où la matrice $\Delta_{i,j}$ est le déterminant de la matrice carrée obtenue en supprimant la colonne j et la ligne i. La matrice $\Delta_{i,j}$ s'appelle le *mineur*; le *cofacteur* est le terme $(-1)^{i+j}\Delta_{i,j}$. Si $j \neq k$, alors $\sum_{i=1}^{n}(-1)^{i+j}a_{i,k}\Delta_{i,j} = 0$, car il s'agit du déterminant où deux colonnes sont égales. Ainsi, $\sum_{i=1}^{n}(-1)^{i+j}a_{i,k}\Delta_{i,j} = \det(A)\cdot\delta_{j,k}$, où $\delta_{j,k}$ est le symbole de Kronecker. On pose $b_{j,i} = (-1)^{i+j}\Delta_{i,j}$, et on a donc $\sum_{i=1}^{n}(-1)^{i+j}a_{i,k}\Delta_{i,j} = \sum_{i=1}^{n} b_{i,j}a_{i,k}$. En nommant $(b_{j,i}) = B$, on trouve donc $B \cdot A = \det(A) \cdot I_n$. Si $\det A \neq 0$, alors

$$\left(\frac{1}{\det A}B\right) \cdot A = I_n \qquad \text{et donc} \qquad A^{-1} = \frac{1}{\det A}B.$$

La *comatrice* est la matrice B transposée : $[\operatorname{com} A]_{i,j} = (-1)^{i+j} \cdot \Delta_{i,j}$ Avec cette définition, on a donc

$$A^{-1} = \frac{1}{\det A} \cdot (\operatorname{com} A)^{\top}.$$

Rappel (déterminant):

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}.$$

Exercice 12: 1. *Soient $H \in \mathcal{M}_n(\mathbb{R})$ et I_n la matrice identité. Montrer que*

$$\det(I_n + H) = 1 + \operatorname{tr} H + o(H).$$

2. *En déduire que la fonction* $\det : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ *est différentiable en I_n. Quelle est sa différentielle en I_n ?*
3. *Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice inversible. Montrer que* det *est différentiable en A. Quelle est sa différentielle en A ?*

1. On calcule, en utilisant la formule du déterminant rappelée précédemment :

$$
\begin{aligned}
\det(I_n + H) &= \begin{vmatrix}
1 + h_{11} & h_{12} & \dots & h_{1n} \\
h_{21} & 1 + h_{22} & \dots & h_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
h_{n1} & h_{n2} & \dots & 1 + h_{nn}
\end{vmatrix} \\
&= 1 + (h_{11} + h_{22} + \cdots + h_{nn}) + \text{termes d'ordre supérieur ou égal à 2 en } h \\
&= 1 + \operatorname{tr} H + o(H)
\end{aligned}
$$

2. Or, tr est une forme linéaire, donc det est différentiable en I_n et $\mathrm{d}\det(I_n) \cdot H = \operatorname{tr} H$ (oui, c'est moche).
3. Comme $A \in \mathrm{GL}_n(\mathbb{R})$, on a

$$
\begin{aligned}
\det(A + H) = \det\left(A \cdot (I_n + A^{-1} \cdot H)\right) &= \det A \times \det(I_n + A^{-1} \cdot H) \\
&= \det A \times (1 + \operatorname{tr}(A^{-1}H) + \text{reste}') \\
&= \det A + \det A \times \operatorname{tr}(A^{-1}H) + \text{reste}
\end{aligned}
$$

où reste $= \det(A) \cdot \|A^{-1}H\|\varepsilon(A^{-1}H) = \|H\|\varepsilon'(H)$ car $\|A^{-1}H\| \leqslant \|A^{-1}\| \times \|H\|$ en choisissant une norme sous-multiplicative et $\|A^{-1}\| \times \det(A)\varepsilon(A^{-1}H) \longrightarrow 0$ quand $H \longrightarrow 0$ car $A^{-1}H \longrightarrow 0$, car $0 \leqslant \|A^{-1}H\| \leqslant \|H\| \longrightarrow 0$ quand $\|H\| \longrightarrow 0$ avec la norme sous-multiplicative. Donc, det est différentiable en $A \in \mathrm{GL}_n(\mathbb{R})$ et

$$\mathrm{d}\det(A) \cdot H = \det(A) \times \mathrm{tr}(A^{-1}H).$$

PROPOSITION 13:

$$\begin{tikzcd}
f \text{ est de classe } \mathscr{C}^1 \arrow[d, Rightarrow, "\text{Taylor Young (plus tard...)}"] \arrow[dr, Rightarrow, bend left] & \\
f \text{ est différentiable en } \vec{a} \arrow[r, Rightarrow] \arrow[d, Rightarrow, shift left] & f \text{ est continue en } \vec{a} \\
f \text{ possède des dérivées partielles en } \vec{a} \arrow[u, Rightarrow, shift left, "\times" description] \arrow[ur, Rightarrow, bend right, "\times" description] &
\end{tikzcd} \cdot$$

DÉMONSTRATION:
Supposons f différentiable en $\vec{a}$. Montrons que f est continue en $\vec{a}$. On sait que $f(\vec{a}+\vec{h}) = f(\vec{a})+\mathrm{d}f(\vec{a})\cdot\vec{h}+\|\vec{h}\|\varepsilon(\vec{h})$. Montrons que $f(\vec{a}+\vec{h}) \longrightarrow f(\vec{a})$ quand $\vec{a} \longrightarrow \vec{h}$. Montrons donc que $f(\vec{a}+\vec{h}) - f(\vec{a}) = \hat{\varepsilon}(\vec{h})$. Or, $f(\vec{a}+\vec{h}) - f(\vec{a}) = \mathrm{d}f(\vec{a}) \cdot \vec{h} + \|\vec{h}\|\varepsilon(\vec{h})$. On a $\|\vec{h}\|\varepsilon(\vec{h}) \longrightarrow 0$ quand $\vec{h} \longrightarrow \vec{0}$. Et, $\mathrm{d}f(\vec{a})$ est linéaire sur un espace vectoriel de dimension finie, donc $\mathrm{d}f(\vec{a})$ est continue, d'où $\mathrm{d}f(\vec{a}) \cdot \vec{h} \longrightarrow \mathrm{d}f(\vec{a}) \cdot \vec{0} = \vec{0}$ car $\mathrm{d}f(\vec{a})$ est linéaire. D'où,

$$f(\vec{a}+\vec{h}) - f(\vec{a}) = \underbrace{\mathrm{d}f(\vec{a}) \cdot \vec{h}}_{\to 0} + \underbrace{\|\vec{h}\|\varepsilon(\vec{h})}_{\to 0} \xrightarrow[\vec{h}\to\vec{0}]{} \vec{0}.$$

EXERCICE 14 (Une fonction qui possède des dérivées partielles, mais pas continue):
Soit la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *définie par*

$$f(x,y) = \frac{xy}{x^2+y^2} \text{ si } (x,y) \neq (0,0) \qquad \text{et} \qquad f(0,0) = 0.$$

Montrer que la fonction f possède des dérivées partielles $\partial_1 f(0,0)$ et $\partial_2 f(0,0)$ en $(0,0)$ mais que f n'est pas continue en $(0,0)$.

La fonction f n'est pas continue en $(0,0)$ car $f(x,x) \longrightarrow \frac{1}{2} \neq f(0,0)$ quand $x \longrightarrow 0$. D'une part, $\big(f(0+h,0) - f(0,0)\big)/h = \big(f(h,0)\big)/h = 0/h = 0 \longrightarrow 0$, lorsque $h \longrightarrow 0$. D'où, $\partial_1 f(0,0)$ existe et $\partial_1 f(0,0) = 0$. De même, $\partial_2 f(0,0) = 0$ par symétrie.

REMARQUE 15:
Soit a un élément de $I \subset \mathbb{R}$ un intervalle ouvert. Une fonction $f : I \longrightarrow \mathbb{R}$ est dérivable en a si, et seulement si elle est différentiable en a. Et alors, $f'(a) = \mathrm{d}f(a) \cdot 1$.

DÉFINITION 16:
Si $f : E \longrightarrow F$ est une fonction différentiable en chaque point $\vec{a} \in U$ d'une partie $U \subset E$, alors on dit que f est *différentiable* sur U et l'application

$$\begin{aligned} \mathrm{d}f : U &\longrightarrow \mathscr{L}(E,F) \\ \vec{a} &\longmapsto \mathrm{d}f(\vec{a}) \end{aligned}$$

est appelée la *différentielle* de f sur U. Si chacune des n fonctions f_i est de classe $\mathscr{C}^1$ sur U, alors on dit que f est de classe $\mathscr{C}^1$ sur U.

PROPOSITION 17: 1.

$$\begin{tikzcd}
f \text{ est de classe } \mathscr{C}^1 \text{ sur } U \arrow[r, Rightarrow, bend left] & f \text{ est différentiable sur } U \arrow[l, Rightarrow, bend left, "\times" description]
\end{tikzcd} .$$

2. La fonction f est $\mathscr{C}^1$ sur U si, et seulement si f est différentiable sur U et sa différentielle $\mathrm{d}f$ est continue sur U.

EXERCICE 18 (Une fonction différentiable mais pas $\mathscr{C}^1$):
Montrer que la fonction f définie à l'exercice 5 est différentiable sur $\mathbb{R}^2$ mais n'est pas $\mathscr{C}^1$ sur $\mathbb{R}^2$.

On a déjà montré que f n'est pas $\mathscr{C}^1$, *i.e.* $\partial_1 f$ ou $\partial_2 f$ n'est pas continue (*c.f.* exercice 5). Mais f est différentiable. En effet, montrons qu'il existe α et β deux réels tels que pour tout vecteur $\vec{h} = (h,k) \in \mathbb{R}^2$, $f(a+h, b+k) = f(a,b) + \alpha h + \beta k + o(\vec{h})$.

— si $a \neq 0$, alors f est $\mathscr{C}^1$ et, d'après la formule de TAYLOR-YOUNG, f est différentiable.

— si $a = 0$, alors $f(0+h, b+k) = h \times h \sin\left(\frac{b+k}{h}\right) = h \times o(h)$ car $\sin\left(\frac{b+k}{h}\right)$ est bornée. Ainsi, $f(h, b+k) = f(0,b) + 0h + 0k + o(h)$.

La fonction f est donc différentiable sur $\mathbb{R}^2$.

EXEMPLE 19: 1. La norme « deux » définie sur $\mathbb{R}^p$ par $N(\vec{x}) = \sqrt{{x_1}^2 + \cdots + {x_p}^2}$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^p \setminus \{\vec{0}\}$ et

$$\forall \vec{a} \in \mathbb{R}^p \setminus \{\vec{0}\},\ \forall \vec{h} \in \mathbb{R}^p, \qquad \mathrm{d}N(\vec{a}) \cdot \vec{h} = \frac{\langle\, \vec{a} \mid \vec{h}\,\rangle}{N(\vec{a})},$$

où $\langle\, \vec{a} \mid \vec{h}\,\rangle = \sum_{i=1}^p a_i h_i$ est le produit scalaire des vecteurs $\vec{a}$ et $\vec{b}$.

2. La fonction $\det : \mathscr{M}_n(\mathbb{R}) \to \mathbb{R}$ est de classe $\mathscr{C}^1$ et

$$\forall A \in \mathscr{M}_n(\mathbb{R}),\ \forall H \in \mathscr{M}_n(\mathbb{R}), \qquad \det(A+H) = \det A + \operatorname{tr}(B^\top \cdot H) + o(H),$$

où $B = \operatorname{com} A$ est la comatrice de A.
On a déjà montré le résultat pour $A \in \mathrm{GL}_n(\mathbb{R})$ (exercice 12). En effet, $B^\top = \det(A)A^{-1}$, et on conclut par linéarité de tr.
Montrons maintenant le résultat pour A quelconque. On sait déjà que det est de classe $\mathscr{C}^\infty$ car $\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$ (par les « théorèmes généraux »). En développant selon la colonne j, pour la ligne i, on trouve que le terme devant $a_{i,j}$ est $(-1)^{i+j}\Delta_{i,j}$. En effet, aucun terme ne contient $a_{i,j}$ car on supprime la j-ième colonne. Ainsi,

$$\frac{\partial \det}{\partial a_{i,j}} = (-1)^{i+j}\Delta_{i,j},$$

qui est une fonction continue. Ainsi, det est de classe $\mathscr{C}^1$, et, d'après le formule de TAYLOR-YOUNG,

$$\begin{aligned}
\det(A+H) &= \det A + \sum_{i=1}^n \sum_{j=1}^n h_{i,j} \overbrace{(-1)^{i+j}\Delta_{i,j}}^{[\operatorname{com} A]_{i,j}} + o(H) \\
&= \det A + \sum_{(i,j) \in [\![1,n]\!]^2} h_{i,j} b_{i,j} + o(H) \\
&= \det A + \operatorname{tr}(B^\top \cdot H) + o(H)
\end{aligned}$$

car on reconnaît la formule du produit scalaire sur $\mathscr{M}_n(\mathbb{R})$.

REMARQUE 20:
De ce dernier exemple, il en résulte que la fonction $\det : \mathscr{M}_n(\mathbb{R}) \to \mathbb{R}$ est continue (car $\mathscr{C}^1 \implies$ différentiable $\implies \mathscr{C}^0$). D'où, $\mathrm{GL}_n(\mathbb{R}) = \det^{-1}(\mathbb{R}^\star)$ est l'image réciproque, par la fonction continue det, de l'ouvert $\mathbb{R}^\star$. Ainsi, $\mathrm{GL}_n(\mathbb{R})$ est un ouvert de $\mathscr{M}_n(\mathbb{R})$. On avait aussi montré que $\mathrm{GL}_n(\mathbb{R})$ est dense dans $\mathscr{M}_n(\mathbb{R})$ (exercice 26 du chapitre 13).

4 Le gradient

L'espace vectoriel E était, jusqu'ici, normé et de dimension finie. On le munit désormais d'un produit scalaire ; E est donc un espace euclidien.

PROPOSITION – DÉFINITION 21:
Si $f : E \longrightarrow \mathbb{R}$ est différentiable en $\vec{a} \in E$, alors il existe un unique vecteur de E, appelé le *gradient* de f en $\vec{a}$, et noté $\nabla f(\vec{a})$ ou grad $f(\vec{a})$, tel que

$$\forall \vec{h} \in E, \qquad \mathrm{d}f(\vec{a}) \cdot \vec{h} = \langle \vec{h} \mid \nabla f(\vec{a}) \rangle$$

est le le produit scalaire du vecteur déplacement $\vec{h}$ et du gradient de f en $\vec{a}$. Si $(\vec{e}_1, \dots, \vec{e}_p)$ est une base orthonormée de E, alors $\nabla f(\vec{a}) = \partial_1 f(\vec{a})\, \vec{e}_1 + \cdots + \partial_p f(\vec{a})\, \vec{e}_p$.

EXEMPLE 22:
La fonction f définie par $f(x, y) = \ln(x^2 + y^2)$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2 \setminus \{(0, 0)\}$ et

$$\forall (x, y) \neq (0, 0), \qquad \nabla f(x, y) = \left(\frac{2x}{x^2 + y^2}, \frac{2y}{x^2 + y^2} \right).$$

La figure ci-dessous représente, en chaque point (x, y) différent de l'origine, le gradient de f en (x, y). On obtient ainsi un champ de vecteurs.

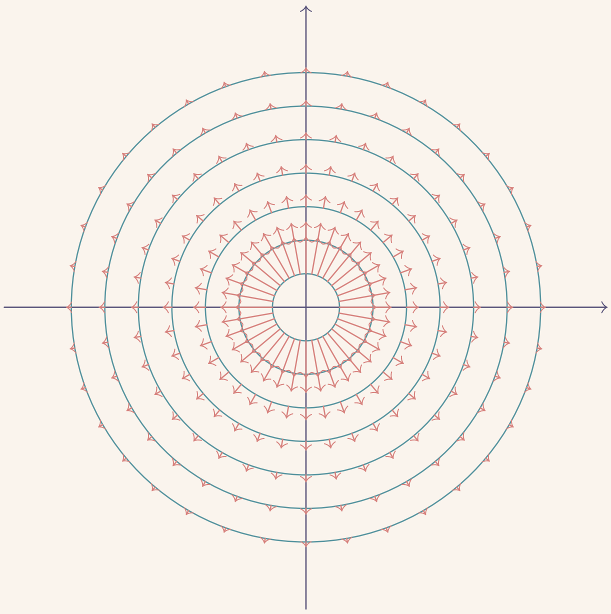


FIGURE 6 – Le champ des gradients de $(x, y) \mapsto \ln(x^2 + y^2)$

On se déplace dans un espace euclidien E, le long d'une courbe paramétrée par $M : t \mapsto M(t)$, et on évalue, à chaque instant t, la valeur de $f\big(M(t)\big)$ prise par une fonction scalaire f en un point $M(t)$.

LEMME 23 (Règle de la chaîne):
Soit $f : U \longrightarrow \mathbb{R}$ définie sur un ouvert U de l'espace euclidien E. (Dans ce lemme, on se place en dimension 2, mais ce résultat est vrai dans un espace E de dimension p.) Soit $M : I \longrightarrow E$ définie sur un intervalle I de $\mathbb{R}$. Si $M(I) \subset U$ et les fonctions f et M sont différentiables, alors

$f \circ M : t \mapsto f\big(M(t)\big)$ est différentiable et, pout tout $t \in I$,

$$\begin{aligned}(f \circ M)'(t) &= \frac{\mathrm{d}}{\mathrm{d}t} f\big(M(t)\big) \\ &= x'(t) \cdot \frac{\partial f}{\partial x}\big(M(t)\big) + y'(t)\frac{\partial f}{\partial y}\big(M(t)\big) \\ &= \langle\, M'(t) \mid \nabla f(M(t)) \,\rangle \\ &= \mathrm{d}f\big(M(t)\big) \cdot M'(t)\end{aligned}$$

est le produit scalaire du vecteur vitesse $M'(t)$ et du gradient de f en M. De plus, si f et M sont de classe $\mathscr{C}^1$, alors $f \circ M$ l'est aussi.

Démonstration:
On pose, pour tout réel t, $M(t) = \big(x(t), y(t)\big)$. On suppose f différentiable, et M dérivable (par rapport à t). Montrons que $f \circ M$ est dérivable, et calculer $(f \circ M)'(t)$. On calcule $f \circ M(t+u) - f \circ M(t) = f\big(x(t+u), y(t+u)\big) - f\big(x(t), y(t)\big)$. Or, comme x est dérivable, et $x(t+u) = x(t) + u\,x'(t) + u\varepsilon_1(u)$. En effet, cette expression est équivalente à $\big(x(t+u) - x(t)\big)/u = x'(t) + \varepsilon_1(u)$, qui tend vers $x'(t)$ quand $u \longrightarrow 0$. De même, $y(t+u) = y(t) + u\,y'(t) + u\,\varepsilon_2(u)$. D'où,

$$f\big(x(t+u), y(t+u)\big) = f\big(x(t) + \underbrace{u\,x'(t) + u\,\varepsilon_1(u)}_{h}, y(t) + \underbrace{u\,y'(t) + u\,\varepsilon_2(u)}_{k}\big).$$

Or, par hypothèse, f est différentiable, d'où $\delta = f(x(t), y(t)) + \mathrm{d}f(x(t), (y)) \cdot \vec{h} + \|\vec{h}\|\,\varepsilon(\vec{h})$, en posant $\vec{h} = (h, k)$. Ainsi,

$$\begin{aligned}f \circ M(t+u) - f \circ M(t) &= \mathrm{d}f\big(x(t), y(t)\big) \cdot \vec{h} + \|\vec{h}\| \cdot \varepsilon(\vec{h}) \\ &= \big\langle \nabla f(M(t)) \,\big|\, (u\,x'(t) + u\,\varepsilon_1(u), u\,y'(t) + u\,\varepsilon_2(u)) \big\rangle \\ &+ \|(u\,x'(t) + u\,\varepsilon_1(u), u\,y'(t) + u\,\varepsilon_2(u))\| + \varepsilon\big((u\,x'(t) + u\,\varepsilon_1(u), u\,y'(t) + u\,\varepsilon_2(u))\big) \\ &= u\Big(\big\langle \nabla f(M(t)) \,\big|\, (x'(t) + \varepsilon_1(u), y'(t) + \varepsilon_2(u)) \big\rangle \\ &+ \|(x'(t) + \varepsilon_1(u), y'(t) + \varepsilon_2(u))\| + \varepsilon\big((x'(t) + \varepsilon_1(u), y'(t) + \varepsilon_2(u))\big)\Big)\end{aligned}$$

D'où,

$$\frac{f \circ M(t+u) - f \circ M(t)}{u} \xrightarrow[u \to 0]{} \big\langle \nabla f\big(M(t)\big) \,\big|\, M'(t) \big\rangle.$$

Dans le lemme précédent, les fonctions f et M ont la représentation suivante.

$$\begin{array}{ccccc} t & \xmapsto{M} & M(t) & \xmapsto{f} & f\big(M(t)\big) \\ I & \longrightarrow & U & \longrightarrow & \mathbb{R} \\ t & & \xmapsto{\quad f \circ M \quad} & & f \circ M(t) \end{array}$$

On en déduit que $(f \circ M)'(t) = 0$ si, et seulement si, le vecteur vitesse $M'(t)$ est orthogonal au gradient de f en $M(t)$, à un instant t. En particulier, si la fonction f est constante le long de la trajectoire, *i.e.* si la trajectoire $M(I)$ est incluse dans la courbe de niveau de f, alors le gradient de f est orthogonal au vecteur vitesse en chaque point de la trajectoire.

Définition 24:
On dit qu'un vecteur $\vec{v} \in E$ est *tangent* à une partie $C \subset E$ en un point $\vec{a} \in C$ s'il existe un application dérivable $M : I \longrightarrow E$ telle que $M(I) \subset C$, et s'il existe $t_0 \in I$ tel que $M(t_0) = \vec{a}$ et $M'(t_0) = \vec{v}$. On note $\mathrm{T}_{\vec{a}}C$ l'ensemble des vecteurs tangents à C en $\vec{a}$, il est nommé *l'espace tangent*.

On dit alors qu'un vecteur est *orthogonal* à la partie C s'il est orthogonal à $\mathrm{T}_{\vec{a}}C$.

« Un vecteur est tangent s'il est un vecteur vitesse. »

THÉORÈME 25 (admis):
Soient $f : E \to \mathbb{R}$ une fonction de classe $\mathscr{C}^1$, et une partie $C \subset E$, telle que f est constante sur C. Soit $\vec{a} \in C$. Si $\nabla f(\vec{a}) \neq \vec{0}$, alors un vecteur $\vec{v} \in E$ est tangent à C si, et seulement si $\vec{v}$ est orthogonal à $\nabla f(\vec{a})$.

Autrement dit : $\mathrm{T}_{\vec{a}}C$ est l'hyperplan $\left[\mathrm{Vect}\left(\nabla f(\vec{a})\right)\right]^{\perp} = \mathrm{Ker}\, \mathrm{d}f(\vec{a})$.

Le gradient de f en $\vec{a}$ est donc orthogonal à C : $\nabla f(\vec{a}) \perp \mathrm{T}_{\vec{a}}C$. Ainsi, le champ électrostatique est orthogonal aux équipotentielles, la ligne de plus petite pente est orthogonal aux lignes de niveau, *etc.*

EXERCICE 26:
Déterminer une équation de la droite tangente à l'hyperbole d'équation $x^2 - y^2 = 1$ au point $(\sqrt{2}, 1)$.

On nomme $\mathscr{H}$ l'hyperbole donnée dans l'énoncé. On a bien $A = (\sqrt{2}, 1) \in \mathscr{H}$ car $\sqrt{2}^2 - 1^2 = 1$. On cherche une équation de $\mathrm{T}_A\mathscr{H}$.

— **Avec le théorème 25.** Soit la fonction f définie comme
$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^2 - y^2. \end{aligned}$$
Cette fonction f est de classe $\mathscr{C}^1$ et est constante sur $\mathscr{H}$. On calcule le gradient de f en A : $\nabla f(A) = (\partial f(\sqrt{2}, 1)/\partial x, \partial f(\sqrt{2}, 1)/\partial y) = 2(\sqrt{2}, -1)$.
$$\begin{aligned} N = (x, y) \in \mathrm{T}_A\mathscr{H} &\iff \overrightarrow{AN} \perp \nabla f(A) \\ &\iff (x - \sqrt{2}, y - 1) \perp (\sqrt{2}, -1) \\ &\iff \sqrt{2}(x - \sqrt{2}) - 1(y - 1) = 0 \\ &\iff \sqrt{2}x - y = 1 \\ &\iff y = \sqrt{2}x - 1 \end{aligned}$$

— **Avec la définition 24.** On rappelle que
$$(x, y) \in \mathscr{H}_+ \iff (x > 0 \text{ et } x^2 - y^2 = 1) \iff \exists t \in \mathbb{R},\ (x = \operatorname{ch} t \text{ et } y = \operatorname{sh} t).$$
Soit $M : t \mapsto \big(x(t), y(t)\big) = (\operatorname{ch} t, \operatorname{sh} t)$ un point de $\mathscr{H}_+$. Il existe un instant $t_0 \in \mathbb{R}$ tel que $M(t_0) = (\sqrt{2}, 1)$. La fonction M est dérivable et $M'(t) = (\operatorname{sh} t, \operatorname{ch} t) = (1, \sqrt{2})$ est tangent à $\mathscr{H}$.
$$N = (x, y) \in \mathrm{T}_A\mathscr{H} \iff \overrightarrow{AN} \mathbin{//} M'(t_0) \iff \begin{vmatrix} x - \sqrt{2} & 1 \\ y - 1 & \sqrt{2} \end{vmatrix} = 0 \iff y = \sqrt{2}x - 1.$$

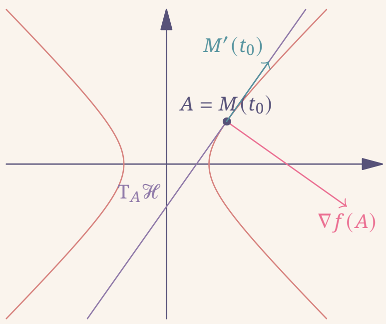


FIGURE 7 – Hyperbole $\mathscr{H}$, gradient en A et tangente en A.

EXEMPLE 27:
La sphère $S \subset \mathbb{R}^3$ de rayon 1 et de centre $\vec{0} = (0, 0, 0)$ est la surface paramétrée par
$$\forall (\varphi, \theta) \in \mathbb{R}^2, \qquad \begin{cases} x(\varphi, \theta) = \cos\theta \cos\varphi \\ y(\varphi, \theta) = \cos\theta \sin\varphi \\ z(\varphi, \theta) = \sin\theta. \end{cases}$$

Les vecteurs $\partial M(\varphi,\theta)/\partial\varphi) = (-\cos\theta\sin\varphi, \cos\theta\cos\varphi, 0)$ et $\partial M(\varphi,\theta)/\partial\theta = (-\sin\theta\cos\varphi, -\sin\theta\sin\varphi, \cos\theta)$ sont tangents à la sphère en le point $M(\varphi,\theta)$. Le plan tangent à la sphère au point (a,b,c) est orthogonal au vecteur

$$\frac{\partial M}{\partial\varphi}(\varphi,\theta) \wedge \frac{\partial M}{\partial\theta}(\varphi,\theta) = (\cos^2\theta\cos\varphi, \cos^\theta\sin\varphi, \cos\theta\sin\varphi) = \cos\theta\,(a,b,c),$$

qui est non nul si $\cos\theta \neq 0$ (*i.e.* si le point $M(\varphi,\theta) = (a,b,c)$ n'est, si au pôle Sud, ni au pôle Nord).

En effet, on cherche une équation du plan P tangent.

$$\begin{aligned} N(x,y,z) \in P &\iff \det(\overrightarrow{MN}, \partial M/\partial\varphi, \partial M/\partial\theta) = 0 \\ &\iff \begin{vmatrix} x-a & ? & ? \\ y-b & ? & ? \\ z-c & ? & ? \end{vmatrix} = 0 \\ &\iff ?\cdot(x-a) + ?\cdot(y-b) + ?\cdot(z-c) = 0 \end{aligned}$$

Autre méthode : $$N(x,y,z) \in P \iff \overrightarrow{MN} \perp \left(\frac{\partial M}{\partial\varphi} \wedge \frac{\partial M}{\partial\theta}\right) \iff \left\langle \vec{MN} \,\middle|\, \frac{\partial M}{\partial\varphi} \wedge \frac{\partial M}{\partial\theta} \right\rangle = 0.$$

Autre autre méthode : la même sphère S est la surface de niveau 1 de la fonction $f : \mathbb{R}^3 \to \mathbb{R}$, $(x,y,z) \mapsto x^2+y^2+z^2$. En chaque point (a,b,c) de la sphère le plan tangent a pour équation $ax+by+cz=0$. On calcule le gradient de f au point (a,b,c) :

$$\nabla f(a,b,c) = \big(\partial_1 f(a,b,c), \partial_2 f(a,b,c), \partial_3 f(a,b,c)\big) = (2a, 2b, 2c) = 2(a,b,c).$$

Ainsi,

$$\begin{aligned} N(x,y,z) \in P \iff \overrightarrow{MN} \perp \nabla f(M) &\iff \left\langle \begin{pmatrix} x-a \\ y-b \\ z-c \end{pmatrix} \middle| \begin{pmatrix} a \\ b \\ c \end{pmatrix} \right\rangle = 0 \\ &\iff a(x-a) + b(y-b) + c(z-c) = 0 \\ &\iff ax + by + cz = a^2+b^2+c^2 = 1 \end{aligned}$$

Définition 28: 1. On dit d'une partie D d'un espace vectoriel qu'elle est *convexe* si, pour tous vecteurs $\vec{a}$ et $\vec{b}$ de D, pour tout $t \in [0,1]$, alors $\vec{a} + t(\vec{b}-\vec{a}) = (1-t)\vec{a} - t\vec{b}$.

« D est convexe si tout segment $[\vec{a}, \vec{b}]$ est inclus dans D. »

2. On dit d'une partie D d'un espace vectoriel normé quelle est *connexe par arcs* si, pour tous vecteurs $\vec{a}$ et $\vec{b}$ de D, il existe une application continue $M : [0,1] \longrightarrow E$ telle que $M(0) = a$, $M(1) = b$ et $\forall t \in [0,1]$, $M(t) \in D$.

« D est connexe par arcs s'il existe un chemin de $\vec{a}$ vers $\vec{b}$ pour tous vecteurs $\vec{a}$ et $\vec{b}$. »

Remarque 29: 1. Le segment $[\vec{a}, \vec{b}]$ est l'ensemble des vecteurs $M(t) = \vec{a} + t(\vec{b}-\vec{a}) = (1-t)\vec{a} + t\vec{b}$, où $t \in [0,1]$.

2. $$\begin{tikzcd} D \text{ est convexe} \arrow[r, Rightarrow, bend left] & D \text{ est connexe par arcs} \arrow[l, Rightarrow, bend left, "\times"] \end{tikzcd}.$$

Mais, **dans** $\mathbb{R}$, on a bien D connexe $\iff$ D connexe par arcs $\iff$ D est un intervalle.

3. L'image directe $f(D)$ d'une partie D connexe par arcs de E par une application continue $f : E \longrightarrow F$ est une partie connexe par arcs de F.

Démonstration:
En effet, soient $\vec{a}'$ et $\vec{b}'$ deux vecteurs de $f(D)$. Il existe $\vec{a}$ et $\vec{b}$ deux vecteurs de D tels que $f(\vec{a}) = \vec{a}'$ et $f(\vec{b}) = \vec{b}'$. Par hypothèse, il existe une application $M : [0,1] \to E$ continue telle que $M([0,1]) \subset D$, $M(\vec{0}) = \vec{a}$ et $M(1) = \vec{b}$.

$$t \xmapsto{f\circ M} f\circ M(t)$$

$$\xrightarrow{\qquad\qquad} F \quad .$$

$$t \xmapsto{M} M(t) \xmapsto{f} f(M(t))$$

Or, $f \circ M(0) = f(M(0)) = f(\vec{a}) = \vec{a}'$, $f \circ M(1) = f(M(1)) = f(\vec{b}) = \vec{b}'$. De plus $f \circ M$ est continue car c'est la composée de deux fonctions continues f et M (par hypothèse). Ainsi, $f \circ M([a,b]) \subset f(D)$. On en déduit que $f(D)$ est connexe par arcs.

4. D'après les deux points précédents, toute fonction continue sur partie connexe par arcs vérifie le théorème des valeurs intermédiaires.

Rappel (Théorème des valeurs intermédiaires):
Soit $f : \mathbb{R} \supset [a,b] \longrightarrow \mathbb{R}$ une fonction continue sur $[a,b]$. Pour tout $y \in [f(a), f(b)] \cup [f(b), f(a)]$, il existe $x \in [a,b]$ tel que $y = f(x)$.

> Théorème:
> Le théorème des valeurs intermédiaires devient donc, si $f : D \longrightarrow \mathbb{R}$ est continue, où $\vec{a}, \vec{b} \in D$ est une partie connexe par arcs de E, alors, pour tout $y \in [f(\vec{a}), f(\vec{b})] \cup [f(\vec{b}), f(\vec{a})]$, il existe $\vec{c} \in D$ tel que $y = f(\vec{c})$.

Démonstration:
De $\vec{a}$ à $\vec{b}$, il existe un chemin : il existe une fonction $M : [0,1] \longrightarrow D$ continue telle que $M(0) = \vec{a}$, $M(1) = \vec{b}$, et $\forall t \in [0,1]$, $M(t) \in D$. La fonction $f \circ M : t \mapsto f(M(t))$ est continue par composition. On conclut par le théorème des valeurs intermédiaires appliqué à $f \circ M$.

Exercice 30: 1. *Montrer que toute boule de tout espace vectoriel normé est convexe.*

2. *Montrer que l'hyperbole $\mathscr{H}$ d'équation $y^2 - x^2 = 1$ n'est pas connexe par arcs.*
3. *Montrer que le produit cartésien $D_1 \times D_2$ de deux parties connexes par arcs $D_1 \subset E_1$ et $D_2 \subset E_2$ est une partie connexe par arcs de $E_1 \times E_2$. En déduire que la sphère $\{\vec{x} \in \mathbb{R}^3 \mid \|\vec{x}\|_2 = 1\}$ est une partie connexe par arcs de $\mathbb{R}^3$, et que, pour aller d'un pôle à l'autre, il faut traverser l'équateur.*

1. On considère la boule $B(\vec{\Omega}, R)$, où $\vec{\Omega}$ est un vecteur d'un espace vectoriel normé $(E, \|\cdot\|)$, et $R > 0$. Soient $\vec{a}, \vec{b} \in B(\vec{\Omega}, R)$; ainsi, $\|\vec{a} - \vec{\Omega}\| < R$ et $\|\vec{a} - \vec{\Omega}\| < R$. Montrons que, pour tout $t \in [0,1]$, $(1-t)\vec{a} + t\vec{b} \in B(\vec{\Omega}, R)$.

$$\begin{aligned}
\|(1-t)\vec{a} + t\vec{b} - \vec{\Omega}\| &= \|(1-t)\vec{a} + t\vec{b} - \vec{\Omega}(1-t) - t\vec{\Omega}\| \\
&= \|(1-t)(\vec{a} - \vec{\Omega}) + t(\vec{b} - \vec{\Omega})\| \\
&\leqslant (1-t)\|\vec{a} - \vec{\Omega}\| + t\|\vec{b} - \vec{\Omega}\| \\
&\leqslant (1-t) \cdot R + t \cdot R \\
&= R
\end{aligned}$$

D'où, $\big((1-t)\vec{a} + t\vec{b}\big) \in B(\vec{\Omega}, R)$.

2. Par l'absurde, on suppose $\mathscr{H}$ connexe. Soit la fonction $f : (x,y) \mapsto y$ continue. L'image d'un connexe par une fonction continue est connexe, d'où $f(\mathscr{H})$ est connexe. Or, $f(\mathscr{H}) = \mathbb{R} \setminus \left]-1, 1\right[= \left]-\infty, -1\right[\cup \left]-1, \infty\right[$, qui n'est pas connexe.

> Proposition 31:
> Soit f une fonction de classe $\mathscr{C}^1$ sur un **convexe** D. La fonction f est constante sur D si, et seulement si, son gradient $\nabla f(\vec{x})$ est nul en tout point $\vec{x} \in D$. De même si D est connexe par arcs.

Démonstration:
L'implication est évidente, une fonction constante a un gradient nul. Montrons la réciproque. D'après la règle de la chaîne, pour toute fonction dérivable M, on a $(f \circ M)'(t) = \langle \nabla f(M(t)) \mid M'(t) \rangle = 0$ car $\nabla f(M(t))$ est nul par hypothèse. Montrons que, pour tout vecteurs $\vec{a}$ et $\vec{b}$ de D, $f(\vec{a}) = f(\vec{b})$. On suppose f de classe $\mathscr{C}^1$. On pose M un chemin de $\vec{a}$ vers $\vec{b}$, on admet que M est de classe $\mathscr{C}^1$, et on a

$$f(\vec{b}) - f(\vec{a}) = f \circ M(1) - f \circ M(0) = \int_0^1 (f \circ M)'(t)\, \mathrm{d}t = \int_0^1 0\, \mathrm{d}t = 0.$$

5 La différentielle d'une composée

Proposition 32:
Soient E, F et G trois espaces vectoriels normés, et soient $U \subset E$ et $V \subset F$ des ouverts. Soit $f : U \to F$ et $g : V \to G$ deux fonctions telles que $f(U \subset V)$. Ainsi,

$$\begin{array}{ccccc} E \supset U & \longrightarrow & F \supset V & \longrightarrow & G \\ \vec{x} & \overset{f}{\longmapsto} & \vec{y} = f(\vec{x}) & \overset{g}{\longmapsto} & g(\vec{y}) = (g \circ f)(\vec{x}) \end{array}.$$

Si f est différentiable en $\vec{a} \in U$, et g est différentiable en $\vec{b} = f(\vec{a})$, alors $g \circ f$ est différentiable en $\vec{a}$, et

$$\boxed{\mathrm{d}(g \circ f)(\vec{a}) = \mathrm{d}g(\vec{b}) \circ \mathrm{d}f(\vec{a}).}$$

Démonstration:
On suppose f et g différentiable. Montrons que $g \circ f$ est différentiable. La fonction f est différentiable, on calcule donc $f(\vec{a}+\vec{h}) = f(\vec{a}) + \mathrm{d}f(\vec{a}) \cdot \vec{h} + \|\vec{h}\|\, \varepsilon_1(\vec{h}) = \vec{b} + \vec{k}$, en posant $\vec{b} = f(\vec{a})$. Et,

$$\begin{aligned} g(\vec{b}+\vec{k}) &= g(\vec{b}) + \mathrm{d}g(\vec{b}) \cdot \vec{k} + \|\vec{k}\|\, \varepsilon_2(\vec{k}) \\ &= g \circ f(\vec{a}) + \mathrm{d}g(\vec{b}) \cdot \left[\mathrm{d}f(\vec{a}) \cdot \vec{h} + \|\vec{h}\|\, \varepsilon_1(\vec{h})\right] + \|\mathrm{d}f(\vec{a}) \cdot \vec{h} + \vec{h}\, \varepsilon_1(\vec{h})\|\, \varepsilon_2(\mathrm{d}f(\vec{a}) \cdot \vec{h} + \|\vec{h}\|\, \varepsilon_1(\vec{h})) \\ &= g \circ f(\vec{a}) + \left[\mathrm{d}g(\vec{b}) \circ \mathrm{d}f(\vec{a})\right] \cdot \vec{h} + o(\vec{h}) \end{aligned}$$

En effet, en nommant $\vec{R}$ le (futur) reste de $g(\vec{b}+\vec{k})$, on a $\vec{R} = \mathrm{d}f(\vec{b}) \cdot [\,\|\vec{h}\|\varepsilon_1(\vec{h})\,] + \|\mathrm{d}f(\vec{a}) \cdot \vec{h} + \|\vec{h}\|\varepsilon_1(\vec{h})\,\|\varepsilon_2\big(\mathrm{d}f(\vec{a}) \cdot \vec{h} + \|\vec{h}\|\varepsilon_1(\vec{h})\big)$.
D'où,

$$\begin{aligned} 0 &\leqslant \left\| \frac{\mathrm{d}g(\vec{b}) \cdot \left[\|\vec{h}\|\, \varepsilon_1(\vec{h})\right] + \left\|\mathrm{d}f(\vec{a}) \cdot \vec{h} + \|\vec{h}\|\, \varepsilon_1(\vec{h})\right\| \varepsilon_2(\mathrm{d}f(\vec{a}) \cdot \vec{h} + \|\vec{h}\|\, \varepsilon_1(\vec{h}))}{\|\vec{h}\|} \right\| \\ &\leqslant \frac{\left\|\mathrm{d}g(\vec{b}) \cdot \left[\|\vec{h}\|\, \varepsilon_1(\vec{h})\right]\right\| + \left\|\mathrm{d}f(\vec{a}) \cdot \vec{h} + \|\vec{h}\|\, \varepsilon_1(\vec{h})\right\| \left\|\varepsilon_2(\mathrm{d}f(\vec{a}) \cdot \vec{h} + \|\vec{h}\|\, \varepsilon_1(\vec{h}))\right\|}{\|\vec{h}\|} \qquad \text{par inégalité triangulaire} \\ &\leqslant \frac{\|\vec{h}\| \cdot \left\|\mathrm{d}g(\vec{b}) \cdot \left[\, \varepsilon_1(\vec{h})\right]\right\| + \left(\left\|\mathrm{d}f(\vec{a}) \cdot \vec{h}\right\| + \left\|\|\vec{h}\|\, \varepsilon_1(\vec{h})\right\|\right) \left\|\varepsilon_2(\mathrm{d}f(\vec{a}) \cdot \vec{h} + \|\vec{h}\|\, \varepsilon_1(\vec{h}))\right\|}{\|\vec{h}\|} \\ &\leqslant \frac{\|\vec{h}\| \cdot \left\|\mathrm{d}g(\vec{b}) \cdot \left[\, \varepsilon_1(\vec{h})\right]\right\| + \left(|||\mathrm{d}f(\vec{a})||| \cdot \|\vec{h}\| + \|\varepsilon_1(\vec{h})\| \cdot \|\vec{h}\|\right) \left\|\varepsilon_2(\mathrm{d}f(\vec{a}) \cdot \vec{h} + \|\vec{h}\|\, \varepsilon_1(\vec{h}))\right\|}{\|\vec{h}\|} \\ &\leqslant \|\mathrm{d}g(\vec{b}) \cdot \varepsilon_1(\vec{h})\| + \|(|||\mathrm{d}f(\vec{a})||| + \|\varepsilon_1(\vec{h})\|) \cdot \varepsilon_2(\ldots)\| \\ &\leqslant |||\mathrm{d}f(\vec{b})||| \cdot \|\varepsilon_1(\vec{h})\| + (|||\mathrm{d}f(\vec{a})||| + \|\varepsilon_1(\vec{h})\|) \cdot \|\varepsilon_2(\ldots)\| \xrightarrow[\vec{h}\to\vec{0}]{} 0 \end{aligned}$$

En généralisant la règle de la chaîne

$$\frac{\mathrm{d}}{\mathrm{d}t} f \circ M(t) = \frac{\partial f}{\partial x} \cdot \frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y} \cdot \frac{\mathrm{d}y}{\mathrm{d}t},$$

on obtient le corolaire suivant.

Corollaire 33:
Avec les hypothèses précédentes, en notant $p = \dim E$, $q = \dim F$, $n = \dim G$, et $(y_1, \ldots, y_k) = \vec{y} = f(\vec{x})$, on a

$$\forall j \in [\![1, p]\!], \qquad \frac{\partial}{\partial x_j} g \circ f(\vec{a}) = \sum_{k=0}^{q} \underbrace{\frac{\partial g}{\partial y_k}}_{\partial_k g} \cdot \frac{\partial f_k}{\partial x_k}.$$

On peut exprimer le même résultat avec la jacobienne :

$$\boxed{\mathrm{J}_{g\circ f}(\vec{a}) = \mathrm{J}_g(\vec{b}) \cdot \mathrm{J}_f(\vec{a}),}$$

qui représente matriciellement l'égalité $\mathrm{d}(g \circ f)(\vec{a}) = \mathrm{d}g(\vec{b}) \circ \mathrm{d}f(\vec{a})$.

DÉMONSTRATION:
On a

$$\underbrace{\begin{pmatrix} \partial_1(g\circ f)_1 & \partial_2(g\circ f)_1 & \dots & \partial_p(g\circ f)_1 \\ \partial_1(g\circ f)_2 & \partial_2(g\circ f)_2 & \dots & \partial_p(g\circ f)_2 \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1(g\circ f)_p & \partial_2(g\circ f)_p & \dots & \partial_p(g\circ f)_p \end{pmatrix}}_{\mathrm{J}_{g\circ f}} = \underbrace{\begin{pmatrix} \partial_1 g_1 & \partial_2 g_2 & \dots & \partial_q g_1 \\ \partial_1 g_2 & \partial_2 g_2 & \dots & \partial_q g_2 \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 g_n & \partial_2 g_n & \dots & \partial_q g_n \end{pmatrix}}_{\mathrm{J}_g} \cdot \underbrace{\begin{pmatrix} \partial_1 f_1 & \partial_2 f_1 & \dots & \partial_p f_1 \\ \partial_1 f_2 & \partial_2 f_2 & \dots & \partial_p f_2 \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 f_q & \partial_2 f_q & \dots & \partial_p f_q \end{pmatrix}}_{\mathrm{J}_f},$$

ce qui donne l'autre expression en « regardant coordonnées par coordonnées. »

EXEMPLE 34 (Le gradient en coordonnées polaires):
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction différentiable. La fonction F définie par $F(r,\varphi) = f(r\cos\varphi, r\sin\varphi)$ pour $r > 0$ et $\varphi \in \mathbb{R}$. Cette fonction s'écrit également $f \circ M$, avec la fonction M définie à l'exercice 9.

$$(r,\varphi) \xmapsto{M} M(r,\varphi) = (x,y) \xmapsto{f} f(x,y)(r,\varphi) \xmapsto{F=f\circ M} F(r,\varphi) \ .$$

La fonction F est différentiable car f et M le sont.[2] Ainsi,

$$\left.\begin{aligned} \frac{\partial F}{\partial r} &= \frac{\partial x}{\partial r}\cdot\frac{\partial f}{\partial x} + \frac{\partial y}{\partial r}\cdot\frac{\partial f}{\partial y} && = \cos\varphi\cdot\frac{\partial f}{\partial x} + \sin\varphi\cdot\frac{\partial f}{\partial y} \\ \frac{\partial F}{\partial \varphi} &= \frac{\partial x}{\partial \varphi}\cdot\frac{\partial f}{\partial x} + \frac{\partial y}{\partial \varphi}\cdot\frac{\partial f}{\partial y} && = -r\sin\varphi\cdot\frac{\partial f}{\partial x} + r\cos\varphi\cdot\frac{\partial f}{\partial y} \end{aligned}\right\} \underset{\substack{L_1\leftarrow\cos\varphi L_1 - \sin\varphi L_2/r \\ \sin\varphi L_1 + \cos\varphi L_2/r}}{\Longleftrightarrow} \left\{\begin{aligned} \frac{\partial f}{\partial x} &= \cos\varphi\cdot\frac{\partial F}{\partial r} - \frac{\sin\varphi}{r}\cdot\frac{\partial F}{\partial \varphi} \\ \frac{\partial f}{\partial y} &= \sin\varphi\cdot\frac{\partial F}{\partial r} - \frac{\cos\varphi}{r}\cdot\frac{\partial F}{\partial \varphi} \end{aligned}\right.$$

On retrouve donc

$$\nabla f(x,y) = \frac{\partial F}{\partial r}\cdot\vec{u}_r + \frac{1}{r}\cdot\frac{\partial F}{\partial \varphi}\cdot\vec{u}_\varphi,$$

en posant $\vec{u}_r = (\cos\varphi, \sin\varphi)$ et $\vec{u}_\theta = (-\sin\varphi, \cos\varphi)$.

6 Dérivées secondes

DÉFINITION 35:
Soit $p \in \mathbb{N}^\star$, un ouvert $D \subset \mathbb{R}^p$ et une fonction $f : D \to \mathbb{R}$, $(x_1, \dots, x_p) \mapsto f(x_1, \dots, x_p)$.

1. Soit un point $\vec{a} = (a_1, \dots, a_p) \in D$. Si la i-ème dérivée partielle $\partial_i f$ existe sur D et possède une j-ième dérivée partielle en $\vec{a}$, alors le nombre réel $\partial_j(\partial_i f)(\vec{a})$ est une *dérivée partielle* en $\vec{a}$, et est noté $\partial_j\partial_i f(\vec{a})$ ou $\partial^2 f(\vec{a})/\partial x_j \partial x_i$.[3] Cette dérivée seconde s'écrit aussi parfois $\partial_{1,2} f$.
2. On dit que f est de classe $\mathscr{C}^2$ sur D si les p^2 fonctions $\partial_j\partial_i f$ sont continues sur D.

A priori, $\partial_1\partial_2 f$ est différent de $\partial_2\partial_1 f$.

EXERCICE 36:
Soit la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x,y) = xy\cdot\frac{x^2-y^2}{x^2+y^2} \text{ si } (x,y) \neq (0,0) \qquad \text{et} \qquad f(0,0) = 0.$$

Montrer que $\dfrac{\partial^2 f}{\partial x\,\partial y}(0,0) = +1$ *et* $\dfrac{\partial^2 f}{\partial y\,\partial x}(0,0) = -1$.

Si $(x,y) \neq (0,0)$, alors

$$\begin{aligned} \partial_1 f(x,y) &= \frac{\frac{\partial}{\partial x}\left(xy(x^2-y^2)\right)\cdot(x^2+y^2) - xy(x^2-y^2)\cdot\frac{\partial}{\partial x}(x^2-y^2)}{(x^2+y^2)^2} \\ &= \frac{[y(x^2-y^2)+2x^2y](x^2+y^2) - 2x^2y(x^2-y^2)}{(x^2+y^2)^2} \end{aligned}$$

2. f différentiable par hypothèse, et M est de classe $\mathscr{C}^1$, donc différentiable.
3. Cette notation vient de $\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)$.

Et, pour $h > 0$, $\big(f(h,0) - f(0,0)\big)/h = 0 \longrightarrow 0$ quand $h \longrightarrow 0$, d'où $\partial_1 f(0,0) = 0$. De plus,

$$\frac{\frac{\partial}{\partial x} f(0,h) - \frac{\partial}{\partial x} f(0,0)}{h} = \frac{-h^5/h^4}{h} = -1 \xrightarrow[h\to 0]{} -1.$$

D'où, $\dfrac{\partial f}{\partial y \partial x}(0,0) = -1$.

On procède de même pour montrer que $\dfrac{\partial f}{\partial x \partial y}(0,0) = 1$ car $f(x,y) = -f(y,x)$.

> Proposition 37 (Théorème de Schwarz):
> Soit D un ouvert de $\mathbb{R}^p$, et $f : D \to \mathbb{R}$ une fonction de classe $\mathscr{C}^2$. Alors,
>
> $$\partial_i \partial_j f = \partial_j \partial_i f, \qquad \text{pour tout } (i,j) \in [\![1,p]\!]^2.$$

> Théorème 38 (Formule de Taylor & Young à l'ordre 2, admise):
> (On se place dans $\mathbb{R}^2$, mais ce résultat s'étend à $\mathbb{R}^p$ de la même manière.) Si $f : D \to \mathbb{R}$ est de classe $\mathscr{C}^2$, alors
>
> $$\begin{aligned} f(a+h, b+k) = f(a,b) & & \text{ordre 0}\\ + h\,\partial_1 f(a,b) + k\,\partial_2 f(a,b) & & \text{ordre 1}\\ + \frac{h^2}{2}\,\partial_1\partial_1 f(a,b) + hk\,\partial_1\partial_2 f(a,b) + \frac{k^2}{2}\,\partial_2\partial_2 f(a,b) & & \text{ordre 2}\\ + \|\vec{h}\|^2\,\varepsilon(\vec{h}) & & \text{reste} \end{aligned}$$
>
> où $\vec{h} = (h,k)$, et $\varepsilon(\vec{h}) \longrightarrow \vec{0}$.

On note le développement à l'ordre 2 comme $\alpha h^2 + 2\beta hk + \gamma k^2$.[4] On note cette application $q_{(a,b)}$. On peut calculer le terme d'ordre 2 par produit matriciel :

$$\underbrace{\begin{pmatrix} h & k \end{pmatrix}}_{[\vec{h}]^\top} \cdot \underbrace{\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}}_{\mathrm{H}_{f}(a,b)} \cdot \underbrace{\begin{pmatrix} h \\ k \end{pmatrix}}_{[\vec{h}]}.$$

> Définition 39:
> Soit f une fonction admettant des dérivées partielles secondes en $\vec{a}$. La *hessienne* de f en $\vec{a}$ est la matrice
>
> $$\mathrm{H}_f(\vec{a}) = \big(\partial_i \partial_j f\big)_{i,j} = \begin{pmatrix} \dfrac{\partial^2 f}{\partial {x_1}^2}(\vec{a}) & \dfrac{\partial^2 f}{\partial x_1\,\partial x_2}(\vec{a}) & \dots & \dfrac{\partial^2 f}{\partial x_1\,\partial x_p}(\vec{a}) \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(\vec{a}) & \dfrac{\partial^2 f}{\partial {x_2}^2}(\vec{a}) & \dots & \dfrac{\partial^2 f}{\partial x_2\,\partial x_p}(\vec{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_p\,\partial x_1}(\vec{a}) & \dfrac{\partial^2 f}{\partial x_p\,\partial x_2}(\vec{a}) & \dots & \dfrac{\partial^2 f}{\partial {x_p}^2}(\vec{a}) \end{pmatrix} \in \mathscr{M}_{pp}(\mathbb{R}).$$

Si la fonction f est de classe $\mathscr{C}^2$, alors, par le théorème de Schwarz, la hessienne de f est une matrice symétrique.

7 Optimisation

4. C'est une forme quadratique.

Définition 40:
Soit $\vec{a}$ un vecteur d'une partie $D \subset \mathbb{R}^p$, où $p \in \mathbb{N}^\star$. Soit $f : D \longrightarrow \mathbb{R}$ une fonction scalaire. On dit que

1. la fonction f possède un *minimum global* sur D en $\vec{a}$ si $\quad \forall \vec{x} \in D, \quad f(\vec{x}) \geqslant f(\vec{a})$,
2. la fonction f possède un *minimum global* en $\vec{a}$ s'il existe $\varepsilon > 0$ tel que pour tout $\vec{x} \in D \cap \bar{B}(\vec{a}, \varepsilon)$, $f(\vec{x}) \geqslant f(\vec{a})$.

On définit de même un *maximum local* en $\vec{a}$ et un *maximum global* sur D en $\vec{a}$. On dit que f possède un *extremum* (*local* en $\vec{a}$, *global* sur D) si f possède un maximum ou un minimum (local en $\vec{a}$, global en D). Un extremum global est *a fortiori* local.

Théorème 41:
Soit $f : E \longrightarrow \mathbb{R}$. Si D est une partie fermée et bornée de E et si f est une fonction continue sur D, alors f possède un maximum et un minimum globaux. Autrement dit, toute fonction réelle continue sur un fermé borné est bornée et atteint ses bornes.

La preuve sera faite dans l'annexe B.

Proposition – Définition 42:
Soient $D \subset \mathbb{R}^p$ et $\vec{a}$ un point intérieur de D. Soit $f : D \to \mathbb{R}$ une fonction scalaire différentiable en $\vec{a}$.

1. On dit que $\vec{a}$ est un *point critique* de f si $\nabla f(\vec{a}) = \vec{0}$ (*i.e.* si $\mathrm{d}f(\vec{a})$ est l'application nulle).
2.
$$\begin{tikzcd} \begin{array}{c} f \text{ possède un extremum local} \\ \text{en un \underline{point intérieur} } \vec{a} \end{array} \arrow[r, Rightarrow, bend right] & \begin{array}{c} \vec{a} \text{ est un point critique de } f \\ \nabla f(\vec{a}) = \vec{0} \end{array} \arrow[l, Rightarrow, bend right, "\times"'] \end{tikzcd}.$$

Démonstration:
Généralisation d'un théorème vu précédemment, la preuve est dans la section 14.4 du cours de *MP2I*.

Proposition 43:
Soit $\vec{a} \in D$ un point d'un ouvert de $\mathbb{R}^p$, et soit $f : D \longrightarrow \mathbb{R}^2$ une fonction scalaire de classe $\mathscr{C}^2$. Soient $(\lambda_1, \ldots, \lambda_p)$ les valeurs propres de la hessienne de f en $\vec{a}$.

1. **Condition <u>nécessaire</u> de minimum local en un point <u>intérieur</u>.** Si f possède un minimum local en $\vec{a}$, alors
$$\nabla f(\vec{a}) = \vec{0} \quad \text{et} \quad \forall i \in [\![1, p]\!],\ \lambda_i \geqslant 0.$$
Autrement dit, $\nabla f(\vec{a}) = \vec{0}$ et $\mathrm{H}_f(\vec{a}) \in \mathcal{S}_p^+$.
2. **Condition <u>suffisante</u> de minimum local en un point <u>intérieur</u>.** Si $\nabla f(\vec{a}) = \vec{0}$, et $\forall i \in [\![1, p]\!]$, $\lambda_i > 0$, autrement dit si $\nabla f(\vec{a}) = \vec{0}$ et $\mathrm{H}_f(\vec{a}) \in \mathcal{S}_p^{++}$, alors f possède un minimum local en $\vec{a}$.

Démonstration: — Si f possède un minimum local, alors $\vec{a}$ est un point critique, donc $\nabla f(\vec{a}) = \vec{0}$. D'où, $f(\vec{a}+\vec{h}) = f(\vec{a}) + \frac{1}{2} q_{\vec{a}}(\vec{h}) + \|h\|^2 \varepsilon(\vec{h})$. Comme $\vec{a}$ est un minimum local, alors $\frac{1}{2} q_{\vec{a}}(\vec{h}) + \|\vec{h}\|^2 \varepsilon(\vec{h}) \geqslant 0$, pour tout $\vec{h}$ suffisamment petit (le minimum est local). Or, $\frac{1}{2} q_{\vec{a}}(\vec{h}) + \|\vec{h}\|^2 \varepsilon(\vec{h}) = \frac{1}{2} \left[\, \vec{h} \,\right]^\top \cdot \mathrm{H}_f(\vec{a}) \cdot \left[\, \vec{h} \,\right] + \|\vec{h}\|^2 \varepsilon(\vec{h})$. Et, la matrice $\mathrm{H}_f(\vec{a})$ est symétrique et à coefficients réels d'où, d'après le théorème spectral, elle est diagonalisable dans une base $\mathscr{B}$ orthonormée formée de vecteurs propres. D'où, $\frac{1}{2} q_{\vec{a}}(\vec{h}) + \|\vec{h}\|^2 \varepsilon(\vec{h}) = \frac{1}{2} \left[\vec{h}\right]_{\mathscr{B}}^\top \cdot \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \cdot \left[\vec{h}\right]_{\mathscr{B}} + \|\vec{h}\|^2 \varepsilon(\vec{h}) = \frac{1}{2}(\lambda_1 h_1^2 + \cdots + \lambda_p h_p^2) + (h_1^2 + \cdots + h_p^2)\, \varepsilon(\vec{h}) \geqslant 0$, pour tout vecteur $\vec{h}$ suffisamment petit. Montrons que $\forall i$, $\lambda_i \geqslant 0$. Par l'absurde, supposons qu'il existe $i \in [\![1, p]\!]$ tel que $\lambda_i \leqslant 0$. On considère le vecteur $\vec{h} = (0, \ldots, 0, h, 0, \ldots, 0)$ qui est non nul à la i-ème coordonnée. Alors,
$$\frac{1}{2} \lambda_i\, h^2 + h^2 \varepsilon(\vec{h}) = h^2 \left(\frac{1}{2} \lambda_i + \varepsilon(\vec{h})\right) \geqslant 0.$$

Or, $\varepsilon(\vec{h}) \to 0$ lorsque $\vec{h} \to 0$. D'où, $\frac{1}{2}\lambda_i + \varepsilon(\vec{h}) < 0$ pour $\vec{h}$ assez petit. Absurde.

Ceci reste vrai pour un maximum en replaçant f par $-f$.

MÉTHODE 44 (Démontrer les extrema locaux en dimension 2):

1. On détermine les points critique de la fonction f.
2. Pour chaque point critique $\vec{a}$, on calcule le déterminant et la trace de la hessienne $\mathrm{H}_f(\vec{a})$.
   (a) Si le déterminant est strictement positif, alors il y a un extremum local en $\vec{a}$:
   - si la trace est strictement positive, alors f admet un minimum local en $\vec{a}$,
   - si la trace est strictement négative, alors f admet un maximum local en $\vec{a}$.

   (b) Si le déterminant est strictement négatif, alors f n'admet pas d'extremum local en $\vec{a}$.

   (c) Si le déterminant est nul, alors la proposition ne permet pas de conclure.

EXERCICE 45:
Étudier les extrema de la fonction

$$\begin{aligned} f : \bar{B}(\vec{0}, 1) &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x^2 - y^2 \end{aligned}$$

définie sur la boule $\bar{B}(\vec{0}, 1) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leqslant 1\}$ *fermée de centre* $\vec{0}$ *et de rayon* 1.

La fonction f est continue sur le fermé borné $\bar{B}(\vec{0}, 1)$, donc elle possède un maximum et un minimum globaux.

- À l'intérieur de la boule $\bar{B}(\vec{0}, 1)$, on procède par analyse–synthèse.
  - **Analyse.** S'il existe un extremum en un point $(a, b) \in \mathring{B}(\vec{0}, 1)$, alors $\nabla f(a, b) = (2a, -2b) = (0, 0)$ d'où $(a, b) = \vec{0}$.
  - **Synthèse.** En $(a, b) = \vec{0}$, alors $f(a, b) = f(0, 0) = 0$. Or, $f(0, k) = -k^2 < 0$, pour tout $k > 0$. Et, $f(h, 0) = h^2 > 0$, pour tout $h > 0$. Le point $\vec{0}$ n'est, ni un minimum local, ni un maximum local,pour la fonction f.
  - **Suite de l'analyse.** (deuxième méthode) On calcule la hessienne de f en $\vec{0}$. On trouve

    $$\mathrm{H}_f(\vec{0}) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

    Les deux valeurs propres ont des signes différents, il n'y a donc ni maximum, ni minium en $\vec{0}$.
- Et au bord, il y a un minimum global et un maximum global par continuité de f. Sur le bord, $x^2 + y^2 = 1$, donc $f(x, y) = x^2 - y^2 = x^2 - (1 - y^2) = 2x^2 - 1$, pour tout $x \in [-1, 1]$. Soit $g : x \mapsto x^2 - 1$ définie sur $[-1, 1]$.
  - Sur l'ouvert $]-1, 1[$, s'il y a un extremum local en x, alors $g'(x) = 4x = 0$ et donc $x = 0$. Et alors, $f(x, y) = g(x) = -1$, et c'est un minimum global (au vu de la fonction f) atteint aux points $(0, 1)$ et $(0, -1)$.
  - Et, au bord de $[-1, 1]$, alors : si $x = 1$ ou si $x = -1$, alors $y = 0$ et donc $f(x, y) = 1$, et c'est un maximum global.

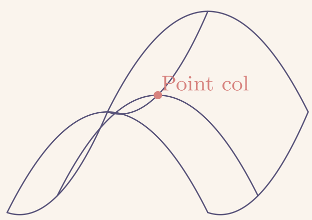


FIGURE 8 – Représentation graphique de la fonction f

PROPOSITION 46 (Théorème des extrema liés : c'est une condition nécessaire):
Soient f et g deux foncions de classe $\mathscr{C}^1$ de E vers $\mathbb{R}$. Si f est constante sur une partie $C \subset E$, et si la restriction de g à C admet un extremum local en $\vec{x} \in C$, et si $\vec{x}$ n'est pas un point critique de f, alors

$$\exists \lambda \in \mathbb{R}, \qquad \nabla g(\vec{x}) = \lambda \, \nabla f(\vec{x}).$$

DÉMONSTRATION:
La fonction g est extrémale sous la contrainte $f(x_1, \dots, x_n) = \mathrm{C}^{\underline{\mathrm{te}}}$. Le gradient est orthogonal aux lignes de niveaux de la fonction. En particulier, $\nabla f(\vec{a})$ est orthogonal à C en tout point $\vec{a} \in C$. On considère un point $M(t) \in C$. À un instant t_0, on aura $M(t_0) = \vec{a}$. Si $g\big|_C$ possède un extremum local en $\vec{a}$, alors, à la date t_0, $\frac{\mathrm{d}}{\mathrm{d}t} g \circ M(t_0) = 0$. Or, d'après la règle de la chaîne, $\frac{\mathrm{d}}{\mathrm{d}t} g \circ M(t_0) = \langle \nabla g(M(t_0)) \mid M'(t_0) \rangle = 0$.

EXEMPLE 47:
Soit C une courbe de niveau d'une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathscr{C}^1$. Soit A un point du plan. Si la distance AM du point A à un point non critique $M \in C$ possède un extremum local, alors le vecteur $\overrightarrow{AM}$ est orthogonal à la courbe C.

En effet, s'il y a un extremum local de g en (x, y) sous la contrainte $f(x, y) = \mathrm{C}^{\underline{\mathrm{te}}}$, alors $\nabla g(x, y) \mathbin{/\!/} \nabla f(a, b)$.

EXERCICE 48:
Déterminer les points du cercle d'équation $x^2 + y^2 = 5$ qui rendent extrémale la valeur $2x + y$.

Soient f et g deux fonctions de classe $\mathscr{C}^1$ définies par $f(x, y) = x^2 + y^2$ et $g(x, y) = 2x + y$. On applique le théorème des extrema liés : s'il existe un extremum en (a, b), alors $\nabla g(a, b) \mathbin{/\!/} \nabla f(a, b)$, *i.e.* $(2, 1) \mathbin{/\!/} (2a, 2b)$. D'où, $\begin{vmatrix} 2 & 2a \\ 1 & 2b \end{vmatrix} = 0$, d'où $2b - a = 0$. De plus, $a^2 + b^2 = 5$.