Chapitre 11

FONCTIONS VECTORIELLES DE PLUSIEURS VARIABLES

Dans tout ce paragraphe, soient $m, n \in \mathbb{N}^*$ et X une partie ouverte de $\mathbb{R}^n$.
Sauf mention expresse du contraire,
nous munirons $\mathbb{R}^n$ et $\mathbb{R}^m$ de la norme euclidienne
$|\cdot| := |\cdot|_2$.

Version du 30 novembre 2005

11.1 Courbes paramétrées

Une application d'un intervalle de $\mathbb{R}$ dans $\mathbb{K}^n$ peut être interprétée de manière plus géométrique en accordant plus d'importance à son image qu'à son graphe. Mais on peut aussi étudier certaines parties de $\mathbb{K}^n$ en les représentant comme l'image d'une telle application.

DEFINITION 1 Une partie C de $\mathbb{K}^n$ est dite une *courbe* s'il existe un intervalle J de $\mathbb{R}$ tel que $J^\circ \neq \emptyset$ et une fonction continue $\gamma : J \longrightarrow \mathbb{K}^n$ telle que $C = \gamma(J)$. On dit que γ est une *paramétrage* de C , ou encore que γ est une *courbe paramétrée* de $\mathbb{K}^n$.

LEMME *Si on écrit*

$$\gamma(t) = (\gamma_1(t), \ldots, \gamma_n(t)) \quad \textit{pour tout } t \in J \ ,$$

alors γ est continue si, et seulement si, chaque fonction composante

$$\gamma_j : J \longrightarrow \mathbb{K} \quad \textit{pour } j = 1, \ldots, n$$

est continue.

Cela découle de la proposition 7.3.(iii). □

REMARQUE 1 On interprète souvent une courbe paramétrée comme la représentation d'un point se déplaçant dans $\mathbb{K}^n$ au cours du temps. C'est l'aspect cinématique de la mécanique.

EXEMPLE 1 Le cercle dans $\mathbb{R}^2$ de centre $(u, v) \in \mathbb{R}^2$ et de rayon $r \in \mathbb{R}_+$:

$$\left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; |x - u|^2 + |y - v|^2 = r^2 \right\} \ ,$$

est paramétré par

$$[0, 1] \longrightarrow \mathbb{R}^2 : t \longmapsto (u + r \cdot \cos(2\pi \cdot t), v + r \cdot \sin(2\pi \cdot t)) \ .$$

Le cercle dans $\mathbb{C}$ de centre $w \in \mathbb{C}$ et de rayon $r \in \mathbb{R}_+$:

$$\{ z \in \mathbb{C} \mid |z - w| = r \} \ ,$$

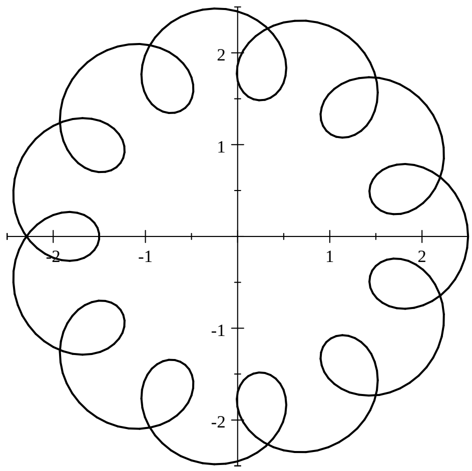
est paramétré par

$$[0, 1] \longrightarrow \mathbb{C} : t \longmapsto w + r \cdot e^{2\pi i \cdot t} \ .$$

EXEMPLE 2 La courbe décrite par la lune par rapport au soleil placé en $(0, 0)$ peut être paramétrée par

$$t \longmapsto R \cdot e^{2\pi i \cdot \Omega t} + r \cdot e^{2\pi i \cdot \omega t} \ ,$$

les trajectoires de la terre autour du soleil et de la lune autour de la terre étant considérées comme circulaires et contenues dans un plan.



$$2 \cdot e^{2\pi i \cdot \mathrm{id}} + \frac{1}{2} \cdot e^{24\pi i \cdot \mathrm{id}}$$

EXEMPLE 3 La droite dans $\mathbb{R}^n$ passant par $u \in \mathbb{R}^n$ et de direction $v \in \mathbb{R}^n \setminus \{0\}$ peut être paramétrée par

$$t \longmapsto u + t \cdot v : \mathbb{R} \longrightarrow \mathbb{R}^n \ .$$

EXEMPLE 4 Si $\gamma : J \longrightarrow \mathbb{K}^n$ est une courbe paramétrée de $\mathbb{K}^n$, alors

$$\delta : t \longmapsto (t, \gamma(t)) : J \longrightarrow \mathbb{R} \times \mathbb{K}^n$$

est une courbe paramétrée de $\mathbb{R} \times \mathbb{K}^n$ et $\delta(J)$ est le graphe de γ .

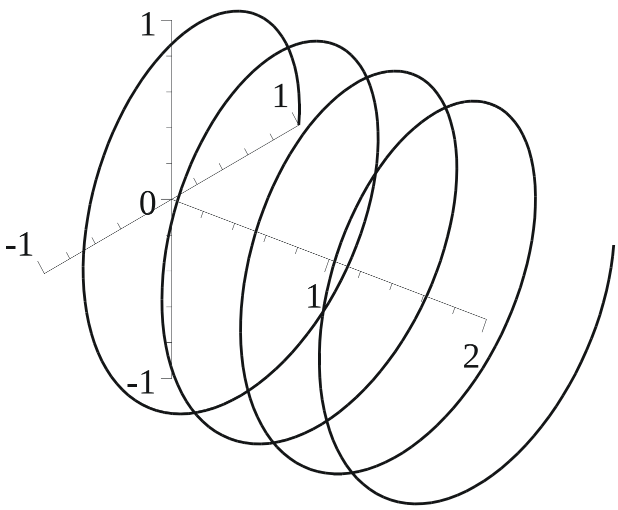
Par exemple si

$$\gamma : t \longmapsto (\cos(4\pi \cdot t), \sin(4\pi \cdot t)) : \mathbb{R} \longrightarrow \mathbb{R}^2$$

est un paramétrage du cercle de centre 0 et de rayon 1 , alors

$$\delta : t \longmapsto (t, \cos(4\pi \cdot t), \sin(4\pi \cdot t)) : \mathbb{R} \longrightarrow \mathbb{R} \times \mathbb{R}^2$$

est une paramétrage d'une vis autour du premier axe dans $\mathbb{R}^3$.



REMARQUE 2 Si $\gamma : J \longrightarrow \mathbb{K}^n$ est une courbe paramétrée, on décrit souvent la courbe $\gamma(J)$ sous la forme

$$\gamma(J) = \{ z \in \mathbb{K}^n \mid F(z) = 0 \} = \{ (z_1, \ldots, z_n) \in \mathbb{K}^n \mid F(z_1, \ldots, z_n) = 0 \} \ ,$$

où $F : \mathbb{K}^n \longrightarrow \mathbb{K}$ est une fonction continue. On dit que $F(z) = 0$ est une *équation de la courbe* $\gamma(J)$ ou que $\gamma(J)$ est d'équation $F(z) = 0$.

Par exemple, le cercle dans $\mathbb{R}^2$ de centre $(u, v) \in \mathbb{R}^2$ et de rayon $r \in \mathbb{R}_+$ est d'équation

$$(x-u)^2 + (y-v)^2 - r^2 = 0 \ .$$

Celui dans $\mathbb{C}$ de centre $w \in \mathbb{C}$ et de rayon $r \in \mathbb{R}_+$ est d'équation

$$|z - w| = r \ .$$

En écrivant une équation il faut toujours bien préciser ce que sont les variables.

DEFINITION 2 On dit que la courbe paramétrée $\gamma : J \longrightarrow \mathbb{K}^n$ est *dérivable* en $t \in J$ si chaque composante γ_j est dérivable en t et on dit que

$$\gamma'(t) := (\gamma_1'(t), \ldots, \gamma_n'(t))$$

est la *dérivée* de γ en t ou le *vecteur tangent* à la courbe γ en t . On dit que γ est *continûment dérivable* dans J , si γ est dérivable dans J et $\gamma' : J \longrightarrow \mathbb{K}^n$ est continue, i.e. si chaque composante γ_j est continûment dérivable dans J .

Si γ est dérivable en $t \in J$, on a

$$\gamma'(t) = \left(\lim_{t \neq s \to t} \frac{\gamma_1(s) - \gamma_1(t)}{s - t}, \ldots, \lim_{t \neq s \to t} \frac{\gamma_n(s) - \gamma_n(t)}{s - t} \right) =$$

$$= \lim_{t \neq s \to t} \left(\frac{\gamma_1(s) - \gamma_1(t)}{s - t}, \ldots, \frac{\gamma_n(s) - \gamma_n(t)}{s - t} \right) = \lim_{t \neq s \to t} \frac{\gamma(s) - \gamma(t)}{s - t} \ ,$$

ce qui montre que $\gamma'(t)$ est limite de vecteurs sécants.

Dans l'interprétation cinématique d'une courbe, $\gamma'(t)$ est le *vecteur vitesse* au temps t .

EXEMPLE 5 Une courbe paramétrée n'est pas nécessairement injective. En un même point de l'espace on peut avoir deux vecteurs tangents.

Soit

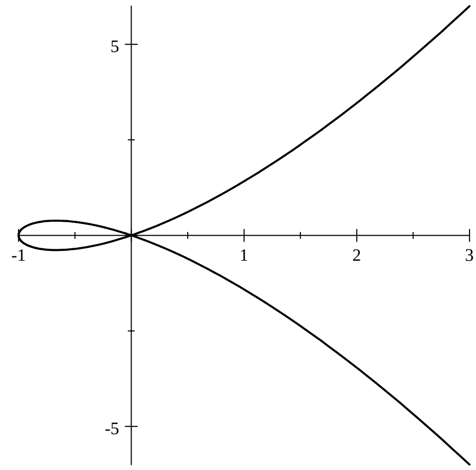
$$\gamma : t \longmapsto \left(t^2 - 1, t^3 - t\right) : \mathbb{R} \longrightarrow \mathbb{R}^2 \ .$$

On a

$$\gamma(\mathbb{R}) = \left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; y^2 = x^2 + x^3 \right\} \ ,$$

en particulier

$$\gamma(-1) = \gamma(1) = (0, 0) \quad , \quad \gamma'(-1) = (-2, 2) \quad \text{et} \quad \gamma'(1) = (2, 2) \ .$$



EXEMPLE 6 Parabole de Neil Une courbe paramétrée dérivable peut avoir une pointe. On dit aussi un *point de rebroussement* . Soit

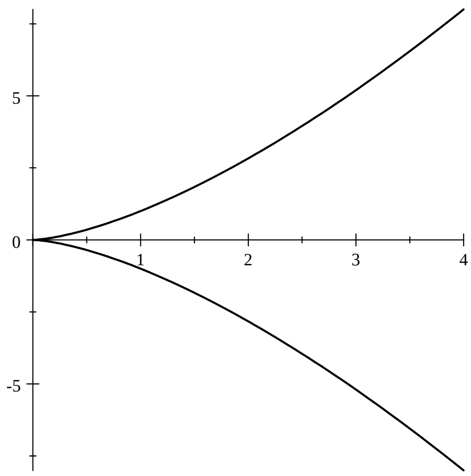
$$\gamma : t \longmapsto \left(t^2, t^3\right) : \mathbb{R} \longrightarrow \mathbb{R}^2 \ .$$

On a

$$\gamma\left(\mathbb{R}\right) = \left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; y^2 = x^3 \right\}$$

et $\gamma\left(0\right) = (0, 0)$, $\gamma'\left(0\right) = (0, 0)$. Un autre paramétrage de la parabole de Neil est donnée par

$$t \longmapsto \begin{cases} \left((t+r)^2, (t+r)^3\right) & t < -r \\ 0 & \text{si} \quad t \in [-r, r] \\ \left((t-r)^2, (t-r)^3\right) & r < t \end{cases} \ .$$



EXERCICE On considère un point matériel de masse m suspendu en 0 à un élastique de longueur L et dont la constante est a . On dit que c'est le pendule élastique. Sa trajectoire γ est gouvernée par l'équation différentielle

$$m \cdot \gamma'' = m \cdot \vec{g} - a \cdot \left(\gamma - L \cdot \frac{\gamma}{|\gamma|}\right) \quad , \text{où} \quad \vec{g} = \begin{pmatrix} 0 \\ 0 \\ -g \end{pmatrix} \ .$$

Nous supposerons que la longueur L est négligeable.

(a) Etant donné $\xi, v \in \mathbb{R}^3$, déterminer la trajectoire du pendule tel que $\gamma(0) = \xi$ et $\gamma'(0) = v$.

(b) Déterminer la courbe décrite dans $\mathbb{R}^3$ par le pendule lorsque $v = 0$.

(c) Déterminer la courbe décrite dans $\mathbb{R}^3$ par le pendule lorsque $\xi_2 = v_2 = 0$.

Utiliser l'exercice 9.8.4.

11.2 Longueur d'une courbe

DEFINITION 1 Si $\gamma : J \longrightarrow \mathbb{K}^n$ est une courbe paramétrée continue, pour tout $a, b \in J$, on définit l' *intégrale* (de Riemann) de γ entre a et b par

$$\int_a^b \gamma := \left(\int_a^b \gamma_j\right)_{j=1,\ldots,n} \in \mathbb{K}^n \ .$$

Remarquons, si γ est continûment dérivable, que

$$\int_a^b \gamma' = \gamma(b) - \gamma(a) \ .$$

PROPOSITION *Soient $\gamma : J \longrightarrow \mathbb{K}^n$ une courbe paramétrée continûment dérivable et $[a, b] \subset J$.*

(i) On a

$$\left|\int_a^b \gamma\right| \leqslant \int_a^b |\gamma| \ .$$

(ii) ***Seconde inégalité de la moyenne*** *Pour tout $t \in [a, b]$, on a*

$$|\gamma(b) - \gamma(a) - (b-a) \cdot \gamma'(t)| \leqslant (b-a) \cdot \sup_{s \in [a,b]} |\gamma'(s) - \gamma'(t)| \ .$$

Démonstration de (i) Posons $z := \int_a^b \gamma \in \mathbb{K}^n$. Grâce à la proposition 9.7, il vient alors

$$|z|^2 = (z| z) = \left|\sum_{j=1}^n \overline{z_j} \cdot \int_a^b \gamma_j\right| \leqslant \int_a^b \left|\sum_{j=1}^n \overline{z_j} \cdot \gamma_j(t)\right| dt =$$

$$= \int_a^b |(z| \gamma(t))| \, dt \leqslant \int_a^b |z| \cdot |\gamma(t)| \, dt = |z| \cdot \int_a^b |\gamma|$$

en ayant utilisé l'inégalité de Cauchy-Schwarz 10.2.

Démonstration de (ii) Considérons la fonction

$$f : [a, b] \longrightarrow \mathbb{K}^n : s \longmapsto \gamma(s) - s \cdot \gamma'(t) \ .$$

Elle est continûment dérivable et on a $f'(s) = \gamma'(s) - \gamma'(t)$ .Il vient alors

$$|\gamma(b) - \gamma(a) - (b-a) \cdot \gamma'(t)| = |f(b) - f(a)| = \left|\int_a^b f'\right| = \left|\int_a^b (\gamma'(s) - \gamma'(t)) \, ds\right| \leqslant$$

$$\leqslant \int_a^b |\gamma'(s) - \gamma'(t)| \, ds \leqslant (b-a) \cdot \sup_{s \in [a,b]} |\gamma'(s) - \gamma'(t)| \ .$$

□

REMARQUE Rappelons que l'espace métrique $(\mathbb{C}^n, |\cdot|)$ peut être identifié à $(\mathbb{R}^{2n}, |\cdot|)$ grâce à

$$z = (z_1, \ldots, z_n) \longmapsto (\operatorname{Re} z_1, \operatorname{Im} z_1, \ldots, \operatorname{Re} z_n, \operatorname{Im} z_n) = (x_1, x_2, \ldots, x_{2n-1}, x_{2n}) = x \ .$$

Pour tout $z, w \in \mathbb{C}^n$, identifiés à $x, y \in \mathbb{R}^{2n}$, on a

$$|z - w|^2 = \sum_{j=1}^{n} |z_j - w_j|^2 = \sum_{j=1}^{n} \left(|\operatorname{Re} z_j - \operatorname{Re} w_j|^2 + |\operatorname{Im} z_j - \operatorname{Im} w_j|^2 \right) =$$

$$= \sum_{j=1}^{2n} |x_j - y_j|^2 = |x - y|^2 \ .$$

DEFINITION 2 Soit $[a, b]$ un intervalle de $\mathbb{R}$ et $(t_k)_{k=0,\ldots,m}$ une subdivision de cet intervalle. On dit que

$$\max\nolimits_{k=0,\ldots,m-1} |t_{k+1} - t_k|$$

est la *finesse* de cette subdivision.

Soient $\gamma : J \longrightarrow \mathbb{K}^n$ une courbe paramétrée, $[a, b] \subset J$ et $L \in \mathbb{R}$. On dit que γ est *rectifiable* et de *longueur* L sur $[a, b]$ si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que pour toute subdivision $(t_k)_{k=0,\ldots,m}$ de $[a, b]$ de finesse $\leqslant \delta$, on ait

$$\left| \sum_{k=0}^{m-1} |\gamma(t_{k+1}) - \gamma(t_k)| - L \right| \leqslant \varepsilon \ .$$

THEOREME *Soit $\gamma : J \longrightarrow \mathbb{K}^n$ une courbe paramétrée continûment dérivable. Alors γ est rectifiable sur tout intervalle $[a, b] \subset J$ et sa longueur est*

$$\int_a^b |\gamma'| = \int_a^b \left(|\gamma_1'(t)|^2 + \ldots + |\gamma_n'(t)|^2 \right)^{\frac{1}{2}} dt \ .$$

Si $(t_k)_{k=0,\ldots,m}$ est une subdivision de $[a, b]$, on a

$$\left| \sum_{k=0}^{m-1} |\gamma(t_{k+1}) - \gamma(t_k)| - \int_a^b |\gamma'(t)| \ dt \right| =$$

$$= \left| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} |\gamma(t_{k+1}) - \gamma(t_k)| \ \frac{dt}{t_{k+1} - t_k} - \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} |\gamma'(t)| \ dt \right| =$$

$$= \left| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \Big(|\gamma(t_{k+1}) - \gamma(t_k)| - (t_{k+1} - t_k) \cdot |\gamma'(t)| \Big) \ \frac{dt}{t_{k+1} - t_k} \right| \leqslant$$

$$\leqslant \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \Big| |\gamma(t_{k+1}) - \gamma(t_k)| - (t_{k+1} - t_k) \cdot |\gamma'(t)| \Big| \ \frac{dt}{t_{k+1} - t_k} \leqslant$$

$$\leqslant \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} |\gamma(t_{k+1}) - \gamma(t_k) - (t_{k+1} - t_k) \cdot \gamma'(t)| \ \frac{dt}{t_{k+1} - t_k} \leqslant$$

$$\leqslant \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \sup_{s\in[t_k,t_{k+1}]} |\gamma'(s) - \gamma'(t)| \, dt \ ,$$

en ayant utilisé l'inégalité de la proposition 10.1.

Mais comme γ' est continue sur $[a,b]$, elle est uniformément continue (théorème 10.21). Etant donné $\varepsilon > 0$, il existe donc un $\delta > 0$ tel que l'on ait $|\gamma'(s) - \gamma'(t)| \leqslant \frac{\varepsilon}{b-a}$ pour tout $s,t \in [a,b]$ tels que $|s-t| \leqslant \delta$. Si la subdivision $(t_k)_{k=0,\ldots,m}$ a une finesse $\leqslant \delta$, on obtient finalement

$$\left| \sum_{k=0}^{m-1} |\gamma(t_{k+1}) - \gamma(t_k)| - \int_a^b |\gamma'(t)| \, dt \right| \leqslant \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \frac{\varepsilon}{b-a} dt = \varepsilon \ .$$

□

EXEMPLE 1 Soit $\gamma : [0,x] \longrightarrow \mathbb{R}^2 : t \longmapsto (\cos t, \sin t)$. C'est une paramétrage d'une portion d'arc sur le cercle unité. On a

$$\gamma'(t) = (-\sin t, \cos t) \ ,$$

donc

$$|\gamma'(t)| = \left(\sin^2 t + \cos^2 t\right)^{\frac{1}{2}} = 1 \ .$$

La longueur de cette courbe paramétrée est par suite

$$\int_0^x dt = x$$

(cf. théorème 7.6).

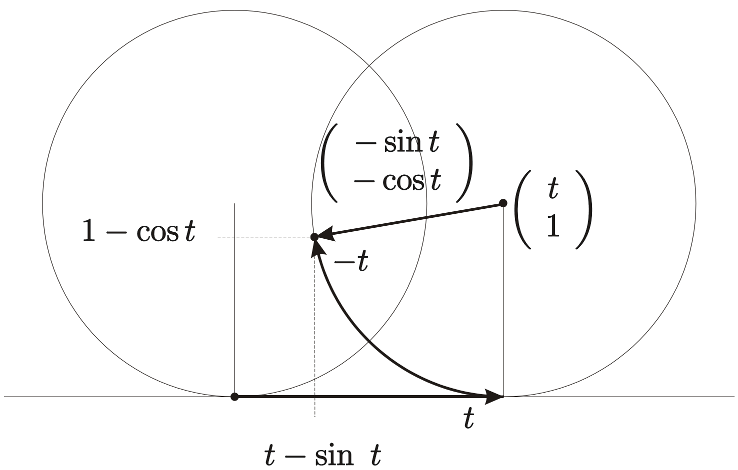
EXEMPLE 2 Cycloïde C'est une courbe de $\mathbb{R}^2$ paramétrée par

$$\gamma : \mathbb{R} \longrightarrow \mathbb{R}^2 : t \longmapsto (t - \sin t, 1 - \cos t) \ .$$

Elle s'obtient en observant un point sur un cercle de rayon 1 roulant sans glisser sur une droite. En partant avec le cercle de rayon 1 et de centre $(0,1)$, lorsque le centre se trouve en $(t,1)$, le point $(0,0)$ se trouve en

$$\begin{pmatrix} t \\ 1 \end{pmatrix} + \begin{pmatrix} \cos(-t) & -\sin(-t) \\ \sin(-t) & \cos(-t) \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} t - \sin t \\ 1 - \cos t \end{pmatrix} \ ,$$

puisque le cercle a effectué une rotation d'angle $-t$.



On a

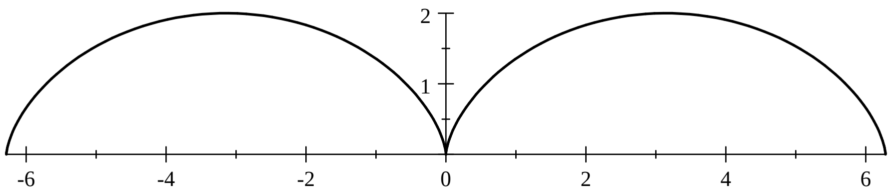
$$\gamma'(t) = (1 - \cos t, \sin t) \ ,$$

donc

$$|\gamma'(t)|^2 = (1 - \cos t)^2 + \sin^2 t = 2 \cdot (1 - \cos t) = 4 \cdot \sin^2 \frac{t}{2} \ .$$

La longueur de cette courbe paramétrée sur $[0, 2\pi]$ est donc

$$\int_0^{2\pi} 2 \cdot \sin \frac{t}{2} \, dt = \left[-4 \cdot \cos \frac{t}{2} \right]_0^{2\pi} = 8 \ .$$



EXERCICE 1 Soient $\gamma : J \longrightarrow \mathbb{K}^n$ courbe paramétrée, $a, b \in J$ et $z \in \mathbb{K}^n$. Montrer que

$$\left(z \middle| \int_a^b \gamma(t) \, dt \right) = \int_a^b (z | \gamma(t)) \, dt \ .$$

EXERCICE 2 Etant donné $c \in \mathbb{R}_+^*$, la courbe paramétrée

$$\gamma : \mathbb{R} \longrightarrow \mathbb{R}^2 : t \longmapsto e^{c \cdot t} \cdot (\cos t, \sin t)$$

s'appelle la *spirale logarithmique* .

(a) Esquisser la courbe $\gamma([-2\pi, 2\pi])$ en prenant $c = \frac{1}{2\pi}$ et calculer γ' dans le cas général.

(b) Pour tout intervalle $[a, b]$ de $\mathbb{R}$, montrer que $\gamma_{|[a,b]}$ est rectifiable et calculer sa longueur $L_{[a,b]}$.

(c) Est-ce que la limite $\lim_{a \to -\infty} L_{a,0}$ existe ?

(d) Montrer que, pour tout $r \in \mathbb{R}_+^*$, la courbe $\gamma(\mathbb{R})$ coupe le cercle d'équation $x^2 + y^2 = r^2$ en un et un seul point. Calculer l'angle d'intersection.

EXERCICE 3 Etant donné $n \in \mathbb{N}^*$ on dit que

$$\mathbb{S}^n := \left\{ x \in \mathbb{R}^{n+1} \ \middle| \ |x|_2 = 1 \right\}$$

est la *n-sphère* dans $\mathbb{R}^{n+1}$. Montrer que

(a) $\mathbb{S}^n$ est compacte.

(b) Quel que soient les points $x, y \in \mathbb{S}^n$, il existe une courbe paramétrée continue

$$\gamma : [0, 1] \longrightarrow \mathbb{S}^n$$

telle que $\gamma(0) = x$ et $\gamma(1) = y$.

On projettera une courbe adéquate de $\mathbb{R}^{n+1} \setminus \{0\}$ sur $\mathbb{S}^n$.

(c) Toute fonction continue $f : \mathbb{S}^n \longrightarrow \mathbb{R}$ possède un point antipodal x , i.e. tel que $f(x) = f(-x)$ [3].

(d) Si $\gamma : J \longrightarrow \mathbb{S}^n$ est dérivable, alors $\gamma'(t)$ est orthogonal à $\gamma(t)$ pour tout $t \in J$ et on a

$$|\gamma''|_2 \geqslant |\gamma'|_2^2$$

partout où γ est deux fois dérivable.

EXERCICE 4 Soit $a \in \,]0, \infty[$. Montrer que la longueur de la courbe paramétrée

$$\gamma : \mathbb{R} \longrightarrow \mathbb{R}^2 : t \longmapsto \left(2 \cdot t, t^2\right)$$

sur $[0, a]$ est

$$a \cdot \sqrt{1 + a^2} + \operatorname{arcsinh} a \ .$$

EXERCICE 5 Soient $a, b \in \mathbb{R}$ tels que $a < b$ et $c, r \in \mathbb{R}_+^*$. Calculer la longueur de la courbe paramétrée

$$\gamma : \mathbb{R} \longrightarrow \mathbb{R}^3 : t \longmapsto (r \cdot \cos t, r \cdot \sin t, c \cdot t)$$

sur $[a, b]$.

[3] Ceci peut par exemple s'interpréter de la manière suivante. Admettons que la température à la surface de la terre soit une fonction continue de la position. Alors il existe deux points aux antipodes l'un de l'autre ayant la même température.

11.3 Dérivées partielles

Soit

$$f : X \longrightarrow \mathbb{K} : x = (x_1, \ldots, x_n) \longmapsto f(x_1, \ldots, x_n) = f(x)$$

une fonction. Son graphe est une partie de $X \times \mathbb{K} \subset \mathbb{R}^n \times \mathbb{K}$, mais même dans le cas $n = 2$ et $\mathbb{K} = \mathbb{R}$ une représentation $3D$ n'est pas complète à cause des zones cachées. Une carte topographique, avec ses courbes de niveaux et ombrage, est en fait idéale. De façon générale cela revient à étudier les ensembles niveaux définis par

$$N_f(h) := \{x \in X \mid f(x) = h\} \quad \text{pour tout } h \in \mathbb{K} .$$

Nous n'étudierons pas cet aspect ici.

Pour étudier une telle fonction, on peut aussi se ramener au cas d'une fonction d'une variable en considérant une courbe paramétrée

$$\gamma : J \longrightarrow X \subset \mathbb{R}^n ,$$

et en étudiant la fonction

$$f \circ \gamma : J \longrightarrow \mathbb{K} .$$

Par exemple, étant donné $x \in X$ et $v \in \mathbb{R}^n$, la courbe paramétrée

$$\gamma : \mathbb{R} \longrightarrow \mathbb{R}^n : t \longmapsto x + t \cdot v$$

est continûment dérivable et $\gamma^{-1}(X)$ est un ouvert de $\mathbb{R}$ contenant 0 . Il existe donc un intervalle J de $\mathbb{R}$ contenant 0 tel que $\gamma(J) \subset X$. On considère en général les droites paramétrées passant par x et parallèles aux axes

$$t \longmapsto x + t \cdot e_j ,$$

où $(e_j)_{j=1,\ldots,n}$ est la base canonique de $\mathbb{R}^n$.

DEFINITION 1 Soient $f : X \longrightarrow \mathbb{K}$, $x \in X$ et $j \in \{1, \ldots, n\}$. On dit que f est en x *partiellement dérivable par rapport à la j-ième variable* si la fonction

$$t \longmapsto f(x + t \cdot e_j)$$

est dérivable en 0 . On pose

$$\partial_j f(x) := \lim_{t \to 0} \frac{f(x + t \cdot e_j) - f(x)}{t}$$

et on dit que c'est la *j-ième dérivée partielle* de f en x .

Seule la j-ième composante de $x + t \cdot e_j$ change par rapport à celles de x . Cela revient donc à considérer les variables autres que la j-ième comme des paramètres et f comme une fonction de cette variable seulement. En d'autre termes $\partial_j f(x)$ est la dérivée de la fonction

$$s \longmapsto f(x_1, \ldots, x_{j-1}, s, x_{j+1}, \ldots, x_n)$$

au point x_j . Ainsi

SCOLIE *Les règles de dérivation habituelles sont aussi valables pour les dérivées partielles.*

DEFINITION 2 On dit que la fonction f est *partiellement dérivable* en x si elle est partiellement dérivable en x par rapport à chaque variable. On dit qu'elle est *partiellement dérivable* sur X si elle est partiellement dérivable en chaque point de X . On dit que f est *continûment partiellement dérivable* (sur X) si en plus chaque dérivée partielle

$$\partial_j f : X \longrightarrow \mathbb{K} : x \longmapsto \partial_j f(x) \quad \text{pour } j \in \{1, \ldots, n\}$$

est continue.

REMARQUE 1 Attention, la condition " f est continûment partiellement dérivable" est plus forte que d'exiger que les fonctions

$$x_j \longmapsto f(x_1, \ldots, x_j, \ldots, x_n)$$

soient continûment dérivables. On dirait "partiellement continûment dérivable" !

Rappelons les notations

$$\mathrm{id} : \mathbb{R}^n \longrightarrow \mathbb{R}^n : x \longmapsto x$$

et

$$\mathrm{pr}_j : \mathbb{R}^n \longrightarrow \mathbb{R} : x \longmapsto x_j \quad \text{pour } j \in \{1, \ldots, n\} \ .$$

EXEMPLE 1 La fonction

$$|\mathrm{id}| : \mathbb{R}^n \longrightarrow \mathbb{R} : x \longmapsto |x| = \left(x_1^2 + \ldots + x_n^2\right)^{\frac{1}{2}}$$

est continûment partiellement dérivable sur $\mathbb{R}^n \setminus \{0\}$ et

$$\partial_j |\mathrm{id}| = \frac{\mathrm{pr}_j}{|\mathrm{id}|} \ .$$

On écrit aussi

$$\partial_j |x| = \frac{x_j}{|x|} \ .$$

Les ensembles niveaux sont

$$N_{|\cdot|}(r) = \{x \in \mathbb{R}^n \mid |x| = r\} =: \mathbb{S}_r^{n-1} \ .$$

EXEMPLE 2 Soient J un intervalle de $\mathbb{R}_+^*$, $g : J \longrightarrow \mathbb{K}$ une fonction (continûment) dérivable et

$$g(|\cdot|) : |\cdot|^{-1}(J) \longrightarrow \mathbb{K} : x \longmapsto g(|x|) \ .$$

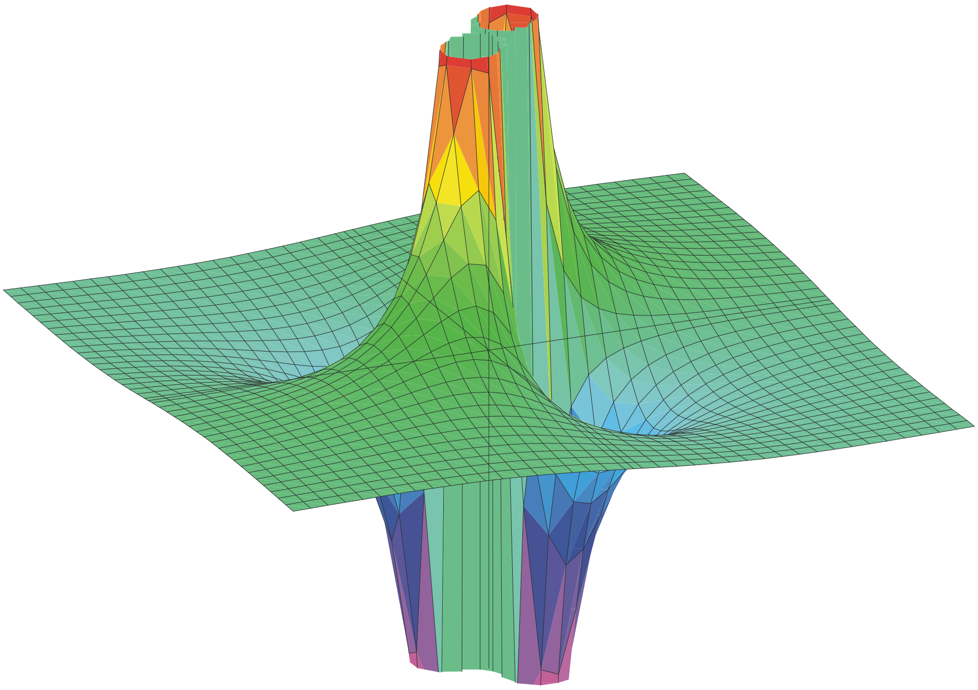
Alors $g(|\cdot|)$ est (continûment) partiellement dérivable sur $|\cdot|^{-1}(J) \subset \mathbb{R}^n \setminus \{0\}$ et

$$\partial_j g(|\mathrm{id}|) = g'(|\mathrm{id}|) \cdot \partial_j |\mathrm{id}| = g'(|\mathrm{id}|) \cdot \frac{\mathrm{pr}_j}{|\mathrm{id}|} \ .$$

On remarquera que $|\cdot|^{-1}\left(\mathbb{R}_+^*\right) = \mathbb{R}^n \setminus \{0\}$.

EXEMPLE 3 Soient $n \geqslant 2$ et

$$f : \mathbb{R}^n \longrightarrow \mathbb{R} : x \longmapsto \begin{cases} \frac{x_1 \cdot x_2 \cdot \ldots \cdot x_n}{|x|^{2n}} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases} \ .$$



$(x, y) \longmapsto \frac{xy}{(x^2+y^2)^2}$

D'après l'exemple précédent cette fonction est continûment partiellement dérivable sur $\mathbb{R}^n \setminus \{0\}$ et on a

$$\partial_1 f(x) = \frac{x_2 \cdot \ldots \cdot x_n}{|x|^{2n}} + x_1 \cdot \ldots \cdot x_n \cdot (-2n) \cdot \frac{1}{|x|^{2n+1}} \cdot \frac{x_1}{|x|} =$$

$$= \frac{x_2 \cdot \ldots \cdot x_n}{|x|^{2n}} \cdot \left(1 - 2n \cdot \frac{x_1^2}{|x|^2}\right) ,$$

avec une formule analogue pour $\partial_j f(x)$.

Qu'en est-il en 0 ? On a $f(t \cdot e_j) = 0$ tout $t \in \mathbb{R}$ et tout $j \in \{1, \ldots, n\}$, donc

$$\partial_j f(0) = 0 .$$

Ceci montre que f est partiellement dérivable sur $\mathbb{R}^n$, mais f n'est pas continue en 0 . En effet soit

$$x_k := \left(\frac{1}{k}, \ldots, \frac{1}{k}\right) \in \mathbb{R}^n .$$

On a $\lim_k x_k = 0$ et $|x_k| = \frac{\sqrt{n}}{k}$, donc

$$\lim_k f(x_k) = \lim_k \frac{\left(\frac{1}{k}\right)^n}{\left(\frac{\sqrt{n}}{k}\right)^{2n}} = \lim_k \left(\frac{k}{n}\right)^n = \infty .$$

Finalement remarquons que f est évidemment "partiellement continûment dérivable" (cf. remarque 1) sur $\mathbb{R}^n \setminus \{0\}$, mais aussi en 0 , puisqu'on a

$$\partial_j f(0, \ldots 0, x_j, 0, \ldots, 0) = 0 \quad \text{pour tout } x_j \in \mathbb{R} .$$

REMARQUE 2 Cet exemple montre que la dérivabilité partielle d'une fonction n'entraîne pas la continuité (globale) de cette fonction. La notion de dérivabilité (totale) traitée en 11.9 supprimera cet inconvénient.

11.4 Gradient

REMARQUE Si l'on doit considérer un élément

$$v = (v_1, \ldots, v_n) \in \mathbb{K}^n$$

comme un vecteur, il est préférable de l'écrire comme un *vecteur colonne* , i.e. une matrice $n \times 1$

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = (v_1, \ldots, v_n)^\intercal \ ,$$

tandis qu'une forme linéaire $\mu \in (\mathbb{K}^n)^* = L(\mathbb{K}^n, \mathbb{K}) \approx \mathbb{K}^n$ est représentée par une matrice $1 \times n$, i.e. un *covecteur* ou *vecteur ligne*

$$\mu = (\mu_1, \ldots, \mu_n)$$

et on a

$$\mu(v) = (\mu_1, \ldots, \mu_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{j=1}^{n} \mu_j \cdot v_j \ .$$

DEFINITION Soit $f : X \longrightarrow \mathbb{K}$ une fonction partiellement dérivable en $x \in X$. Le vecteur

$$\operatorname{grad} f(x) := (\partial_1 f(x), \ldots, \partial_n f(x))^\intercal \in \mathbb{K}^n$$

s'appelle le *gradient* de f en x .

EXEMPLE Avec les notations des exemples 1 et 2 de 11.3, on a

$$\operatorname{grad} |\mathrm{id}| = \frac{\mathrm{id}}{|\mathrm{id}|}$$

et

$$\operatorname{grad} g(|\mathrm{id}|) = g'(|\mathrm{id}|) \cdot \frac{\mathrm{id}}{|\mathrm{id}|} \ .$$

PROPOSITION (Règle du produit) *Si $f, g : X \longrightarrow \mathbb{K}$ sont partiellement dérivables en $x \in X$, alors $f \cdot g$ est aussi partiellement dérivable en x et on a*

$$\operatorname{grad}(f \cdot g)(x) = f(x) \cdot \operatorname{grad} g(x) + g(x) \cdot \operatorname{grad} f(x) \ .$$

En effet, pour tout $j \in \{1, \ldots, n\}$, on a évidemment

$$\partial_j (f \cdot g)(x) = \partial_j f(x) \cdot g(x) + f(x) \cdot \partial_j g(x) \ ,$$

d'où le résultat. $\square$

11.5 Divergence

DEFINITION On dit qu'une application

$$f = (f_k)_{k=1,\ldots,m} : X \longrightarrow \mathbb{R}^m : x \longmapsto (f_1(x), \ldots, f_m(x))^\intercal$$

est *partiellement dérivable* en $x \in X$ si chaque composante f_k , pour $k \in \{1, \ldots, m\}$, est partiellement dérivable en x .

Une application $v : X \longrightarrow \mathbb{R}^n$ s'appelle un *champ de vecteurs* sur X . Si v est partiellement dérivable en x , le scalaire

$$\operatorname{div} v(x) := \sum_{j=1}^{n} \partial_j v_j(x) \ ,$$

s'appelle la *divergence* de v en x .

EXEMPLE 1 Si $f : X \longrightarrow \mathbb{R}$ est partiellement dérivable sur X , alors

$$\operatorname{grad} f : X \longrightarrow \mathbb{R}^n : x \longmapsto \operatorname{grad} f(x)$$

est un champ de vecteur sur X .

PROPOSITION (Règle du produit) *Soient $f : X \longrightarrow \mathbb{R}$, $v : X \longrightarrow \mathbb{R}^n$ et $x \in X$. Si f et v sont partiellement dérivables en x , alors*

$$f \cdot v : X \longrightarrow \mathbb{R}^n : y \longmapsto f(y) \cdot v(y)$$

est un champ de vecteurs partiellement dérivable en x et on a

$$\operatorname{div}(f \cdot v)(x) = (\operatorname{grad} f(x) | \, v(x)) + f(x) \cdot \operatorname{div} v(x) \ .$$

Pour tout $j \in \{1, \ldots, n\}$, on a

$$\partial_j (f \cdot v)_j (x) = \partial_j (f \cdot v_j)(x) = \partial_j f(x) \cdot v_j(x) + f(x) \cdot \partial_j v_j(x) \ ,$$

donc

$$\operatorname{div}(f \cdot v)(x) = \sum_{j=1}^{n} \partial_j (f \cdot v)_j (x) = \sum_{j=1}^{n} [\partial_j f(x) \cdot v_j(x) + f(x) \cdot \partial_j v_j(x)] =$$

$$= \sum_{j=1}^{n} \partial_j f(x) \cdot v_j(x) + \sum_{j=1}^{n} f(x) \cdot \partial_j v_j(x) = (\operatorname{grad} f(x) | \, v(x)) + f(x) \cdot \operatorname{div} v(x) \ .$$

□

EXEMPLE 2 Soit $\operatorname{id} : \mathbb{R}^n \longrightarrow \mathbb{R}^n : x \longmapsto x$ le champ de vecteur identité. On a

$$\operatorname{div} \operatorname{id} = n \ .$$

En effet $\mathrm{id}_j : x \longmapsto x_j$, donc

$$\operatorname{div} x = \sum_{j=1}^{n} \partial_j x_j = \sum_{j=1}^{n} 1 = n \ .$$

$\square$

EXEMPLE 3 Plus généralement si T est une application linéaire dans $\mathbb{R}^n$ dont la matrice dans la base canonique est $(T_{k,l}) \in \mathbb{M}_{\mathbb{R}}(n \times n)$, alors

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^n : x \longmapsto Tx = \left(\sum_{l=1}^{n} T_{k,l} \cdot x_l \right)_{k=1,\ldots,n} ,$$

alors

$$\operatorname{div} T(x) = \operatorname{Tr} T := \sum_{j=1}^{n} T_{j,j}$$

est une fonction constante.

EXEMPLE 4 Soit $\frac{\mathrm{id}}{|\mathrm{id}|} : \mathbb{R}^n \smallsetminus \{0\} \longrightarrow \mathbb{R}^n : x \longmapsto \frac{x}{|x|}$. On a alors

$$\operatorname{div} \frac{\mathrm{id}}{|\mathrm{id}|} = \left(\operatorname{grad} \frac{1}{|\mathrm{id}|} \middle| \mathrm{id} \right) + \frac{1}{|\mathrm{id}|} \cdot \operatorname{div} \mathrm{id} = \left(-\frac{1}{|\mathrm{id}|^2} \cdot \frac{\mathrm{id}}{|\mathrm{id}|} \middle| \mathrm{id} \right) + \frac{n}{|\mathrm{id}|} = \frac{n-1}{|\mathrm{id}|}$$

d'après l'exemple 11.4.

11.6 Dérivées partielles d'ordre supérieur

DEFINITION 1 Soit $f : X \longrightarrow \mathbb{K}$. On dit que f est *deux fois partiellement dérivable* dans X si f est partiellement dérivable et si chaque dérivée partielle $\partial_j f$ est partiellement dérivable dans X .

THEOREME (de Schwarz) *Soient $f : X \longrightarrow \mathbb{K}$ une fonction deux fois partiellement dérivable dans X , $x \in X$ et $k, l \in \{1, \ldots, n\}$. Si $\partial_k \partial_l f$ et $\partial_l \partial_k f$ sont continues en x , alors*

$$\partial_k \partial_l f(x) = \partial_l \partial_k f(x) \ .$$

On peut supposer que $k < l$ et désignons par s et t les variables correspondantes. Nous n'écrirons pas les autres variables qui sont fixes et égales aux composantes respectives de x .

Appliquons tout d'abord la formule des accroissements finis (corollaire 11.7) à la fonction

$$s \longmapsto f(s,t) - f(s, x_l)$$

pour s et t au voisinage de x_k et x_l respectivement. Il existe $\xi(s,t)$ entre x_k et s tel que

$$f(s,t) - f(s,x_l) - [f(x_k,t) - f(x_k,x_l)] = [\partial_k f(\xi(s,t), t) - \partial_k(\xi(s,t), x_l)] \cdot (s - x_k) \ .$$

De même, en considérant la fonction

$$t \longmapsto \partial_k f(s,t) \ ,$$

il existe $\eta(s,t)$ entre x_l et t tel que

$$\partial_k f(s,t) - \partial_k f(s,x_l) = \partial_l \partial_k f(s, \eta(s,t)) \cdot (t - x_l) \ .$$

On a donc

$$f(s,t) - f(s,x_l) - [f(x_k,t) - f(x_k,x_l)] = \partial_l \partial_k f(\xi(s,t), \eta(\xi(s,t), t)) \cdot (t - x_l) \cdot (s - x_k) \ .$$

Par symétrie, il existe $\widetilde{\eta}(s,t)$ entre x_l et t , ainsi que $\widetilde{\xi}(s,t)$ entre x_k et s tels que

$$f(s,t) - f(x_k,t) - [f(s,x_l) - f(x_k,x_l)] = \partial_k \partial_l f\left(\widetilde{\xi}(s, \widetilde{\eta}(s,t)), \widetilde{\eta}(s,t)\right) \cdot (s - x_k) \cdot (t - x_l) \ .$$

Par comparaison, on obtient alors

$$\partial_l \partial_k f(\xi(s,t), \eta(\xi(s,t), t)) = \partial_k \partial_l f\left(\widetilde{\xi}(s, \widetilde{\eta}(s,t)), \widetilde{\eta}(s,t)\right) \ .$$

Mais si s tend vers x_k et t vers x_l , alors $\xi(s,t)$ et $\widetilde{\xi}(s,t)$ tendent vers x_k , tandis que $\eta(s,t)$ et $\widetilde{\eta}(s,t)$ tendent vers x_l . Ainsi $\eta(\xi(s,t), t)$ tend vers x_l et $\widetilde{\xi}(s, \widetilde{\eta}(s,t))$ vers x_k . Par la continuité de $\partial_l \partial_k f$ et $\partial_k \partial_l f$ en (x_k, x_l) , on en déduit évidemment que

$$\partial_l \partial_k f(x_k, x_l) = \partial_k \partial_l f(x_k, x_l) \ ,$$

ce qu'il fallait démontrer. □

DEFINITION 2 On définit les dérivées partielle d'ordre supérieur par récurrence. Une telle dérivée partielle est de la forme

$$\partial_{k_1}\left(\partial_{k_2}\left(\cdots \partial_{k_{m-1}}(\partial_{k_m} f)\right)\right)$$

pour une suite $(k_j)_{j=1,\ldots,m} \subset \{1,\ldots,n\}$. On dit que m est l' *ordre* de cette dérivation et que la fonction f est *m-fois partiellement dérivable* si toutes les dérivées partielles d'ordre $\leqslant m$ sont définies.

En supposant que **toutes** les dérivées partielles d'ordre $\leqslant m$ sont continues (cf. remarque 11.3), la proposition montre que l'ordre dans lequel ces dérivées sont effectuées n'a pas d'importance.

Dans ce cas toute dérivée partielle s'écrit de la manière suivante :

DEFINITION 3 Pour tout $\alpha = (\alpha_1,\ldots,\alpha_n) \in \mathbb{N}^n$, on dit que c'est un *multi-indice* , on pose

$$\partial^\alpha f := \partial_1^{\alpha_1}\partial_2^{\alpha_2}\ldots\partial_n^{\alpha_n} f \ ,$$

où $\partial_j^{\alpha_j} f$ désigne la *dérivée partielle* d'ordre α_j par rapport à la j-ième variable. L'entier

$$|\alpha|_1 := \alpha_1 + \alpha_2 + \ldots + \alpha_n$$

est évidemment l'ordre de ∂^α .

EXEMPLE Si X est une partie ouverte de $\mathbb{R}^3$ et $v : X \longrightarrow \mathbb{R}^3$ un champ de vecteurs partiellement dérivable en $x \in X$, le vecteur

$$\operatorname{rot} v\,(x) := \left(\partial_2 v_3\,(x) - \partial_3 v_2\,(x)\,, \partial_3 v_1\,(x) - \partial_1 v_3\,(x)\,, \partial_1 v_2\,(x) - \partial_2 v_1\,(x)\right)^{\mathsf{T}}$$

s'appelle le *rotationnel* de v en x .

Si $f : X \longrightarrow \mathbb{R}$ est deux fois continûment partiellement dérivable, alors

$$\operatorname{rot}\operatorname{grad} f = 0 \ .$$

En effet la première composante de $\operatorname{rot}\operatorname{grad} f$ est

$$\partial_2\partial_3 f - \partial_3\partial_2 f = 0 \ .$$

Les autres s'obtiennent par permutation circulaire. □

Beaucoup de champ v que l'on rencontre en physique s'écrivent comme le gradient d'une fonction f , dite *fonction potentiel* . Une condition nécessaire pour qu'il en soit ainsi, si v est continûment partiellement dérivable, est donc $\operatorname{rot} v = 0$. Cette condition est suffisante dans certain cas, par exemple dans $\mathbb{R}^3$, mais pas dans $\mathbb{R}^3 \setminus \{(0,0)\} \times \mathbb{R}$ (cf. exemple 15.7.3).

EXERCICE 1 Soit

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x,y) \longmapsto \begin{cases} \frac{x\cdot y^3}{x^2+y^2} & (x,y) \neq (0,0) \\ & \text{si} \\ 0 & (x,y) = (0,0) \end{cases} \ .$$

Montrer

(a) f est continue.

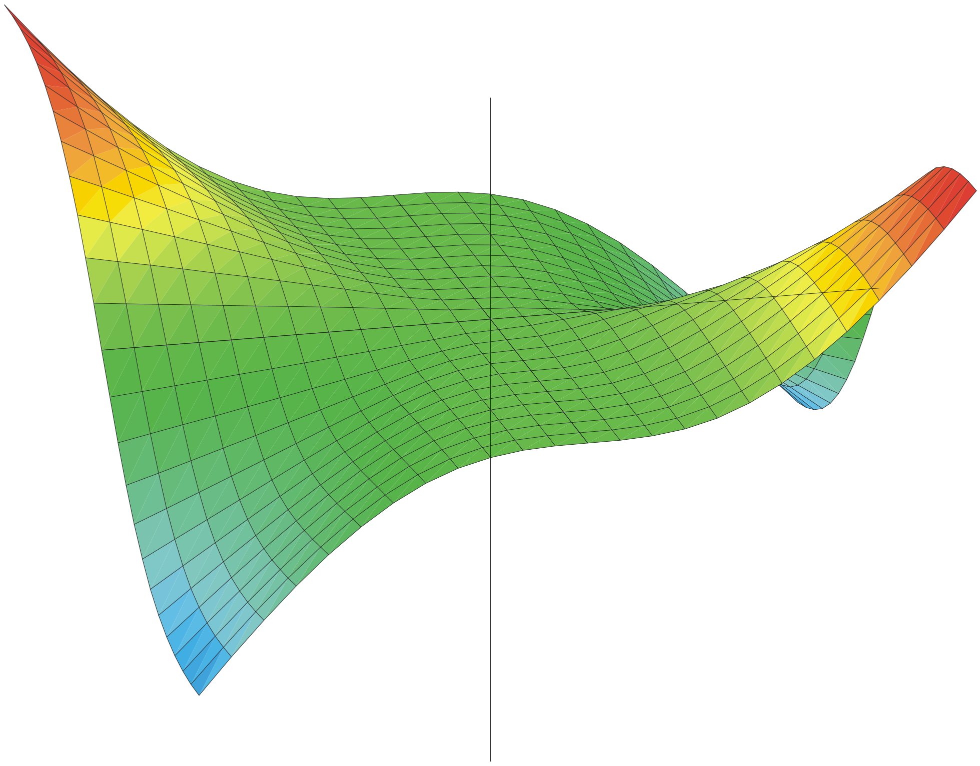
(b) f est continûment partiellement dérivable.

(c) f est deux fois partiellement dérivable.

(d) On a

$$\partial_1\partial_2 f\,(0,0) \neq \partial_2\partial_1 f\,(0,0) \ .$$

(e) Déterminer la dérivée partielle du deuxième ordre qui n'est pas continue.



$(x, y) \longmapsto \frac{xy^3}{x^2+y^2}$

EXERCICE 2 On considère le champ de vecteurs sur $\mathbb{R}^3$

$$v : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} x^2 + y^2 + z^2 \\ 2 \cdot xy + z \\ 2 \cdot xz + y \end{pmatrix} .$$

Montrer que $\operatorname{rot} v = 0$ et déterminer une fonction partiellement dérivable $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ telle que $v = \operatorname{grad} f$.

11.7 Opérateur de Laplace

DEFINITION Si f est deux fois partiellement dérivable dans X , on définit le *laplacien* de f sur X par

$$\Delta f := \operatorname{div} \operatorname{grad} f = \sum_{j=1}^{n} \partial_j^2 f \ .$$

On dit que Δ est l' *opérateur de Laplace* . L'équation aux dérivées partielle $\Delta f = 0$ s'appelle l' *équation de Laplace* ou *équation du potentiel* . On dit que les solutions de cette équation sont *harmoniques* .

EXEMPLE 1 Soient J un intervalle de $\mathbb{R}_+^*$, $g : J \longrightarrow \mathbb{R}$ deux fois continûment dérivable et

$$g(|\cdot|) : |\cdot|^{-1}(J) \longrightarrow \mathbb{R} \ .$$

Utilisant les exemples 11.4 et 11.5.4 et la règle du produit 11.5, on a

$$\Delta g(|\mathrm{id}|) = \operatorname{div} \operatorname{grad} g(|\mathrm{id}|) = \operatorname{div}\left(g'(|\mathrm{id}|) \cdot \frac{\mathrm{id}}{|\mathrm{id}|}\right) =$$

$$= \left(\operatorname{grad} g'(|\mathrm{id}|) \middle| \frac{\mathrm{id}}{|\mathrm{id}|}\right) + g'(|\mathrm{id}|) \cdot \operatorname{div} \frac{\mathrm{id}}{|\mathrm{id}|} =$$

$$= \left(g''(|\mathrm{id}|) \cdot \frac{\mathrm{id}}{|\mathrm{id}|} \middle| \frac{\mathrm{id}}{|\mathrm{id}|}\right) + g'(|\mathrm{id}|) \cdot \frac{n-1}{|\mathrm{id}|} =$$

$$= g''(|\mathrm{id}|) + \frac{n-1}{|\mathrm{id}|} \cdot g'(|\mathrm{id}|) \ .$$

En particulier dans $\mathbb{R}^n \setminus \{0\}$ on a

$$\Delta \frac{1}{|x|^{n-2}} = (n-2) \cdot (n-1) \cdot \frac{1}{|x|^n} - \frac{n-1}{|x|} \cdot (n-2) \cdot \frac{1}{|x|^{n-1}} = 0 \ .$$

Ceci montre que $\frac{1}{|\mathrm{id}|^{n-2}}$ est une fonction harmonique sur $\mathbb{R}^n \setminus \{0\}$. La fonction $\frac{1}{|\mathrm{id}|}$ est donc harmonique sur $\mathbb{R}^3 \setminus \{0\}$. Remarquons finalement que

$$\operatorname{grad} \frac{1}{|\mathrm{id}|} = -\frac{1}{|\mathrm{id}|^2} \cdot \frac{\mathrm{id}}{|\mathrm{id}|}$$

est un champ attractif proportionnel au carré de la distance à l'origine. C'est la *loi d'attraction de Newton* ou de *Coulomb* .

EXEMPLE 2 Soit X un ouvert de $\mathbb{R}^{1+n}$ et notons $(t, x) = (t, x_1, \ldots, x_n)$ un point de $\mathbb{R}^{1+n}$. On désigne par Δ_x l'opérateur de Laplace par rapport aux variables $x_1, \ldots, x_n$.

Etant donné $c, k \in \mathbb{R}_+^*$, les équations aux dérivées partielles

$$\partial_t^2 f = c^2 \cdot \Delta_x f \quad \text{et} \quad \partial_t f = k \cdot \Delta_x f$$

s'appellent respectivement l' *équation des ondes* et l' *équation de la chaleur* .

EXERCICE 1 Soit $c \in \mathbb{R}_+^*$, $v \in \mathbb{R}^n$ tel que $|v| = 1$ et $g : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction deux fois continûment dérivable. Montrer que la fonction

$$f : \mathbb{R}^{1+n} \longrightarrow \mathbb{R} : (t, x) \longmapsto g\left((v|\, x) - c \cdot t\right)$$

est solution de l'équation des ondes

$$\partial_t^2 f = c^2 \cdot \Delta_x f \ .$$

EXERCICE 2 Montrer que la fonction

$$f : \mathbb{R}_+^* \times \mathbb{R}^n \longrightarrow \mathbb{R} : (t, x) \longmapsto t^{-\frac{n}{2}} \cdot \exp\left(-\frac{|x|^2}{4 \cdot t}\right)$$

est solution de l'équation de la chaleur

$$\partial_t f = \Delta_x f \ .$$

11.8 Norme d'une application linéaire

PROPOSITION *Soient F, G des espace normés et $T : F \longrightarrow G$ une application linéaire. Pour que T soit continue, il faut et il suffit qu'il existe $M \in \mathbb{R}_+$ tel que*

$$\|Tf\|_G \leqslant M \cdot \|f\|_F \quad \textit{pour tout } f \in F \ . \tag{$*$}$$

Si T est continue, elle est en particulier continue en 0 . Il existe donc $\delta > 0$ tel que

$$T(B(0,\delta)) \subset B(0,1) \ .$$

L'inégalité étant trivialement vraie pour $f = 0$, nous pouvons supposer que $f \neq 0$. Mais alors, en posant $M := \frac{1}{\delta}$, on a

$$\frac{1}{M \cdot \|f\|_F} \cdot f \in B(0,\delta) \ ,$$

donc

$$T\left(\frac{1}{M \cdot \|f\|_F} \cdot f\right) \in B(0,1) \ ,$$

i.e.

$$\left\|T\left(\frac{1}{M \cdot \|f\|_F} \cdot f\right)\right\|_G \leqslant 1 \ ,$$

d'où l'on tire l'inégalité.

Réciproquement nous pouvons supposer que $M > 0$. Soit $f_0 \in F$. Pour $\varepsilon > 0$ donné, posons $\delta := \frac{\varepsilon}{M}$. Pour tout $f \in F$ tel que $\|f - f_0\|_F \leqslant \delta$, il vient

$$\|Tf - Tf_0\|_G = \|T(f - f_0)\|_G \leqslant M \cdot \|f - f_0\|_F \leqslant \varepsilon \ ,$$

ce qui prouve la continuité de T en f_0 . □

REMARQUE 1 La plus petite constante $M \in \overline{\mathbb{R}}_+$ satisfaisant à $(*)$ est

$$\|T\| := \sup_{f \in F, \|f\|_F \leqslant 1} \|Tf\|_G \ .$$

Ainsi T est continue si, et seulement si, $\|T\| < \infty$.

En effet $(*)$ est équivalent à

$$\left\|T\left(\frac{f}{\|f\|_F}\right)\right\|_G = \frac{1}{\|f\|_F} \cdot \|Tf\|_G \leqslant M \quad \text{pour tout } f \in F \setminus \{0\} \ ,$$

donc à

$$\|Tf\|_G \leqslant M \quad \text{pour tout } f \in F \text{ tel que } \|f\|_F \leqslant 1 \ .$$

DEFINITION On dit que $\|T\|$ est la *norme* de l'application linéaire T .

REMARQUE 2 Il n'est pas difficile de montrer que l'ensemble $\mathcal{L}(F,G)$ des applications linéaires continues est un sous-espace vectoriel de $L(F,G)$ et que

$$A \longmapsto \|A\| : \mathcal{L}(F,G) \longrightarrow \mathbb{R}_+$$

est une norme, dite la *norme opérateur* .

Si G est un espace de Banach, il en est de même de $\mathcal{L}(F,G)$.

EXEMPLE 1 La forme linéaire

$$\int_a^b : f \longmapsto \int_a^b f : \left(\mathcal{C}([a,b]), \|\cdot\|_\infty\right) \longrightarrow \mathbb{K}$$

est continue de norme $b-a$.

En effet, pour tout $f \in \mathcal{C}([a,b])$, on a

$$\left|\int_a^b f\right| \leqslant \int_a^b |f| \leqslant (b-a)\cdot\|f\|_\infty \ ,$$

$$\int_a^b 1 = b-a \quad \text{et} \quad \|1\|_\infty = 1\ .$$

□

EXEMPLE 2 On montre facilement (exercice) que

$$\|f\|_{(1),\infty} := \|f\|_\infty + \|f'\|_\infty$$

définit une norme sur $\mathcal{C}^{(1)}([a,b])$ et en fait un espace de Banach. L'application linéaire

$$\partial : f \longmapsto f' : \left(\mathcal{C}^{(1)}([a,b]), \|\cdot\|_{(1),\infty}\right) \longrightarrow \left(\mathcal{C}([a,b]), \|\cdot\|_\infty\right)$$

est continue de norme 1 .

En effet, pour tout $f \in \mathcal{C}^{(1)}([a,b])$, on a

$$\|\partial f\|_\infty = \|f'\|_\infty \leqslant \|f\|_{(1),\infty} \ ,$$

$$\left\|\partial \frac{1}{k\cdot(b-a)^{k-1}}\cdot(\mathrm{id}-a)^k\right\|_\infty = \left\|\frac{1}{(b-a)^{k-1}}\cdot(\mathrm{id}-a)^{k-1}\right\|_\infty = 1$$

et

$$\left\|\frac{1}{k\cdot(b-a)^{k-1}}\cdot(\mathrm{id}-a)^k\right\|_{(1),\infty} = \left\|\frac{1}{k\cdot(b-a)^{k-1}}\cdot(\mathrm{id}-a)^k\right\|_\infty + \left\|\frac{1}{(b-a)^{k-1}}\cdot(\mathrm{id}-a)^{k-1}\right\|_\infty =$$

$$= \frac{b-a}{k} + 1 \to 1 \quad \text{lorsque } k\to\infty\ .$$

□

EXEMPLE 3 Par contre

$$\partial : f \longmapsto f' : \left(\mathcal{C}^{(1)}([a,b]), \|\cdot\|_\infty\right) \longrightarrow \left(\mathcal{C}([a,b]), \|\cdot\|_\infty\right)$$

n'est pas continue, puisque

$$\left\| \frac{1}{(b-a)^k} \cdot (\mathrm{id} - a)^k \right\|_\infty = 1$$

et

$$\left\| \partial \frac{1}{(b-a)^k} \cdot (\mathrm{id} - a)^k \right\|_\infty = \left\| \frac{k}{(b-a)^k} \cdot (\mathrm{id} - a)^{k-1} \right\|_\infty =$$

$$= k \to \infty \quad \text{lorsque } k \to \infty \ .$$

REMARQUE 3 Soit

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

une application linéaire. Rappelons que, en utilisant les base canoniques de $\mathbb{R}^n$ et $\mathbb{R}^m$, on représente T par sa matrice $m \times n$

$$(T_{k,l})_{\substack{k=1,\ldots,m \\ l=1,\ldots,n}} = \begin{pmatrix} T_{1,1} & \cdots & T_{1,n} \\ \vdots & & \vdots \\ T_{n,1} & \cdots & T_{n,n} \end{pmatrix} \ .$$

COROLLAIRE *Toute application linéaire $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ est continue, i.e.*

$$|Tv| \leqslant \|T\| \cdot |v| \quad \textit{pour tout } v \in \mathbb{R}^n \ .$$

Plus précisément, si $T = (T_{k,l})_{\substack{k=1,\ldots,m \\ l=1,\ldots,n}}$, alors en posant $|T|_\infty := \max_{\substack{k=1,\ldots,m \\ l=1,\ldots,n}} |T_{k,l}|$, on a

$$|T|_\infty \leqslant \|T\| \leqslant n\sqrt{m} \cdot |T|_\infty < \infty \ .$$

En effet, pour tout $k = 1, \ldots, m$ et $l = 1, \ldots, n$, on a tout d'abord

$$|T_{k,l}| = |(Te_l)_k| \leqslant |Te_l| \leqslant \|T\| \ ,$$

donc

$$|T|_\infty \leqslant \|T\| \ .$$

D'autre part, pour tout $v \in \mathbb{R}^n$ tel que $|v| \leqslant 1$, grâce au théorème 10.13, on a

$$|(Tv)_k| = \left| \sum_{l=1}^n T_{k,l} \cdot v_l \right| \leqslant \sum_{l=1}^n |T_{k,l}| \cdot |v_l| \leqslant n \cdot |T|_\infty \cdot |v|_\infty \leqslant n \cdot |T|_\infty \cdot |v| \ ,$$

donc $|Tv|_\infty \leqslant n \cdot |T|_\infty$. Par suite

$$|Tv| \leqslant \sqrt{m} \cdot |Tv|_\infty \leqslant n\sqrt{m} \cdot |T|_\infty \ ,$$

donc

$$\|T\| \leqslant n\sqrt{m} \cdot |T|_\infty \ .$$

$\square$

REMARQUE 4 La continuité de T peut se démontrer directement de la manière suivante. Il suffit de montrer que chaque composante

$$T_k : x \longmapsto (Tx)_k = \sum_{l=1}^{n} T_{k,l} \cdot x_l : \mathbb{R}^n \longrightarrow \mathbb{R} \quad , \text{ pour } k = 1, \ldots, m \ ,$$

est continue. Mais cela est évident, puisque

$$T_k = \sum_{l=1}^{n} T_{k,l} \cdot \mathrm{pr}_l \ \ .$$

Remarquons en outre que

$$T : \left(\mathbb{R}^n, |\cdot|_p\right) \longrightarrow \left(\mathbb{R}^m, |\cdot|_q\right)$$

est aussi continue pour tout $p, q \in [1, \infty]$, puisque les normes considérées sont équivalentes.

REMARQUE 5 Il est clair que $T \longmapsto |T|_\infty$ est aussi une norme, puisque l'espace vectoriel $L(\mathbb{R}^n, \mathbb{R}^m)$ s'identifie à $\mathbb{R}^{m \times n}$.

11.9 Dérivabilité totale

Rappelons (théorème 8.1) qu'une fonction $f : J \longrightarrow \mathbb{R}$ est dérivable en $x \in J$ si, et seulement si, il existe $a \in \mathbb{R}$ tel que la fonction $\varphi : J \longrightarrow \mathbb{R}$ définie par

$$f(y) = f(x) + a \cdot (y - x) + \varphi(y)$$

satisfasse à

$$\lim_{x \neq y \to x} \frac{\varphi(x)}{y - x} = 0 \ .$$

Ceci nous conduit à poser la

DEFINITION 1 Soient $f : X \longrightarrow \mathbb{R}^m$ une application et $x \in X$. Nous dirons que f est *(totalement) dérivable* en x s'il existe une application linéaire $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ telle que l'application $\varphi : X \longrightarrow \mathbb{R}^m$ définie par

$$f(y) = f(x) + T(y - x) + \varphi(y)$$

satisfasse à

$$\lim_{x \neq y \to x} \frac{\varphi(y)}{|y - x|} = 0 \ .$$

PROPOSITION *Si f est dérivable en $x \in X$, alors f est continue et partiellement dérivable en x . En outre T et sa matrice $(T_{k,l})_{\substack{k=1,\ldots,m \\ l=1,\ldots,n}}$ sont univoquement déterminées et on a*

$$T_{k,l} = \partial_l f_k(x) \ .$$

On a

$$\lim_{x \neq y \to x} \varphi(y) = \lim_{x \neq y \to x} |y - x| \cdot \frac{\varphi(y)}{|y - x|} = 0 \ ,$$

donc

$$\lim_{y \to x} f(y) = \lim_{x \neq y \to x} f(y) = f(x) + \lim_{x \neq y \to x} [T(y - x) + \varphi(y)] = f(x)$$

par le corollaire 11.8, ce qui prouve que f est continue en x .

D'autre part, pour tout $k = 1, \ldots, m$ et $l = 1, \ldots, n$, on a

$$f_k(x + t \cdot e_l) = f_k(x) + T_{k,l} \cdot t + \varphi_k(x + t \cdot e_l) \quad \text{avec } \lim_{0 \neq t \to 0} \frac{\varphi_k(x + t \cdot e_l)}{t} = 0 \ ,$$

donc

$$\partial_l f_k(x) = T_{k,l}$$

par le théorème 8.1 . □

COROLLAIRE *Supposons que f soit partiellement dérivable en x et soit*

$$f_k(y) = f_k(x) + \sum_{l=1}^{n} \partial_l f_k(x) \cdot (y_l - x_l) + \varphi_k(y) \quad \textit{pour tout } k \in \{1, \ldots, m\} \ .$$

Pour que f soit dérivable en x , il faut et il suffit que

$$\lim_{x\neq y\to x} \frac{\varphi_k(y)}{|y-x|} = 0 \quad \textit{pour tout } k \in \{1,\ldots,m\} \ .$$

DEFINITION 2 Si f est dérivable en x , alors l'unique application linéaire satisfaisant à la condition de la définition 1 s'appelle la *dérivée* de f en x et est notée $Df(x)$. On dit que la matrice correspondante

$$(\partial_l f_k(x))_{\substack{k=1,\ldots,m\\ l=1,\ldots,n}} = \begin{pmatrix} \partial_1 f_1(x) & \cdots & \partial_n f_1(x) \\ \vdots & & \vdots \\ \partial_1 f_m(x) & \cdots & \partial_n f_m(x) \end{pmatrix}$$

est la *matrice jacobienne* de f en x .

Si f est dérivable en tout point $x \in X$, on dit que

$$Df : x \longmapsto Df(x) : X \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \simeq \mathbb{M}_{\mathbb{R}}(m\times n) \simeq \mathbb{R}^{m\times n}$$

est l' *application dérivée* de f .

REMARQUE La notion de dérivabilité (totale) supprime la pathologie rencontrée avec la dérivabilité partielle dans l'exemple 11.3.3. La fonction de cet exemple est partiellement dérivable, mais non-dérivable en 0 , puisqu'elle n'y est pas continue.

EXERCICE 1 Soit $g : \mathbb{R}_+ \longrightarrow \mathbb{R}$ une fonction dérivable en 0 telle que $g'(0) = 0$. Montrer que la fonction

$$g(|\cdot|) : \mathbb{R}^n \longrightarrow \mathbb{R} : x \longmapsto g(|x|)$$

est partiellement dérivable en 0 . Est-elle (totalement) dérivable en 0 ?

EXERCICE 2 Soit $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$ une fonction bornée. Montrer que la fonction

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x,y) \longmapsto x \cdot y \cdot g(x,y)$$

est continue et partiellement dérivable en $(0,0)$. Est-elle totalement dérivable en $(0,0)$?

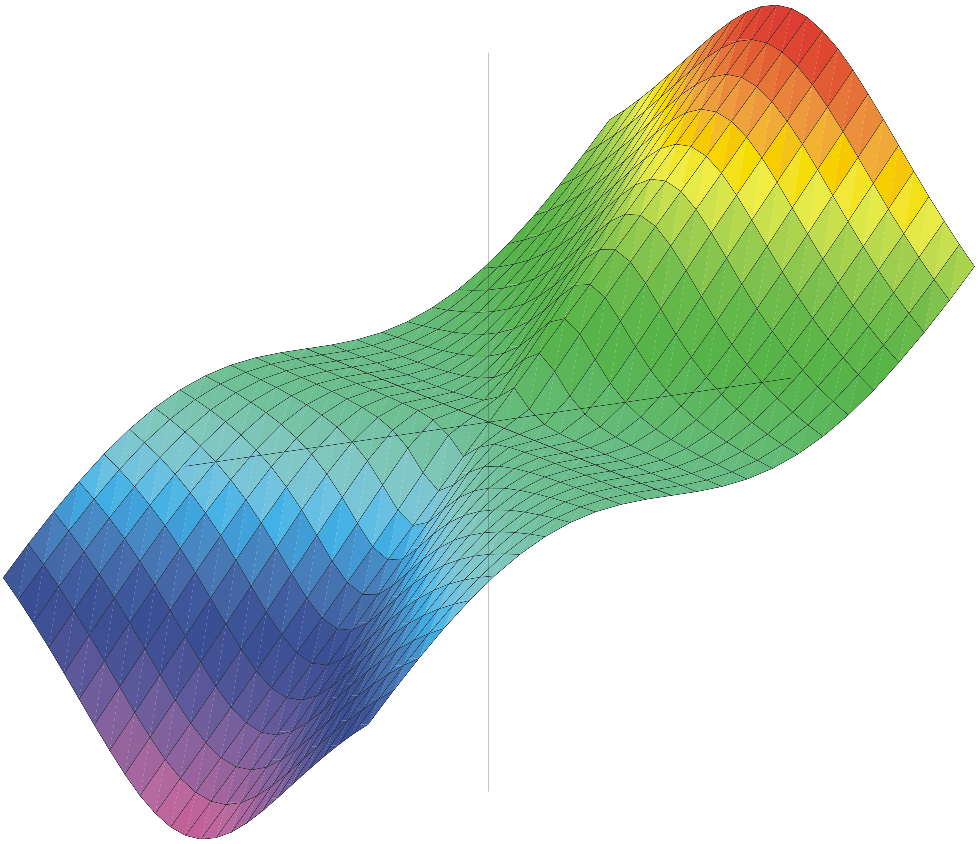
EXERCICE 3 Soit

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x,y) \longmapsto \begin{cases} \frac{y^3}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \phantom{\text{si }} (x,y) = (0,0) \end{cases} \ .$$

Montrer

(a) f est continue et partiellement dérivable en $(0,0)$, mais pas (totalement) dérivable.

(b) Si $\gamma : J \longrightarrow \mathbb{R}^2$ est une courbe paramétrée dérivable telle que $\gamma(0) = (0,0)$ et $\gamma'(0) \neq (0,0)$, alors $f \circ \gamma$ est dérivable en 0 .



$(x, y) \longmapsto \frac{y^3}{x^2+y^2}$

11.10 Applications continûment dérivables

DEFINITION On dit que $f : X \longrightarrow \mathbb{R}^m$ est *continûment dérivable* si

$$Df : X \longrightarrow \mathcal{L}\left(\mathbb{R}^n, \mathbb{R}^m\right)$$

est continue.

Si f est continûment dérivable, alors les dérivées partielles sont continues. Nous allons voir que la réciproque est vraie. Plus précisément, on a le

THEOREME *Soient $f : X \longrightarrow \mathbb{R}^m$ une application partiellement dérivable et $x \in X$. Si toutes les dérivées partielles $\partial_l f_k$ sont continues en x , alors f est dérivable en x .*

Soit $\varepsilon > 0$ tel que $B\left(x, \varepsilon, |\cdot|_\infty\right) \subset X$. Pour tout $y \in B\left(x, \varepsilon, |\cdot|_\infty\right)$, écrivons

$$y = x + \sum_{j=1}^{n} t_j \cdot e_j$$

et, pour $p = 0, \ldots, n$, posons

$$y_p := x + \sum_{j=1}^{p} t_j \cdot e_j \ .$$

On a donc $y_0 = x$ et $y_n = y$. Comme

$$y_p = y_{p-1} + t_p \cdot e_p \ ,$$

la formule des accroissements finis appliquée à la fonction

$$t \longmapsto f_k\left(y_{p-1} + t \cdot e_p\right) \ ,$$

pour $k = 0, \ldots, m$, entraîne l'existence d'un $\xi_{k,p}(y)$ entre 0 et t_p tel que

$$f_k\left(y_p\right) - f_k\left(y_{p-1}\right) = \partial_p f_k\left(y_p + \xi_{k,p}(y) \cdot e_p\right) \cdot t_p \ .$$

On en déduit que

$$f_k(y) - f_k(x) = \sum_{p=1}^{n} \left[f_k\left(y_p\right) - f_k\left(y_{p-1}\right)\right] = \sum_{p=1}^{n} \partial_p f_k\left(y_{p-1} + \xi_{k,p}(y) \cdot e_p\right) \cdot t_p =$$

$$= \sum_{p=1}^{n} \partial_p f_k(x) \cdot t_p + \varphi_k(y)$$

avec

$$\varphi_k(y) := \sum_{p=1}^{n} \left[\partial_p f_k\left(y_{p-1} + \xi_{k,p}(y) \cdot e_p\right) - \partial_p f_k(x)\right] \cdot t_p \ .$$

Mais puisque $|t_p| \leqslant |y - x|$, il vient

$$\left|\frac{\varphi_k(y)}{|y-x|}\right| = \left|\sum_{p=1}^{n}\left[\partial_p f_k\left(y_{p-1} + \xi_{k,p}(y)\cdot e_p\right) - \partial_p f_k(x)\right]\cdot\frac{t_p}{|y-x|}\right| \leqslant$$

$$\leqslant \sum_{p=1}^{n}\left|\partial_p f_k\left(y_{p-1} + \xi_{k,p}(y)\cdot e_p\right) - \partial_p f_k(x)\right| \ .$$

La continuité des dérivées partielles montre alors que

$$\lim\nolimits_{x\neq y\to x}\frac{\varphi_k(y)}{|y-x|} = 0 \ ,$$

ce qu'il fallait démontrer par le corollaire 11.9. □

COROLLAIRE *Une application $f : X \longrightarrow \mathbb{R}^m$ est continûment dérivable si, et seulement si, f est continûment partiellement dérivable.*

REMARQUE Entre les propriétés d'une application $f : X \longrightarrow \mathbb{R}^m$, on a les implications suivantes :

$$\begin{array}{ccc}
\text{continûment dérivable} & \Longleftrightarrow & \text{continûment partiellement dérivable} \\
\Downarrow & & \Downarrow \\
\text{dérivable} & \Longrightarrow & \text{partiellement dérivable} \\
\Downarrow & & \Downarrow \\
\text{continue} & \Longrightarrow & \text{séparément continue}
\end{array}$$

11.11 Exemples

EXEMPLE 1 Considérons tout d'abord le cas d'une courbe paramétrée $\gamma : J \longrightarrow \mathbb{R}^m$.

γ est (totalement) dérivable si, et seulement si, γ est dérivable (au sens de la définition 11.1.2, p. 370). On a

$$D\gamma(t) = \gamma'(t) \quad \text{pour tout } t \in J \ .$$

Remarquons que $D\gamma(t)$ et $\gamma'(t)$ sont des matrices à une colonne (cf. remarque 11.4).

Cela découle du corollaire 11.9 et du théorème 8.1 . □

EXEMPLE 2 Considérons maintenant le cas d'une fonction $f : X \longrightarrow \mathbb{R}$. Si f est (totalement) dérivable en $x \in X$, la dérivée $Df(x)$ est une matrice à une ligne

$$Df(x) = (\partial_1 f(x), \ldots, \partial_n f(x)) \ .$$

Ainsi

$$\operatorname{grad} f(x) = Df(x)^{\intercal} \ .$$

En outre, pour tout $v \in \mathbb{R}^n$, on a

$$Df(x)\,v = (\partial_1 f(x), \ldots, \partial_n f(x)) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{j=1}^n \partial_j f(x) \cdot v_j = (\operatorname{grad} f(x)|\, v) \ .$$

Pour tout $y \in X$, on a alors

$$f(y) = f(x) + (\operatorname{grad} f(x)|\, y - x) + \varphi(y) = f(x) + \sum_{j=1}^n \partial_j f(x) \cdot (y_j - x_j) + \varphi(y)$$

avec

$$\lim_{x \neq y \to x} \frac{\varphi(y)}{|y - x|} = 0 \ .$$

Le graphe de la fonction

$$y \longmapsto f(x) + (\operatorname{grad} f(x)|\, y - x) : \mathbb{R}^n \longrightarrow \mathbb{R}$$

est un hyperplan tangent au graphe de f en $(x, f(x))$ car, comme dans la remarque 8.1.3, cette fonction fournit la meilleure approximation affine de f au voisinage de x .

Pour $j \in \{1, \ldots, n\}$, les courbes paramétrées

$$t \longmapsto (x + t \cdot e_j, f(x + t \cdot e_j)) \in \mathbb{R}^n \times \mathbb{R} \ ,$$

sont contenues dans $\operatorname{Gr} f$ et les vecteurs tangents en 0

$$(e_j, \partial_j f(x)) \in \mathbb{R}^n \times \mathbb{R}$$

sont contenus dans l'hyperplan tangent en $(x, f(x))$. On en déduit :

PROPOSITION *L'hyperplan tangent à* $\mathrm{Gr}\, f$ *est horizontal si, et seulement si,*

$$\operatorname{grad} f(x) = 0 \ .$$

Lors de considérations heuristiques on écrit volontiers

$$f(y) \simeq f(x) + (\operatorname{grad} f(x) |\, y - x) = f(x) + \sum_{j=1}^{n} \partial_j f(x) \cdot (y_j - x_j) \ ,$$

mais ceci cache toute l'information sur l'erreur que l'on fait.

EXEMPLE 3 Soit $\alpha \in \mathbb{N}^n$ un multi-indice. On dit que la fonction

$$x \longmapsto x^\alpha := x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \ldots \cdot x_n^{\alpha_n} : \mathbb{R}^n \longrightarrow \mathbb{R}$$

est un *monôme* . Une fonction du type

$$x \longmapsto \sum_{|\alpha|_1 \leqslant k} c_\alpha \cdot x^\alpha : \mathbb{R}^n \longrightarrow \mathbb{R}$$

s'appelle un *polynôme de degré* k s'il existe un α tel que

$$|\alpha|_1 = k \quad \text{et} \quad c_\alpha \neq 0 \ .$$

Rappelons que

$$|\alpha|_1 = \alpha_1 + \ldots + \alpha_n \ .$$

Soit f un polynome de degré $\leqslant 2$ sur $\mathbb{R}^n$, i.e.

$$f(x) = a + \sum_{j=1}^{n} b_j \cdot x_j + \sum_{k,l=1}^{n} T_{k,l} \cdot x_k x_l = a + (b|\, x) + (x|\, Tx)$$

avec

$$b = (b_j)_{j=1,\ldots,n} \in \mathbb{R}^n \quad \text{et} \quad T = (T_{k,l})_{k,l=1,\ldots,n} \in \mathcal{L}(\mathbb{R}^n) \ .$$

On peut supposer que T est symétrique en remplaçant T au besoin par $S := \frac{1}{2} \cdot (T + T^\intercal)$.

En effet, on a

$$(x|\, Tx) = (T^\intercal x|\, x) = (x\, |T^\intercal x) \ ,$$

donc

$$(x|\, Tx) = \frac{1}{2} \cdot \Big((x|\, Tx) + (x\, |T^\intercal x) \Big) = \left(x \,\middle|\, \frac{1}{2} \cdot (T + T^\intercal) \right) \ .$$

Pour tout $v \in \mathbb{R}^n$, il vient alors

$$f(x+v) = a + (b|\, x+v) + (x+v|\, T(x+v)) =$$

$$= f(x) + (b|\, v) + (x|\, Tv) + (Tx|\, v) + (v|\, Tv) =$$

$$= f(x) + (b + 2 \cdot Tx|\, v) + (v|\, Tv) \ .$$

D'après l'inégalité de Cauchy-Schwarz, on obtient

$$\left| \frac{(v|\, Tv)}{|v|} \right| \leqslant |Tv| \leqslant \|T\| \cdot |v| \ ,$$

par le corollaire 11.8, donc

$$\lim_{0 \neq v \to 0} \frac{(v|\, Tv)}{|v|} = 0 \ .$$

Ainsi la fonction

$$f : x \longmapsto a + (b|\, x) + (x|\, Tx) : \mathbb{R}^n \longrightarrow \mathbb{R}$$

est dérivable en x et

$$\operatorname{grad} f\,(x) = b + 2 \cdot Tx \ .$$

EXERCICE 1 Etant donné $b \in \mathbb{R}^n$ et $T \in \mathcal{L}\,(\mathbb{R}^n)$, calculer la dérivée de

$$\mathbb{R}^n \longrightarrow \mathbb{R}^n : x \longmapsto Tx + b \ .$$

EXERCICE 2 Calculer la dérivée de

$$\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} : (x, y) \longmapsto (x|\, y) \ .$$

11.12 Dérivation des applications composées

THEOREME *Soient $p \in \mathbb{N}^*$, Y un ouvert de $\mathbb{R}^p$, $g : Y \longrightarrow X$ et $f : X \longrightarrow \mathbb{R}^m$. Si g est dérivable en $y \in Y$ et si f est dérivable en $g(y)$, alors $f \circ g$ est dérivable en y et on a*

$$D(f \circ g)(y) = Df(g(y)) \circ Dg(y) \ .$$

$$\begin{array}{ccccc} & g & & f & \\ Y & \longrightarrow & X & \longrightarrow & \mathbb{R}^m \\ \cap & & \cap & & \\ \mathbb{R}^p & \longrightarrow & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & Dg(y) & & Df(g(y)) & \end{array}$$

Posons $U := Df(g(y)) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ et $V := Dg(y) : \mathbb{R}^p \longrightarrow \mathbb{R}^n$. Par définition, pour tout $v \in \mathbb{R}^p$ et $u \in \mathbb{R}^n$, on a

$$f(g(y)+u) = f(g(y)) + Uu + \varphi(u) \quad \text{avec } \lim_{0\neq u\to 0} \frac{\varphi(u)}{|u|} \ ,$$

et

$$g(y+v) = g(y) + Vv + \psi(v) \quad \text{avec } \lim_{0\neq v\to 0} \frac{\psi(v)}{|v|} \ .$$

Il vient alors

$$f \circ g(y+v) = f(g(y+v)) = f(g(y) + Vv + \psi(v)) =$$

$$= f(g(y)) + U(Vv + \psi(v)) + \varphi(Vv + \psi(v)) =$$

$$= f(g(y)) + UVv + U\psi(v) + \varphi(Vv + \psi(v)) \ .$$

Il nous suffit donc de montrer que

$$\lim_{0\neq v\to 0} \frac{1}{|v|} \cdot [U\psi(v) + \varphi(Vv + \psi(v))] = 0 \ .$$

Mais comme U est continue, on a tout d'abord

$$\lim_{0\neq v\to 0} \frac{1}{|v|} \cdot U\psi(v) = U\left(\lim_{0\neq v\to 0} \frac{1}{|v|} \cdot \psi(v)\right) = 0 \ .$$

Définissons maintenant

$$\tilde{\varphi}(u) := \begin{cases} \frac{\varphi(u)}{|u|} & \\ & \text{si} \\ 0 & \end{cases} \begin{array}{c} u \neq 0 \\ \\ u = 0 \end{array} \ .$$

On a $\varphi(u) = |u| \cdot \tilde{\varphi}(u)$ et $\lim_{u\to 0} \tilde{\varphi}(u) = 0$. En posant $M := \sup_{|v|\leqslant 1} |Vv|$ et grâce au corollaire 11.8, il vient en particulier

$$\frac{1}{|v|} \cdot |\varphi(Vv + \psi(v))| = \frac{1}{|v|} \cdot |Vv + \psi(v)| \cdot |\tilde{\varphi}(Vv + \psi(v))| \leqslant$$

$$\leqslant \frac{1}{|v|} \cdot (M \cdot |v| + |\psi(v)|) \cdot |\tilde{\varphi}(Vv + \psi(v))| \leqslant \left(M + \frac{|\psi(v)|}{|v|}\right) \cdot |\tilde{\varphi}(Vv + \psi(v))| \ .$$

On alors

$$\lim_{0 \neq v \to 0} \frac{1}{|v|} \cdot |\varphi(Vv + \psi(v))| = 0 \ ,$$

puisque la continuité de V montre que

$$\lim_{0 \neq v \to 0} (Vv + \psi(v)) = \lim_{0 \neq v \to 0} Vv + \lim_{0 \neq v \to 0} |v| \cdot \frac{|\psi(v)|}{|v|} = 0 \ .$$

Ceci finit de prouver le théorème. □

EXEMPLE Considérons le cas $f : X \longrightarrow \mathbb{R}$ et $g : Y \longrightarrow X$. On a donc $f \circ g : Y \longrightarrow \mathbb{R}$ et il vient

$$D(f \circ g)(y) = (\partial_1 f(g(y)), \ldots, \partial_n f(g(y))) \begin{pmatrix} \partial_1 g_1(y) & \cdots & \partial_p g_1(y) \\ \vdots & & \vdots \\ \partial_1 g_n(y) & \cdots & \partial_p g_n(y) \end{pmatrix} =$$

$$= \left(\sum_{j=1}^{n} \partial_j f(g(y)) \cdot \partial_q g_j(y)\right)_{q=1,\ldots,p} \ ,$$

ou bien en utilisant l'exemple 11.11.2

$$\operatorname{grad}(f \circ g)(y) = D(f \circ g)(y)^{\intercal} = [Df(g(y)) \circ Dg(y)]^{\intercal} =$$

$$= Dg(y)^{\intercal} \circ Df(g(y))^{\intercal} = Dg(y)^{\intercal} \operatorname{grad} f(g(y)) \ ,$$

ou encore

$$\partial_q(f \circ g)(y) = \sum_{j=1}^{n} \partial_j f(g(y)) \cdot \partial_q g_j(y) \ .$$

On peut se souvenir de cette formule en l'écrivant sous la forme ancienne :

$$x_j = x_j(y_1, \ldots, y_p) \quad , \quad f(x_1, \ldots, x_n) = f(x_1(y_1, \ldots, y_p), \ldots, x_n(y_1, \ldots, y_p))$$

et

$$\frac{\partial f}{\partial y_q} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \cdot \frac{\partial x_j}{\partial y_q} \ .$$

Le cas le plus important est celui où $p = n$ et $\Phi : Y \longrightarrow X$ est une bijection. On dit que l'on a fait le *changement de variables* $x = \Phi(y)$. Cette formule montre donc comment calculer les dérivées partielles par rapport aux nouvelles variables, connaissant les dérivées partielles par rapport aux anciennes.

EXERCICE 1 Etant donné $b \in \mathbb{R}^n$, calculer la dérivée de

$$\mathbb{R}^n \longrightarrow \mathbb{R} : x \longmapsto |x - b|^2 \ .$$

EXERCICE 2 Soient $f, g : X \longrightarrow \mathbb{R}^m$ des applications dérivables. Calculer la dérivée de

$$X \longrightarrow \mathbb{R} : x \longmapsto (f(x)| g(x)) \ .$$

EXERCICE 3 Soient $p \in \mathbb{N}^*$ et Y une partie ouverte de $\mathbb{R}^p$, ainsi que

$$\Phi : Y \longrightarrow X \quad , \quad f : X \longrightarrow \mathbb{R}$$

des applications dérivables.

(a) Montrer que, pour tout $q \in \{1, \ldots, p\}$, on a

$$\partial_q (f \circ \Phi) = \sum_{j=1}^{n} \partial_j f \circ \Phi \cdot \partial_q \Phi_j \ .$$

(b) Soient

$$\Phi :]0, \infty[\times \mathbb{R} \longrightarrow \mathbb{R}^2 : (r, \varphi) \longmapsto (r \cdot \cos \varphi, r \cdot \sin \varphi)$$

et

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

une fonction deux fois dérivable. Montrer que

$$(\Delta f) \circ \Phi = \partial_r^2 (f \circ \Phi) + \frac{1}{r} \cdot \partial_r (f \circ \Phi) + \frac{1}{r^2} \cdot \partial_\varphi^2 (f \circ \Phi) \ .$$

11.13 Dérivées directionnelles

PROPOSITION *Soient $f : X \longrightarrow \mathbb{R}$ une fonction dérivable, $x \in X$, $\gamma : J \longrightarrow \mathbb{R}^n$ une courbe paramétrée dérivable contenue dans X et $t \in J$ tels que $\gamma(t) = x$. Alors*

$$(f \circ \gamma)'(t) = (\operatorname{grad} f(x) | \gamma'(t)) \ .$$

En effet on a

$$(f \circ \gamma)'(t) = D(f \circ \gamma)(t) = Df(\gamma(t)) \circ D\gamma(t) =$$

$$= (\partial_1 f(\gamma(t)), \ldots, \partial_n f(\gamma(t))) \begin{pmatrix} \partial\gamma_1(t) \\ \vdots \\ \partial\gamma_n(t) \end{pmatrix} = (\operatorname{grad} f(x) | \gamma'(t)) \ .$$

Par exemple, pour tout $\xi \in \mathbb{R}^n \setminus \{0\}$, on peut appliquer cette formule à la courbe paramétrée dérivable $s \longmapsto x + s \cdot \xi$, qui est définie dans un voisinage de 0 . On obtient

$$\partial_\xi f(x) := \lim_{0 \neq s \to 0} \frac{1}{s} \cdot [f(x + s \cdot \xi) - f(x)] = (\operatorname{grad} f(x) | \xi) \ .$$

DEFINITION On dit que $\partial_\xi f(x)$ est la *dérivée de f dans la direction ξ* .

REMARQUE 1 La proposition montre que le taux de variation de f au voisinage du point x , le long d'une courbe paramétrée γ passant par x , ne dépend que du vecteur tangent à cette courbe en ce point, puisqu'on a

$$(f \circ \gamma)'(t) = (\operatorname{grad} f(x) | \gamma'(t)) = \partial_{\gamma'(t)} f(x) \ .$$

REMARQUE 2 Si $\operatorname{grad} f(x) \neq 0$, alors $(\operatorname{grad} f(x) | \xi)$ est maximal sur

$$\mathbb{S}^{n-1} = \{\xi \in \mathbb{R}^n \mid |\xi| = 1\}$$

si, et seulement si, ξ est parallèle à $\operatorname{grad} f(x)$ et de même direction. Ceci montre que $\operatorname{grad} f(x)$ indique la direction de la plus grande pente.

EXERCICE 1 Soit $\gamma : J \longrightarrow X$ est une courbe paramétrée dérivable de hauteur constante, i.e. il existe $c \in \mathbb{R}$ tel que

$$f(\gamma(t)) = c \quad \text{pour tout } t \in J \ .$$

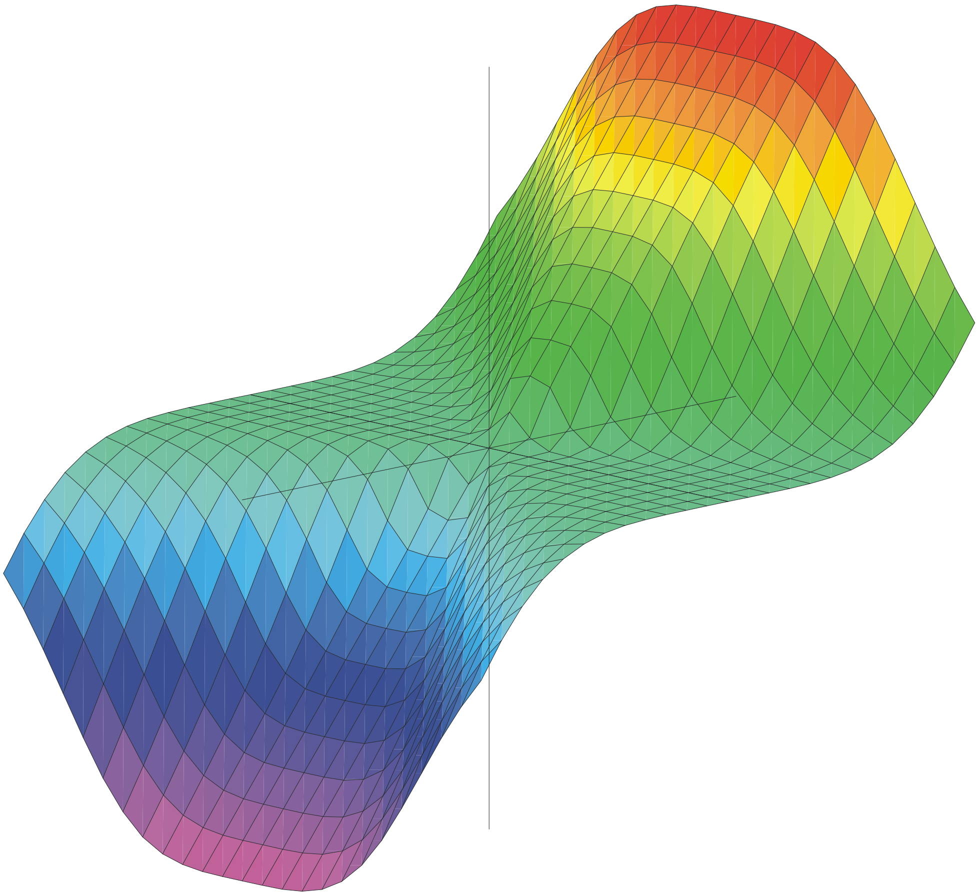
Montrer que

$$(\operatorname{grad} f(\gamma(t)) | \gamma'(t)) = 0 \quad \text{pour tout } t \in J$$

et donner une interprétation géométrique de ce fait.

EXERCICE 2 Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par

$$f(x,y) := \begin{cases} \frac{y^5}{2\cdot x^4+y^4} & \text{si } (x,y) \neq (0,0) \\ 0 & \phantom{\text{si }} (x,y) = (0,0) \end{cases} .$$



$(x,y) \longmapsto \frac{y^5}{2\cdot x^4+y^4}$

(a) Montrer que f est continue, partiellement dérivable et que toutes les dérivées directionnelles de f en $(0,0)$ existent, i.e. pour toute courbe paramétrée dérivable $\gamma : J \longrightarrow \mathbb{R}^2$ telle que $0 \in J$ et $\gamma(0) = (0,0)$, la fonction $f \circ \gamma$ est dérivable en 0 .

(b) Montrer que la formule de dérivation des fonctions composées n'est pas applicable à f et à la courbe paramétrée

$$\delta : \mathbb{R} \longrightarrow \mathbb{R}^2 : t \longmapsto (t,t) \ ,$$

i.e. on n'a pas

$$(f \circ \delta)'(0) = \sum_{j=1}^{2} \partial_j f(\delta(0)) \cdot \delta_j'(0) \ .$$

11.14 Inégalité de la moyenne

Rappelons qu'une matrice $m \times n$ peut être considérée comme un vecteur de $\mathbb{R}^{m \times n}$. Une fonction d'une variable à valeurs matricielles sera donc intégrée coefficient par coefficient (cf. définition 11.2.1, p. 373).

DEFINITION 1 Soit

$$T : J \longrightarrow \mathcal{L}\left(\mathbb{R}^n, \mathbb{R}^m\right) \approx \mathbb{M}_{\mathbb{R}}\left(m \times n\right) : t \longmapsto \left(T_{k,l}\left(t\right)\right)$$

une application continue. Pour tout $a, b \in J$, on définit l' *intégrale* (de Riemann) de T entre a et b par

$$\int_a^b T := \left(\int_a^b T_{k,l}\left(t\right)\, dt\right)_{\substack{k=1,\ldots,m \\ l=1,\ldots,n}} \ .$$

Pour tout $v \in \mathbb{R}^n$, on vérifie immédiatement que

$$\left(\int_a^b T\left(t\right)\, dt\right) v = \int_a^b T\left(t\right) v\, dt \ .$$

PROPOSITION (Formule des accroissements finis) *Soient $f : X \longrightarrow \mathbb{R}^m$ une application continûment dérivable et $x, y \in X$ tels que l'on ait $t \cdot x + [1-t] \cdot y \in X$ pour tout $t \in [0,1]$. On alors*

$$f\left(y\right) - f\left(x\right) = \left(\int_0^1 Df\left(t \cdot x + [1-t] \cdot y\right)\, dt\right)\left(y - x\right) \ .$$

Soit $\gamma : t \longmapsto t \cdot x + [1-t] \cdot y$. on a

$$D\gamma\left(t\right) = \gamma'\left(t\right) = x - y \ ,$$

donc

$$\left(\int_0^1 Df\left(t \cdot x + [1-t] \cdot y\right)\, dt\right)\left(y - x\right) = -\int_0^1 Df\left(t \cdot x + [1-t] \cdot y\right) \circ D\gamma\left(t\right)\, dt =$$

$$= -\int_0^1 D\left(f \circ \gamma\right)\left(t\right)\, dt = -\int_0^1 \left(f \circ \gamma\right)'\left(t\right)\, dt = f \circ \gamma\left(0\right) - f \circ \gamma\left(1\right) = f\left(y\right) - f\left(x\right) \ .$$

□

DEFINITION 2 On dit qu'une partie C de $\mathbb{R}^n$ est *convexe* si l'on a

$$t \cdot x + [1-t] \cdot y \in C \quad \text{pour tout } x, y \in C \text{ et } t \in [0,1] \ ,$$

i.e. si le segment joignant x à y est contenu dans C .

EXEMPLE Quel que soit $p \in [1, \infty]$, les boules fermées et ouvertes

$$B\left(x, r, |\cdot|_p\right) \quad \text{et} \quad D\left(x, r, |\cdot|_p\right)$$

de $\mathbb{R}^n$ sont convexes.

Une partie de $\mathbb{R}$ est convexe si, et seulement si, c'est un intervalle.

La première partie découle immédiatement de l'inégalité triangulaire. Si C est une partie convexe de $\mathbb{R}$, on vérifie immédiatement que

$$]\inf C, \sup C[\ \subset C\ ,$$

donc que C est un intervalle d'extrémités $\inf C$ et $\sup C$. $\square$

THEOREME (Inégalité de la moyenne) *Soient $f : X \longrightarrow \mathbb{R}^m$ une application continûment dérivable et C une partie convexe de X telle que*

$$M := \sup_{z\in C} \|Df(z)\| < \infty\ .$$

Pour tout $x, y \in C$, on a

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

En particulier f est uniformément continue sur C .

D'après la formule des accroissements finis, la proposition 11.2, p. 373, et le corollaire 11.8, on a

$$|f(y) - f(x)| = \left|\int_0^1 Df(t\cdot x + [1-t]\cdot y)(y-x)\,dt\right| \leqslant$$

$$\leqslant \int_0^1 |Df(t\cdot x + [1-t]\cdot y)(y-x)|\,dt \leqslant$$

$$\leqslant \int_0^1 \|Df(t\cdot x + [1-t]\cdot y)\| \cdot |y-x|\,dt \leqslant$$

$$\leqslant M\cdot|y-x|\ .$$

$\square$

REMARQUE D'après le corollaire 11.8, on a

$$\sup_{z\in C} \|Df(z)\| \leqslant n\sqrt{m}\cdot \sup_{z\in C} \max_{\substack{k=1,\ldots,m\\ l=1,\ldots,n}} |\partial_l f_k(z)| = n\sqrt{m}\cdot \max_{\substack{k=1,\ldots,m\\ l=1,\ldots,n}} \|\partial_l f_k\|_{\infty,C}\ .$$

On en déduit, par le corollaire 10.19, si C est compact que

$$\sup_{z\in C} \|Df(z)\| < \infty\ .$$

L'inégalité de la moyenne est donc applicable à toute partie fermée bornée et convexe de $\mathbb{R}^n$ contenue dans X (théorème de Heine-Borel 10.18), en particulier à tout segment joignant deux points de X et contenu dans X .

EXERCICE Soient X une partie ouverte et convexe de $\mathbb{R}^n$ et $f : X \longrightarrow \mathbb{R}^n$ une application continûment dérivable telle que, pour tout $x \in X$, on ait

$$(v|\,Df(x)\,v) > 0 \quad \text{pour tout } v \in \mathbb{R}^n \setminus \{0\}\ .$$

Montrer que f est injective. On utilisera l'exercice 11.2.1, p. 376.

11.15 Formule de Taylor

Si $f : X \longrightarrow \mathbb{R}$ est dérivable, alors Df est une application de X dans

$$\mathcal{L}(\mathbb{R}^n, \mathbb{R}) = (\mathbb{R}^n)^* \ ,$$

dont les éléments sont des matrices à une ligne. Au contraire, on a

$$\operatorname{grad} f = Df^{\intercal} : X \longrightarrow \mathbb{R}^n \ .$$

DEFINITION 1 Nous dirons que f est *deux fois (totalement) dérivable* en $x \in X$, si f est dérivable (dans un voisinage ouvert de x) et si $\operatorname{grad} f$ est dérivable en x . On a

$$\operatorname{Hess} f(x) := D(\operatorname{grad} f)(x) = \begin{pmatrix} \partial_1\partial_1 f(x) & \cdots & \partial_n\partial_1 f(x) \\ \vdots & & \vdots \\ \partial_1\partial_n f(x) & \cdots & \partial_n\partial_n f(x) \end{pmatrix} = (\partial_l\partial_k f(x))_{\substack{k=1,\ldots,n \\ l=1,\ldots,n}} \ .$$

On dit que c'est la *matrice hessienne* de f en x .

La proposition 11.9 montre que $\operatorname{grad} f$ est continue en x . D'autre part, grâce au corollaire 11.10, f est deux fois continûment dérivable si, et seulement si, f est deux fois continûment partiellement dérivable. Dans ce cas la matrice $\operatorname{Hess} f(x)$est symétrique par le théorème de Schwarz 11.6.

On remarquera que $D^2 f(x) := D(Df)(x)$ est matrice ligne formée de vecteurs ligne.

REMARQUE Plus généralement on peut considérer une application $f : X \longrightarrow \mathbb{R}^m$. Sa deuxième dérivée est une application

$$D^2 f := D(Df) : X \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)) \ .$$

On définit par récurrence la dérivabilité (totale) d'ordre supérieur. Il est clair que

SCOLIE *f est k-fois continûment (totalement) dérivable si, et seulement si, f est k-fois continûment partiellement dérivable, i.e. si toutes les dérivées partielles d'ordre $\leqslant k$ sont continues.*

DEFINITION 2 Si $\alpha \in \mathbb{N}^n$ est un multi-indice, on pose

$$\alpha! := \alpha_1! \cdot \alpha_2! \cdot \ldots \cdot \alpha_n! \ .$$

PROPOSITION *Soient $f : X \longrightarrow \mathbb{R}$ une fonction k-fois continûment dérivable, $x \in X$ et $v \in \mathbb{R}^n \setminus \{0\}$. Alors la fonction*

$$g : t \longmapsto f(x + t \cdot v) \ ,$$

définie dans un voisinage de 0 *, est* k*-fois continûment dérivable et on a*

$$g^{(k)}(t) = \sum_{j_1,\ldots,j_k=1}^{n} \partial_{j_k} \ldots \partial_{j_1} f(x+t\cdot v)\cdot v_{j_k}\cdot\ldots\cdot v_{j_1} = \sum_{|\alpha|_1=k} \frac{k!}{\alpha!}\cdot\partial^\alpha f(x+t\cdot v)\cdot v^\alpha \ .$$

D'après la proposition 11.13, on a

$$g^{(1)}(t) = g'(t) = (\operatorname{grad} f(x+t\cdot v)|\, v) = \sum_{j_1=1}^{n} \partial_{j_1} f(x+t\cdot v)\cdot v_{j_1} = \sum_{|\alpha|_1=1} \partial^\alpha f(x+t\cdot v)\cdot v^\alpha \ ,$$

et par suite

$$g^{(2)}(t) = \sum_{j_1=1}^{n}\left(\sum_{j_2=1}^{n} \partial_{j_2}\partial_{j_1} f(x+t\cdot v)\cdot v_{j_2}\right)\cdot v_{j_1} \ .$$

Par récurrence on a donc

$$g^{(k)}(t) = \sum_{j_1,\ldots,j_k=1}^{n} \partial_{j_k}\ldots\partial_{j_1} f(x+t\cdot v)\cdot v_{j_k}\cdot\ldots\cdot v_{j_1} \ ,$$

et chaque dérivée partielle intervenant est d'ordre k .

Etant donné $\alpha \in \mathbb{N}^n$ tel que $|\alpha|_1 = k$, on doit déterminer le nombre $N(\alpha)$ de k-uples $(j_1,\ldots,j_k)$ tels que 1 apparaisse α_1 fois, .. et n apparaisse α_n fois, i.e. tels que

$$\partial_{j_k}\ldots\partial_{j_1} f = \partial^\alpha f$$

(théorème de Schwarz 11.6). Comme il y a $\alpha_1!$ possibilités de rendre les 1 distinguables les uns des autres, etc... , et $k!$ possibilités de permuter k objets (cf. 3.10), on obtient

$$N(\alpha)\cdot\alpha_1!\cdot\ldots\cdot\alpha_n! = k! \ .$$

Ainsi

$$N(\alpha) = \frac{k!}{\alpha!} \ ,$$

d'où le résultat. □

THEOREME *Soient* $f : X \longrightarrow \mathbb{R}$ *est une fonction* $(k+1)$*-fois continûment dérivable et* x, y *des points de* X *tels que le segment les joignant soit contenu dans* X *. Alors*

$$f(y) = \sum_{|\alpha|_1\leqslant k} \frac{1}{\alpha!}\cdot\partial^\alpha f(x)\cdot(y-x)^\alpha + \sum_{|\alpha|_1=k+1} \frac{1}{\alpha!}\cdot\partial^\alpha f(x+\theta\cdot[y-x])\cdot(y-x)^\alpha$$

pour un certain $\theta \in [0,1]$ *.*

Considérons la fonction g de la proposition précédente avec $v := y-x$. Utilisant la formule de Taylor avec reste de Lagrange 8.9, il existe $\theta \in [0,1]$ tel que

$$g(1) = \sum_{l=0}^{k} \frac{g^{(l)}(0)}{l!} + \frac{g^{(k+1)}(\theta)}{(k+1)!} \ .$$

Mais comme $f(y) = g(1)$,

$$g^{(l)}(0) = \sum_{|\alpha|_1=l} \frac{l!}{\alpha!}\cdot\partial^\alpha f(x)\cdot(y-x)^\alpha$$

et

$$g^{(k+1)}(\theta) = \sum_{|\alpha|_1 = k+1} \frac{(k+1)!}{\alpha!} \cdot \partial^\alpha f(x + \theta \cdot [y-x]) \cdot (y-x)^\alpha \ ,$$

le résultat en découle immédiatement. □

COROLLAIRE *Soient $f : X \longrightarrow \mathbb{R}$ une fonction k-fois continûment dérivable et $x \in X$. La fonction $\varphi : X \longrightarrow \mathbb{R}$ définie par*

$$f(y) = \sum_{|\alpha|_1 \leqslant k} \frac{1}{\alpha!} \cdot \partial^\alpha f(x) \cdot (y-x)^\alpha + \varphi(y) \quad \textit{pour tout } y \in X$$

satisfait à

$$\lim_{x \neq y \to x} \frac{\varphi(y)}{|y-x|^k} = 0 \ .$$

Pour tout y suffisamment proche de x , i.e. $y \in B(x,\varepsilon)$ pour $\varepsilon > 0$ suffisamment petit, il existe un $\theta(y) \in [0,1]$ tel que

$$f(y) = \sum_{|\alpha|_1 < k} \frac{1}{\alpha!} \cdot \partial^\alpha f(x) \cdot (y-x)^\alpha + \sum_{|\alpha|_1 = k} \frac{1}{\alpha!} \cdot \partial^\alpha f(x + \theta \cdot [y-x]) \cdot (y-x)^\alpha \ ,$$

donc

$$\varphi(y) = \sum_{|\alpha|_1 = k} \frac{1}{\alpha!} \cdot [\partial^\alpha f(x + \theta(y) \cdot [y-x]) - \partial^\alpha f(x)] \cdot (y-x)^\alpha \ .$$

Mais comme $|y_j - x_j| \leqslant |y-x|$ pour tout $j \in \{1,\ldots,n\}$, on a

$$|(y-x)^\alpha| \leqslant |y-x|^k \ ,$$

donc

$$\left| \frac{\varphi(y)}{|y-x|^k} \right| \leqslant \sum_{|\alpha|_1 = k} \frac{1}{\alpha!} \cdot |\partial^\alpha f(x + \theta(y) \cdot [y-x]) - \partial^\alpha f(x)| \ .$$

Le résultat en découle puisque les dérivées partielles sont continues et que

$$|\theta(y) \cdot (y-x)| \leqslant |y-x|$$

tend vers 0 lorsque y tend vers x . □

EXEMPLE (Cas de l'ordre 2) Si $f : X \longrightarrow \mathbb{R}$ est deux fois continûment dérivable et $x \in X$, en écrivant pour tout $y \in X$

$$f(y) = f(x) + (\operatorname{grad} f(x) | \, y-x) + \frac{1}{2} \cdot (y-x | \operatorname{Hess} f(x)(y-x)) + \varphi(y) \ ,$$

ou bien

$$f(y) = f(x) + \sum_{j=1}^n \partial_j f(x) \cdot (y_j - x_j) + \frac{1}{2} \cdot \sum_{k,l=1}^n \partial_l \partial_k f(x) \cdot (y_l - x_l) \cdot (y_k - x_k) + \varphi(y) \ ,$$

on a

$$\lim_{x \neq y \to x} \frac{\varphi(y)}{|y-x|^2} = 0 \ .$$

C'est immédiat, en remarquant que les multi-indices α tels que $|\alpha|_1 = 2$ sont soit de la forme $2 \cdot e_j$ pour un $j \in \{1, \ldots, n\}$ et $(2 \cdot e_j)! = 2$, soit de la forme $e_k + e_l$ pour $k, l \in \{1, \ldots, n\}$ tels que $k \neq l$ et $(e_k + e_l)! = 1$. Mais comme $\partial_l \partial_k f(x) = \partial_k \partial_l f(x)$ sont associés au même multi-indice $e_k + e_l = e_l + e_k$, on obtient le facteur $\frac{1}{2}$. On peut aussi reprendre la démonstration du théorème! $\square$

EXERCICE Déterminer le développement limité de Taylor à l'ordre 3 de la fonction

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto x^2 \cdot y + 2 \cdot y^3 + y^2 \cdot \cos x$$

au point $(1, 2)$.

11.16 Discussion locale d'une fonction

DEFINITION 1 Soient Y une partie quelconque de $\mathbb{R}^n$, $f : Y \longrightarrow \mathbb{R}$ une fonction et $x \in Y$. Nous dirons que f possède un *maximum* , respectivement *minimum local en* x , s'il existe un voisinage V de x (dans Y) tel que l'on ait

$$f(y) \leqslant f(x) \quad \text{, resp.} \quad f(y) \geqslant f(x) \quad \text{, pour tout } y \in V \ .$$

Pour simplifier on parle d'un *extremum local* si l'on ne veut pas préciser. Si l'on a

$$f(y) < f(x) \quad \text{, resp.} \quad f(y) > f(x) \quad \text{, pour tout } y \in V \smallsetminus \{x\} \ ,$$

on dit que ce minimum, respectivement maximum, est *strict* ou *isolé* .

PROPOSITION *Si $f : X \longrightarrow \mathbb{R}$ est partiellement dérivable et possède en $x \in X$ un minimum ou un maximum local, alors*

$$Df(x) = 0 \quad \textit{, i.e.} \quad \operatorname{grad} f(x) = 0 \ .$$

En effet, pour tout $j \in \{1, \ldots, n\}$, la fonction

$$t \longmapsto f(x + t \cdot e_j)$$

possède en 0 un minimum ou un maximum local. On a donc $\partial_j f(x) = 0$ par la proposition 8.4, d'où le résultat. □

DEFINITION 2 Si $f : X \longrightarrow \mathbb{R}$ est partiellement dérivable, nous dirons que $x \in X$ est un *point critique* de f si $\operatorname{grad} f(x) = 0$.

Ainsi tout point où f possède un extremum local est un point critique de f . Mais comme nous l'avons vu dans la remarque 8.4.2, la réciproque est déjà fausse en dimension 1 !

DEFINITION 3 Soit $S \in \mathcal{L}(\mathbb{R}^n) \approx \mathbb{M}_{\mathbb{R}}(n \times n)$ une application linéaire ou une matrice symétrique, donc telle que $S^{\intercal} = S$. Nous dirons qu'elle est *définie positive* , respectivement *définie négative* si, pour tout $v \in \mathbb{R}^n \smallsetminus \{0\}$, on a

$$(v|\, Sv) > 0 \quad \text{, resp.} \quad (v|\, Sv) < 0 \ .$$

On dit qu'elle est *(semidéfinie) positive* , respectivement *(semidéfinie) négative* si, pour tout $v \in \mathbb{R}^n$, on a

$$(v|\, Sv) \geqslant 0 \quad \text{, resp.} \quad (v|\, Sv) \leqslant 0 \ .$$

On dit qu'elle est *indéfinie* s'il existe $u, v \in \mathbb{R}^n$ tels que

$$(u|\, Su) > 0 \quad \text{et} \quad (v|\, Sv) < 0 \ .$$

REMARQUE Comme S est symétrique, il existe une base orthonormée $(\epsilon_j)_{j=0,\ldots,n}$ de $\mathbb{R}^n$ diagonalisant S , i.e. formée de vecteurs propres, dont les valeurs propres correspondantes sont $(\lambda_j)_{j=1,\ldots,n}$. En décomposant $v \in \mathbb{R}^n$ dans cette base, i.e.

$$v = \sum_{j=1}^{n} \tilde{v}_j \cdot \epsilon_j \ ,$$

on a

$$(v|\, Sv) = \sum_{j=1}^{n} \lambda_j \cdot \tilde{v}_j^2 \ .$$

On en déduit que S est définie positive, resp. définie négative, semidéfinie positive, semidéfinie négative si, et seulement si, pour tout $j \in \{1,\ldots,n\}$, on a

$$\lambda_j > 0 \quad \text{, resp. } < 0 \ , \ \geqslant 0 \ , \ \leqslant 0 \ .$$

Elle est indéfinie si, et seulement si, il existe $k, l \in \{1,\ldots,n\}$ tels que

$$\lambda_k > 0 \quad \text{et} \quad \lambda_l < 0 \ .$$

THEOREME *Soient $f : X \longrightarrow \mathbb{R}$ une fonction deux fois continûment dérivable et $x \in X$. Si $\operatorname{grad} f(x) = 0$ et $\operatorname{Hess} f(x)$ est définie positive, respectivement définie négative, alors f possède un minimum, respectivement un maximum local strict en x . Si $\operatorname{Hess} f(x)$ est indéfinie, alors f ne possède ni un minimum local, ni un maximum local.*

D'après l'exemple 11.15 en écrivant

$$f(x+v) = f(x) + \frac{1}{2} \cdot (v|\operatorname{Hess} f(x)\, v) + \varphi(v) \ ,$$

pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que l'on ait

$$\left| \frac{\varphi(v)}{|v|^2} \right| \leqslant \varepsilon \quad \text{si } |v| \leqslant \delta \ .$$

Comme $\operatorname{Hess} f(x)$ est définie positive, on a $\lambda := \min_{j=1,\ldots,n} \lambda_j > 0$ et, dans la base orthonormée diagonalisant $\operatorname{Hess} f(x)$, il vient

$$\frac{1}{2} \cdot (v|\operatorname{Hess} f(x)\, v) = \frac{1}{2} \cdot \sum_{j=1}^{n} \lambda_j \cdot v_j^2 \geqslant \frac{\lambda}{2} \cdot \sum_{j=1}^{n} v_j^2 = \frac{\lambda}{2} \cdot |v|^2 \ .$$

On a alors

$$f(x+v) - f(x) = \frac{1}{2} \cdot (v|\operatorname{Hess} f(x)\, v) + \varphi(v) \geqslant \frac{\lambda}{2} \cdot |v|^2 - |\varphi(v)| \geqslant \left(\frac{\lambda}{2} - \varepsilon \right) \cdot |v|^2 \geqslant \frac{\lambda}{4} \cdot |v|^2 > 0 \ ,$$

si $0 < |v| \leqslant \delta$ en ayant choisi $\varepsilon := \frac{\lambda}{4}$. Ceci montre que f possède un minimum local strict en x . On applique ce résultat à $-f$ si $\operatorname{Hess} f(x)$ est définie négative.

Si $\operatorname{Hess} f(x)$ est indéfinie, il existe $u, v \in \mathbb{R}^n$ tels que

$$(u|\operatorname{Hess} f(x)\, u) > 0 \quad \text{et} \quad (v|\operatorname{Hess} f(x)\, v) < 0 \ .$$

On a alors

$$f(x + t \cdot u) - f(x) = \frac{t^2}{2} \cdot (u|\operatorname{Hess} f(x)\, u) + \varphi(t \cdot u) \geqslant \frac{t^2}{2} \cdot (u|\operatorname{Hess} f(x)\, u) - |\varphi(tu)| \geqslant$$

$$\geqslant \left(\frac{(u|\,\mathrm{Hess}\, f(x)\, u)}{2} - \varepsilon \cdot |u|^2 \right) \cdot t^2 = \frac{(u|\,\mathrm{Hess}\, f(x)\, u)}{4} \cdot t^2 > 0 \ ,$$

si $0 < |t| \leqslant \delta$ en ayant choisi $\varepsilon := \frac{(u|\,\mathrm{Hess}\, f(x) u)}{4 \cdot |u|^2}$. On a de même

$$f(x + t \cdot v) - f(x) = \frac{t^2}{2} \cdot (v|\,\mathrm{Hess}\, f(x)\, v) + \varphi(t \cdot v) \leqslant \frac{t^2}{2} \cdot (v|\,\mathrm{Hess}\, f(x)\, v) + |\varphi(tv)| \leqslant$$

$$\leqslant \left(\frac{(v|\,\mathrm{Hess}\, f(x)\, v)}{2} + \varepsilon \cdot |v|^2 \right) \cdot t^2 = \frac{(v|\,\mathrm{Hess}\, f(x)\, v)}{4} \cdot t^2 < 0 \ ,$$

si $0 < |t| \leqslant \delta$ en ayant choisi $\varepsilon := -\frac{(v|\,\mathrm{Hess}\, f(x) v)}{4 \cdot |v|^2}$. □

EXEMPLE 1 Etant donné $a \in \mathbb{R}$, $b \in \mathbb{R}^n$ et $S \in \mathcal{L}(\mathbb{R}^n)$ une application linéaire symétrique, on considère la fonction

$$f : x \longmapsto a + (b|\, x) + (x|\, Sx) : \mathbb{R}^n \longrightarrow \mathbb{R} \ .$$

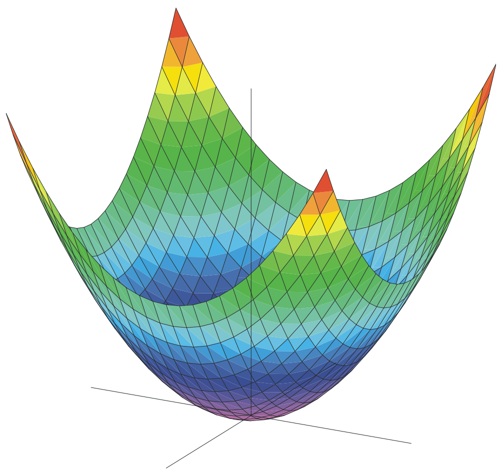
D'après l'exemple 11.11.3, on a

$$\mathrm{grad}\, f(x) = b + 2 \cdot Sx \ .$$

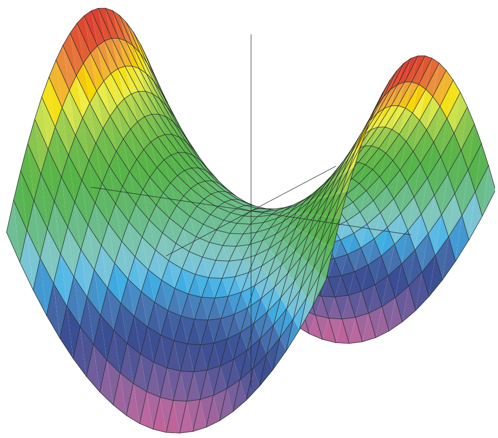
Ainsi $\mathrm{grad}\, f(x) = 0$ si, et seulement si, on a $b + 2 \cdot Sx = 0$. En admettant que S est inversible, ce qui signifie que $\det S \neq 0$ ou que toutes ses valeurs propres sont $\neq 0$, on voit que le seul point où $\mathrm{grad}\, f$ s'annule est

$$-\frac{1}{2} \cdot \overset{-1}{S} b \ .$$

L'interprétation géométrique de cet exemple en liaison avec le théorème est alors simple si l'on se place dans la base orthonormée diagonalisant S .



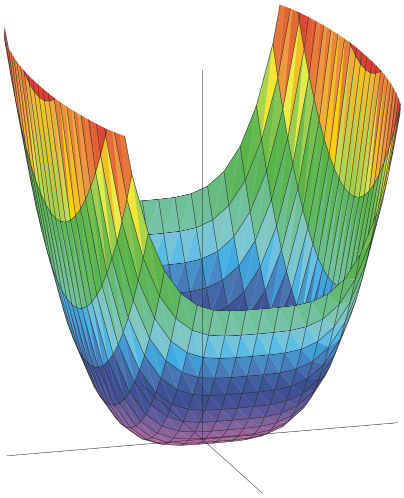
$(x, y) \longmapsto x^2 + y^2$



$(x, y) \longmapsto x^2 - y^2$

EXEMPLE 2 Soit

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto x^2 + y^4 \ .$$



$(x, y) \longmapsto x^2 + y^4$

On a

$$\operatorname{grad} f(x, y) = \left(2 \cdot x, 4 \cdot y^3\right)^{\intercal} \ ,$$

donc $\operatorname{grad} f$ ne s'annule qu'en $(0, 0)$. D'autre part

$$\operatorname{Hess} f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 12 \cdot y^2 \end{pmatrix} \ ,$$

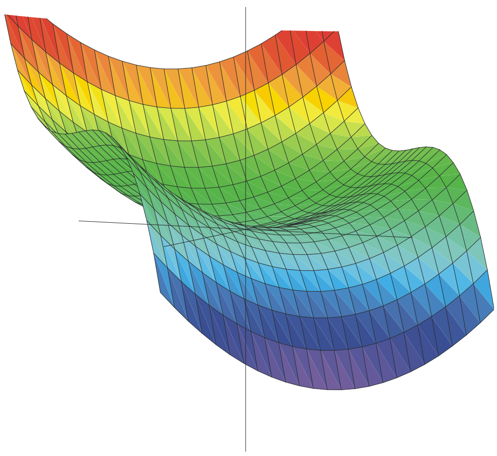
donc

$$\operatorname{Hess} f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

est semidéfinie positive. Comme $x^2 + y^4 > 0$ si $(x, y) \neq (0, 0)$, on voit que $(0, 0)$ est un minimum absolu strict.

EXEMPLE 3 Soit

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto x^2 + y^3 \ .$$



$(x, y) \longmapsto x^2 + y^3$

On a

$$\operatorname{grad} f(x,y) = \left(2\cdot x, 3\cdot y^2\right)^{\intercal} \ ,$$

donc $\operatorname{grad} f$ ne s'annule qu'en $(0,0)$. D'autre part

$$\operatorname{Hess} f(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 6\cdot y \end{pmatrix} \ ,$$

donc

$$\operatorname{Hess} f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

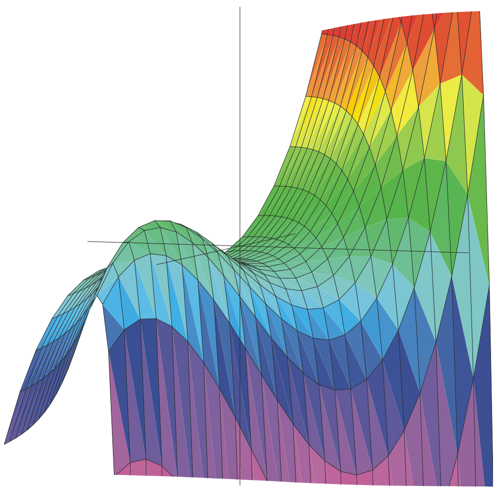
est semidéfinie positive, mais

$$f(x,0) > 0 \text{ si } x \neq 0 \quad \text{et} \quad f(0,y) < 0 \text{ si } y < 0 \ .$$

EXERCICE 1 Montrer que la fonction

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x,y) \longmapsto 3x\cdot e^y - x^3 - e^{3y}$$

est surjective et déterminer ses extrema locaux.

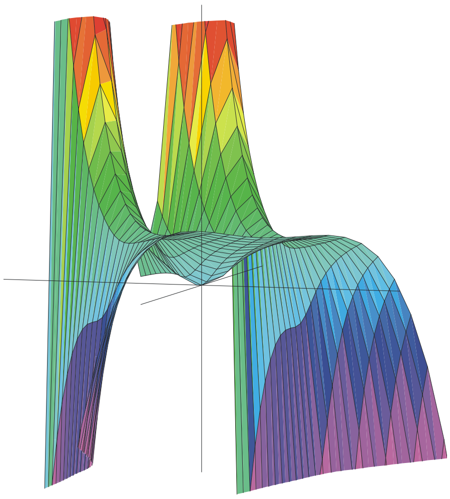


Cet exemple montre que le résultat de l'exercice 8.5.3 ne se généralise pas sans hypothèse supplémentaire à la dimension 2 . Il est tiré de l'article de Ira Rosenholtz et Lowell Smylie [4]. En même temps J. Marshall Ash et Harlan Sexton [5] ont publié avec une légère modification l'exemple suivant :

$$\mathbb{R}^2 \longrightarrow \mathbb{R} : (x,y) \longmapsto \frac{x^2}{1+x^2} + \left(2y^2 - y^4\right)\cdot\left(e^x + \frac{1}{1+x^2}\right) \ .$$

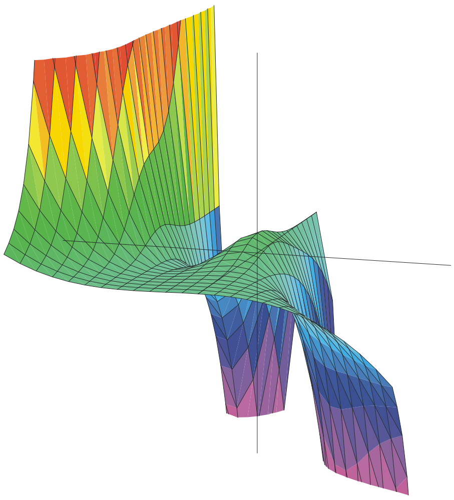
[4] I. Rosenholtz, L. Smylie, *"The only critical point in town" test,* Math. Mag. 58 (1985), p. 149-150.

[5] J.M. Ash, H. Sexton, *A surface with one local minimum,* Math. Mag. 58 (1985), p. 147-149.



Ces auteurs donnent également une condition suffisante pour que le théorème reste vrai. Un autre exemple

$$\mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto \left(2x^3 - 3x^2\right) \cdot e^{-y} + \left(2x^3 - 3x^2 + 1\right) \cdot e^{-y^2}$$



est dû à David A. Smith [6] et a été publié dans le livre de Philip Gillett [7].

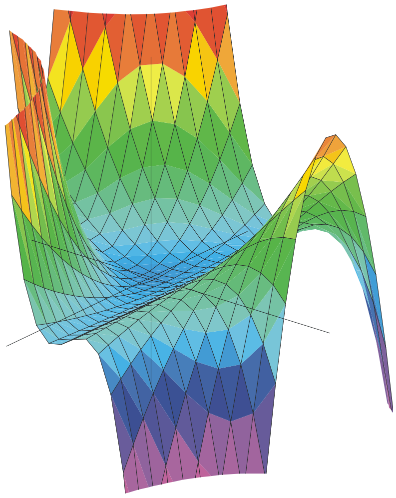
Mais l'exemple le plus simple est celui de Bruce Calvert et M.K. Vamanamurthy [8].

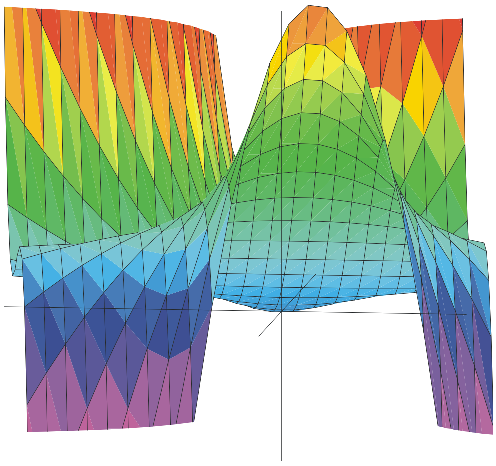
$$\mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto x^2 \cdot (1 + y)^3 + y^2$$

[6] D.A. Smith, *Three observations on a theme : Editorial note,* Math. Mag. 58 (1985), p. 146.

[7] P. Gillett, *Calculus and analytic geometry,* 2nd ed., D.C. Heath, Lexington, Mass., 1984.

[8] B. Calvert, M.K. Vamanamurthy, *Local and global extrema for functions of several variables,* J. Austral. Math. Soc. 29 (1980), p. 362-368.





Ils ont aussi donnés des conditions suffisantes.

EXERCICE 2 Droite de régression par la méthode des moindres carrés de Gauß.

Soit $(x_j, y_j)_{j=1,\ldots,n}$ une suite finie de $\mathbb{R}^2$ et posons

$$\overline{x} := \frac{1}{n} \cdot \sum_{j=1}^{n} x_j \quad , \quad \overline{y} := \frac{1}{n} \cdot \sum_{j=1}^{n} y_j \ .$$

Montrer que la fonction

$$(m, b) \longmapsto \sum_{j=1}^{n} (y_j - m \cdot x_j - b)^2 : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

possède un minimum en

$$m = \frac{\sum_{j=1}^{n} (x_j - \overline{x}) \cdot (y_j - \overline{y})}{\sum_{j=1}^{n} (x_j - \overline{x})^2} \quad \text{et} \quad b = \overline{y} - m \cdot \overline{x} \ .$$

De même

$$m \longmapsto \sum_{j=1}^{n} (y_j - m \cdot x_j)^2 : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

possède un minimum en

$$m = \frac{\sum_{j=1}^{n} x_j \cdot y_j}{\sum_{j=1}^{n} x_j^2} \ .$$

11.17 Minimum et maximum avec condition

DEFINITION Soient $f : X \longrightarrow \mathbb{R}$, $F : X \longrightarrow \mathbb{R}^m$ des fonctions et $\xi \in X$ tel que $F(\xi) = 0$. Nous dirons que f possède un *maximum* , respectivement un *minimum local* en ξ *sous la condition* $F = 0$, si la restriction de f à l'ensemble

$$\{F = 0\}$$

possède un maximum, respectivement un minimum local en ξ . Pour simplifier on parle d'un *extremum local avec condition* si l'on ne veut pas préciser.

Pour démontrer le lemme qui suit nous aurons besoin du théorème de la fonction réciproque que nous démontrerons plus tard en 13.2.

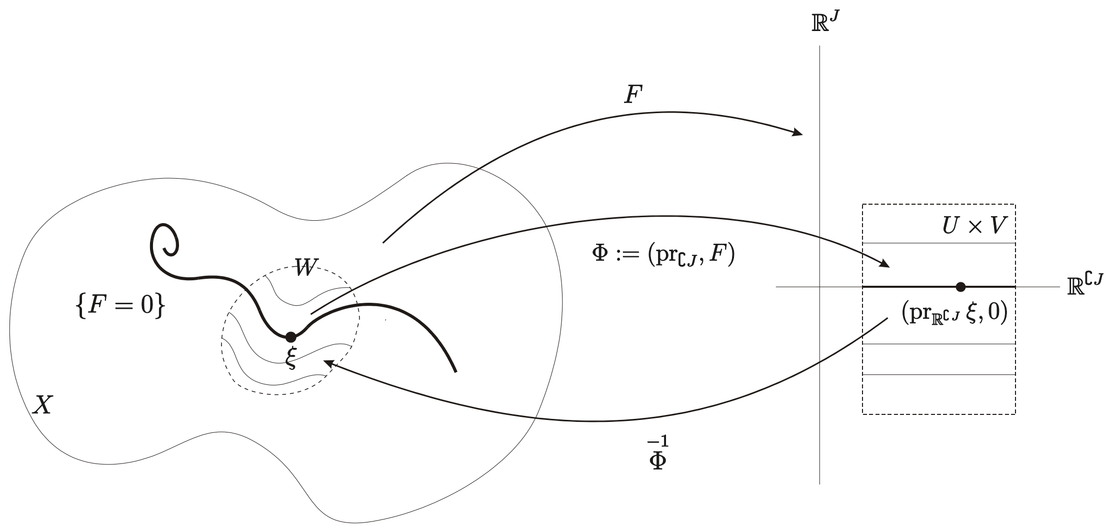
LEMME *Si $F : X \longrightarrow \mathbb{R}^m$ est continûment dérivable, $F(\xi) = 0$ et $DF(\xi) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ est surjective, alors il existe une partie à m éléments $J \subset \{1, \ldots, n\}$, un voisinage ouvert W de ξ dans X et dans la décomposition $\mathbb{R}^n = \mathbb{R}^{\complement J} \times \mathbb{R}^J$ un voisinage ouvert de $(\mathrm{pr}_{\complement J}\, \xi, 0)$ de la forme $U \times V$ tels que l'application*

$$\Phi := (\mathrm{pr}_{\complement J}, F) : X \longrightarrow \mathbb{R}^{\complement J} \times \mathbb{R}^J = \mathbb{R}^n : x \longmapsto (\mathrm{pr}_{\mathbb{R}^{\complement J}}\, x, F(x))$$

soit un difféomorphisme de W sur $U \times V$. En outre

$$\Phi(\{F = h\} \cap W) = U \times \{h\}$$

pour tout $h \in V$.



Puisque $DF(\xi)$ est surjective, $\operatorname{Ker} DF(\xi)$ est de dimension $n - m$. Il existe donc au moins m vecteurs de la base canonique de $\mathbb{R}^n$ qui ne sont pas dans ce sous-espace vectoriel. Soit J l'ensemble des indices de ces vecteurs. Utilisant la décomposition $\mathbb{R}^n = \mathbb{R}^{\complement J} \times \mathbb{R}^J$, on a $\operatorname{Ker} DF(\xi) \cap \left[\{0\} \times \mathbb{R}^J\right] = \{0\}$; l'application linéaire $D_2F(\xi) : \mathbb{R}^J \longrightarrow \mathbb{R}^J$, qui s'identifie à la restriction de $DF(\xi)$ à $\{0\} \times \mathbb{R}^J \approx \mathbb{R}^J$, n'a donc pas de noyau. Elle est donc inversible. La dérivée en ξ de l'application continûment dérivable Φ est

$$D\Phi(\xi) = \begin{pmatrix} \mathrm{Id} & 0 \\ D_1 F(\xi) & D_2 F(\xi) \end{pmatrix} ,$$

donc inversible, puisque $\det D\Phi(\xi) = \det D_2 F(\xi) \neq 0$.

Le théorème de la fonction réciproque 13.2 montre alors qu'il existe un voisinage ouvert W de ξ dans X tel que Φ soit un difféomorphisme de W sur le voisinage ouvert $\Phi(W)$ de $(\mathrm{pr}_{\mathbb{R}^{\complement J}} \xi, 0)$. En choisissant un voisinage ouvert de de la forme $U \times V$, il nous suffit de remplacer W par $\overset{-1}{\Phi}(U \times V)$. Finalement $x \in \{F = h\} \cap W$ est équivalent à $F(x) = h$ et $\Phi(x) \in U \times V$, donc à $\Phi(x) \in U \times \{h\}$. □

THEOREME (Multiplicateur de Lagrange) *On suppose que $f : X \longrightarrow \mathbb{R}$ et $F : X \longrightarrow \mathbb{R}^m$ sont des fonctions continûment dérivables sur X et que $\xi \in X$ est tel que $F(\xi) = 0$ et $DF(\xi) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ soit surjective.*

*Si f possède un extremum local en ξ sous la condition $F = 0$, alors il existe $\lambda = (\lambda_k)_{k=1,\ldots,m} \in \mathbb{R}^m$ (on dit que les λ_k sont les **multiplicateurs de Lagrange**) tel que*

$$\operatorname{grad} f(\xi) = DF(\xi)^{\intercal} \lambda ,$$

i.e.

$$\partial_l f(\xi) = \sum_{k=1}^{m} \partial_l F_k(\xi) \cdot \lambda_k \qquad \textit{pour } l = 1, \ldots, n .$$

Si $m = 1$ cela signifie que

$$\operatorname{grad} f(\xi) = \lambda \cdot \operatorname{grad} F(\xi) .$$

D'après le lemme et l'hypothèse la fonction $f \circ \overset{-1}{\Phi}$ possède un extremum local en $(\eta, 0) := (\mathrm{pr}_{\mathbb{R}^{\complement J}} \xi, 0)$ sous la condition $\mathrm{pr}_{\mathbb{R}^J} = 0$, i.e. $f \circ \overset{-1}{\Phi} \circ j_{\mathbb{R}^{\complement J}}$ possède un extremum local en η , en désignant par $j_{\mathbb{R}^{\complement J}}$ l'injection canonique de $\mathbb{R}^{\complement J}$ dans $\mathbb{R}^n = \mathbb{R}^{\complement J} \times \mathbb{R}^J$. La proposition 11.16 montre alors que

$$0 = D\left(f \circ \overset{-1}{\Phi} \circ j_{\mathbb{R}^{\complement J}}\right)(\eta) = D\left(f \circ \overset{-1}{\Phi}\right)(\eta, 0)\, D\begin{pmatrix} \mathrm{Id} \\ 0 \end{pmatrix}(\eta) =$$

$$= \left(D_1\left(f \circ \overset{-1}{\Phi}\right)(\eta, 0), D_2\left(f \circ \overset{-1}{\Phi}\right)(\eta, 0)\right)\begin{pmatrix} \mathrm{Id} \\ 0 \end{pmatrix} = D_1\left(f \circ \overset{-1}{\Phi}\right)(\eta, 0) .$$

Mais comme

$$Df(\xi) = D\left(f \circ \overset{-1}{\Phi} \circ \Phi\right)(\xi) = D\left(f \circ \overset{-1}{\Phi}\right)(\eta, 0)\, D\Phi(\xi) =$$

$$= \left(D_1\left(f \circ \overset{-1}{\Phi}\right)(\eta, 0), D_2\left(f \circ \overset{-1}{\Phi}\right)(\eta, 0)\right)\begin{pmatrix} D\,\mathrm{pr}_{\mathbb{R}^{\complement J}}(\xi) \\ DF(\xi) \end{pmatrix} =$$

$$= \left(0, D_2\left(f \circ \overset{-1}{\Phi}\right)(\eta, 0)\right)\begin{pmatrix} \mathrm{pr}_{\mathbb{R}^{\complement J}} \\ DF(\xi) \end{pmatrix} =$$

$$= D_2\left(f \circ \overset{-1}{\Phi}\right)(\eta, 0)\, DF(\xi) ,$$

on obtient

$$\operatorname{grad} f(\xi) = Df(\xi)^\intercal = \left(D_2 \left(f \circ \overset{-1}{\Phi} \right) (\eta, 0)\, DF(\xi) \right)^\intercal = DF(\xi)^\intercal\, \lambda$$

en ayant posé $\lambda := \left(D_2 \left(f \circ \overset{-1}{\Phi} \right) (\eta, 0) \right)^\intercal$. $\square$

REMARQUE 1 Cette méthode permet souvent de déterminer les minima et maxima locaux avec condition. Mais il faut tout d'abord s'assurer de leur existence, en général à l'aide d'un argument de compacité. Elle est surtout utile lorsqu'il n'est pas possible de paramétrer simplement l'ensemble $\{F = 0\}$.

EXEMPLE 1 On va déterminer le rapport entre le rayon ρ et la hauteur κ d'une boîte cylindrique d'un certain volume $V > 0$ de telle manière que la surface du fer blanc nécessaire soit minimale. Cet exemple simple montre aussi comment on prouve l'existence d'un minimum avec condition.

La condition $F = 0$ sur le volume est décrite par la fonction

$$F : (r, h) \longmapsto \pi \cdot r^2 \cdot h - V : \mathbb{R}^2 \longrightarrow \mathbb{R} \ ,$$

tandis que la surface l'est par

$$f : (r, h) \longmapsto 2\pi \cdot r \cdot h + 2\pi \cdot r^2 : \mathbb{R}^2 \longrightarrow \mathbb{R} \ .$$

Ces fonctions sont évidemment continûment dérivables. On en déduit que la partie $\{F = 0\}$ est fermée dans $\mathbb{R}^2$. Il en est de même, pour tout $M \in \mathbb{R}_+$, de l'ensemble $\{f \leqslant M\}$, donc de

$$\{F = 0\} \cap \{f \leqslant M\} \cap \mathbb{R}_+^2 \ .$$

Cet ensemble est également borné car, pour tout $(r, h) \in \mathbb{R}_+^2$ tel que

$$\pi \cdot r^2 \cdot h = V \quad \text{et} \quad 2\pi \cdot r \cdot h + 2\pi \cdot r^2 \leqslant M \ ,$$

on obtient

$$2\pi \cdot r^2 \leqslant M \quad \text{et} \quad 4\pi \cdot V \cdot h = 4\pi^2 \cdot r^2 \cdot h^2 \leqslant M^2 \ .$$

Cet ensemble est donc compact, et non-vide si $M \geqslant 2 \cdot V + 2\pi$, puisqu'il contient alors le point $\left(1, \frac{V}{\pi}\right)$. La fonction f atteint donc son minimum sur $\{F = 0\} \cap \{f \leqslant M\} \cap \mathbb{R}_+^2$ en un point (ρ, κ) , mais en ce point elle atteint aussi son minimum sur $\{F = 0\} \cap \mathbb{R}_+^2 \subset \mathbb{R}_+^{*2}$.

On peut donc raisonner dans l'ouvert $\mathbb{R}_+^{*2}$. On a

$$\operatorname{grad} f(r, h) = \begin{pmatrix} 2\pi \cdot h + 4\pi \cdot r \\ 2\pi \cdot r \end{pmatrix}$$

et

$$\operatorname{grad} F(r, h) = \begin{pmatrix} 2\pi \cdot r \cdot h \\ \pi \cdot r^2 \end{pmatrix} \neq (0, 0) \quad \text{si } r \neq 0 \ .$$

Il existe donc $\lambda \in \mathbb{R}$ tel que

$$\begin{pmatrix} 2\pi \cdot \kappa + 4\pi \cdot \rho \\ 2\pi \cdot \rho \end{pmatrix} = \lambda \cdot \begin{pmatrix} 2\pi \cdot \rho \cdot \kappa \\ \pi \cdot \rho^2 \end{pmatrix} \ .$$

On en déduit que $2 = \lambda \cdot \rho$, et par suite que $\kappa + 2 \cdot \rho = 2 \cdot \kappa$, donc que

$$\kappa = 2 \cdot \rho \ .$$

En remplaçant dans la condition, il vient

$$\rho^3 = \frac{V}{2\pi} \quad \text{et} \quad \kappa^3 = \frac{4 \cdot V}{\pi} \ .$$

Remarquons que cette méthode n'est pas la meilleure pour ce problème, puisque l'on peut résoudre la condition par rapport à l'une des variables, puis en substituant, se ramener à un problème à une variable.

EXEMPLE 2 Soient $\eta \in \mathbb{R}^n$ et A une partie non-vide de $\mathbb{R}^n$ décrite par l'équation $F = 0$, i.e.

$$A = \{F = 0\} \quad \text{avec } F : \mathbb{R}^n \longrightarrow \mathbb{R} \text{ continue} \ .$$

On se propose de montrer que la distance de η à A , i.e.

$$d(\eta, A) := \inf_{x \in A} d(\eta, x) \ ,$$

est atteinte en un point $\xi \in A$ et de la calculer.

Il est préférable de considérer $d(\eta, A)^2$, ce qui revient à déterminer le minimum de la fonction continûment dérivable

$$f : x \longmapsto |x - \eta|^2 : \mathbb{R}^n \longrightarrow \mathbb{R}$$

sous la condition $F = 0$. Pour R assez grand, la partie

$$\left\{f \leqslant R^2\right\} \cap \{F = 0\} \subset B(\eta, R)$$

est bornée, et fermée puisque F est continue, donc compacte. Il est alors clair que f atteint son minimum sous la condition $F = 0$ en un point $\xi \in \{F = 0\}$.

Nous supposerons que F est continûment dérivable. On a

$$\operatorname{grad} f(x) = 2 \cdot (x - \eta) \ .$$

Si $\operatorname{grad} F(\xi) \neq 0$, il existe donc $\lambda \in \mathbb{R}$ tel que

$$2 \cdot (\xi - \eta) = \lambda \cdot \operatorname{grad} F(\xi) \ .$$

Considérons le cas particulier d'un hyperplan d'équation

$$F(x) = a + (b|\, x) = 0 \quad \text{avec } a \in \mathbb{R} \text{ et } b \in \mathbb{R}^n \text{ tel que } |b| = 1 \ .$$

Il vient $\operatorname{grad} F(x) = b \neq 0$, donc $2 \cdot (\xi - \eta) = \lambda \cdot b$. On en tire $\xi = \eta + \frac{\lambda}{2} \cdot b$, puis

$$a + \left(b \,\middle|\, \eta + \frac{\lambda}{2} \cdot b\right) = 0 \ ,$$

et par suite

$$\frac{\lambda}{2} = -\Big(a + (b|\, \eta)\Big) \ ,$$

donc finalement

$$\xi = \eta - \Big(a + (b|\, \eta)\Big) \cdot b \ .$$

EXEMPLE 3 Soient $S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ une application linéaire symétrique et

$$f : \mathbb{R}^n \longrightarrow \mathbb{R} : x \longmapsto (x|\, Sx) \ .$$

PROPOSITION *Il existe des vecteurs $\xi, \eta \in \mathbb{S}^{n-1}$, i.e. $|\xi| = |\eta| = 1$, où f atteint son maximum, respectivement minimum sur $\mathbb{S}^{n-1}$. Ce sont des vecteurs propres dont les valeurs propres correspondantes sont respectivement la plus grande et la plus petite valeur propre de S :*

$$\sup\ f\left(\mathbb{S}^{n-1}\right) = (\xi|\, S\xi) \quad \textit{resp.} \quad \inf\ f\left(\mathbb{S}^{n-1}\right) = (\eta|\, S\eta)\ .$$

En outre toutes les valeurs propres de S sont réelles et S est diagonalisable.

Nous devons chercher un maximum de la fonction continûment dérivable f sous la condition $|x|^2 - 1 = 0$. On a

$$\operatorname{grad} f(x) = 2 \cdot Sx$$

(cf. exemple 11.16.1). En considérant la fonction continûment dérivable

$$F : \mathbb{R}^n \longrightarrow \mathbb{R} : x \longmapsto |x|^2 - 1\ ,$$

on a

$$\operatorname{grad} F(x) = 2 \cdot x\ ,$$

donc $\operatorname{grad} F(x) \neq 0$ si $x \in \mathbb{S}^{n-1}$. Comme

$$f : \mathbb{S}^{n-1} \longrightarrow \mathbb{R}$$

est continue, et que $\mathbb{S}^{n-1}$ est une partie fermée bornée de $\mathbb{R}^n$, donc compacte (théorème de Heine-Borel 10.18), nous savons que f possède au moins un maximum $\xi \in \mathbb{S}^{n-1}$ (corollaire 10.19). Il existe donc un $\lambda \in \mathbb{R}$ tel que

$$2 \cdot S\xi = \operatorname{grad} f(\xi) = \lambda \cdot \operatorname{grad} F(\xi) = 2\lambda \cdot \xi\ .$$

Ceci montre que S possède au moins une valeur propre réelle λ de vecteur propre ξ . Si μ est une autre valeur propre réelle de S de vecteur propre η , on peut évidemment supposer que $|\eta| = 1$, alors

$$\mu = (\eta|\, S\eta) \leqslant \max\ f\left(\mathbb{S}^{n-1}\right) = f(\xi) = (\xi|\, S\xi) = \lambda\ ,$$

ce qui montre que λ est la plus grande valeur propre (réelle) de S . On raisonne de la même manière pour le minimum.

En décomposant $\mathbb{R}^n$ sous la forme

$$\mathbb{R}^n = \mathbb{R} \cdot \xi \oplus (\mathbb{R} \cdot \xi)^{\perp}$$

et en raisonnant par récurrence, on montre que S est diagonalisable et que toutes les valeurs propres sont réelles. □

COROLLAIRE *Si $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ est une application linéaire, alors $\|T\|^2$ est la plus grande valeur propre de $T^{\intercal}T$.*

En effet

$$\|T\|^2 = \sup_{v\in\mathbb{R}^n} |Tv|^2 = \sup_{v\in\mathbb{R}^n} (Tv|\, Tv) = \sup_{v\in\mathbb{R}^n} (v|\, T^{\intercal}Tv)$$

et $T^{\intercal}T$ est une application linéaire symétrique. □

REMARQUE 2 Ce corollaire montre que $\|T\|$ s'exprime difficilement à l'aide des coefficients de la matrice de T . Si

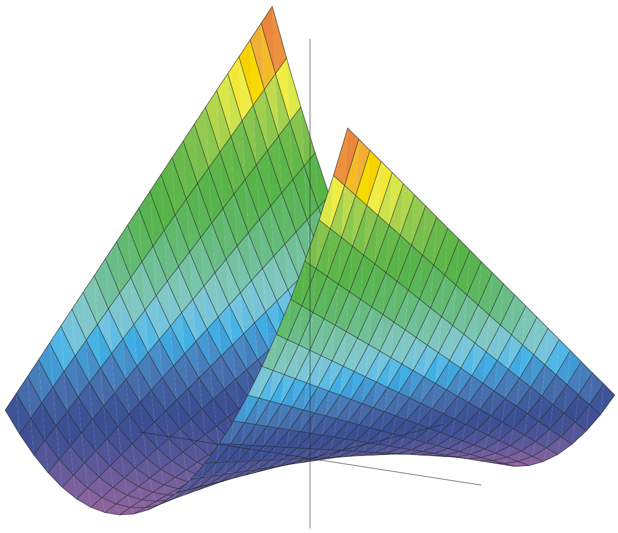
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_{\mathbb{R}}(2 \times 2)\ ,$$

alors

$$\left\|\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right\|^2 = \frac{1}{2}\cdot\left(a^2+b^2+c^2+d^2+\sqrt{\left[(a+d)^2+(b-c)^2\right]\left[(a-d)^2+(b+c)^2\right]}\right) .$$

EXERCICE 1 On considère la fonction

$$f:\mathbb{R}^2 \longrightarrow \mathbb{R} : (x,y) \longmapsto 4\cdot x^2 - 3\cdot x\cdot y .$$



$(x,y) \longmapsto 4\cdot x^2 - 3\cdot x\cdot y$

(a) Déterminer les minima et maxima locaux de f sur $\{(x,y)\in\mathbb{R}^2 \mid x^2+y^2<1\}$.

(b) Déterminer les minima et maxima locaux de f sur $\{(x,y)\in\mathbb{R}^2 \mid x^2+y^2=1\}$.

(c) Déterminer les minima et maxima locaux comme absolus de f sur

$$\left\{(x,y)\in\mathbb{R}^2 \;\middle|\; x^2+y^2\leqslant 1\right\} .$$

EXERCICE 2 Déterminer les extrema locaux et absolus de la fonction

$$f:\mathbb{R}^n \longrightarrow \mathbb{R} : (x_1,\ldots,x_n) \longmapsto \prod_{j=1}^n x_j^2$$

sous la condition

$$\sum_{j=1}^n x_j^2 = 1 .$$

En déduire que, pour tout $a_1,\ldots,a_n\in\mathbb{R}_+$, on a

$$\left(\prod_{j=1}^n x_j^2\right)^{\frac{1}{n}} \leqslant \frac{1}{n}\cdot\sum_{j=1}^n a_j .$$

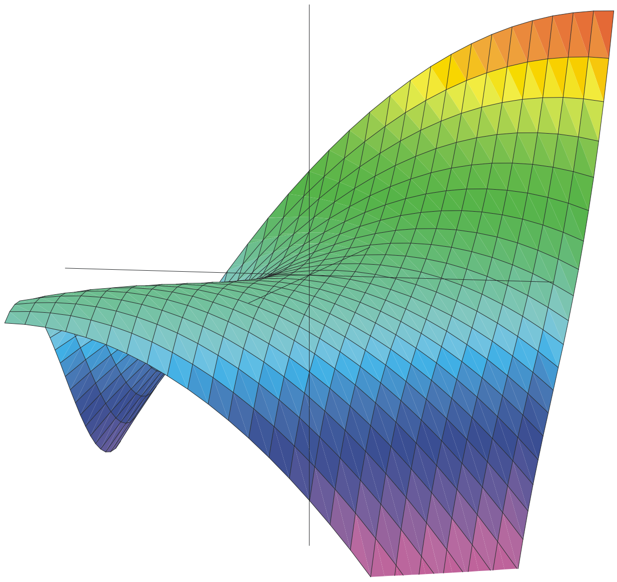
EXERCICE 3 On considère la fonction

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \longmapsto y^3 + y + 4 \cdot x \cdot y - 2 \cdot x^2 \ .$$

(a) Déterminer les minima et maxima locaux de f .

(b) Déterminer les extrema locaux comme absolus de f sur

$$\left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; x \leqslant 0 \text{ et } x + 2 \geqslant -y \geqslant 0 \right\} \ .$$



$(x, y) \longmapsto y^3 + y + 4 \cdot x \cdot y - 2 \cdot x^2$