

Chapitre XXVII

Fonctions de deux variables

On munit dans ce chapitre l'espace euclidien \mathbb{R}^2 de son produit scalaire canonique et de la norme euclidienne associée.

1. Topologie de \mathbb{R}^2

a) Ouverts du plan

Définition XXVII.1. Soit $x \in \mathbb{R}^2$ et $r \in \mathbb{R}_+^*$. On appelle :

— **disque ouvert** de centre x et rayon r l'ensemble :

$$D(x, r) = \{y \in \mathbb{R}^2 \mid \|x - y\| < r\};$$

— **disque fermé** de centre x et rayon r l'ensemble :

$$\bar{D}(x, r) = \{y \in \mathbb{R}^2 \mid \|x - y\| \leq r\}.$$

☞ **Remarque XXVII.1.** Le lecteur motivé pourra démontrer que tout disque de \mathbb{R}^2 est convexe, au sens de la caractérisation vue au corollaire III.12.a. Il existe cependant des convexes qui ne sont pas des disques : par exemple, un "segment "

$$\{(1 - \lambda)a + \lambda b \mid \lambda \in [0, 1]\},$$

pour $a, b \in \mathbb{R}^2$, reste convexe dans \mathbb{R}^2 sans être un disque.

Définition XXVII.2. Une partie $A \subset \mathbb{R}^2$ est dite :

— **ouverte** si :

$$\forall x \in A, \exists r \in \mathbb{R}_+^*, D(x, r) \subset A;$$

— **fermée** si son complémentaire $\mathbb{R}^2 \setminus A$ est une partie ouverte de \mathbb{R}^2 .

☞ **Remarque XXVII.2.** Un ouvert est donc un ensemble dont tous les points peuvent être "entourés" par un disque interne à l'ensemble : aucun de ses points n'est donc "au bord" de l'ensemble.

▣► **Exemple XXVII.1.**

- \mathbb{R}^2 et \emptyset sont à la fois ouverts et fermés (on peut démontrer que ce sont les seuls ensembles de ce type) ;
- les singletons sont fermés ;
- les "segments époinés" $\{(1 - \lambda)a + \lambda b \mid \lambda \in [0, 1[\}$, pour $a, b \in \mathbb{R}^2$, ne sont ni ouverts ni fermés.

✎ **Exercice XXVII.1.** Démontrer que l'ensemble $\Lambda = \{(x, y) \in \mathbb{R}^2 \mid x < y\}$ est un ouvert.

Proposition XXVII.1. Tout disque ouvert (resp. fermé) est ouvert (resp. fermé).

Démonstration. Soit $x \in \mathbb{R}^2$ et soit $r \in \mathbb{R}_+^*$. Alors, si $y \in D(x, r)$, on a $\|x - y\| < r$ et donc si, on pose $\mu = r - \|x - y\|$ on a, pour tout $z \in D(y, \mu)$:

$$\begin{aligned} \|x - z\| &= \|x - y + y - z\| \\ &\leq \|x - y\| + \|y - z\| \\ &< \|x - y\| + r - \|x - y\| \\ &= r \end{aligned}$$

et donc on a bien $D(y, \mu) \subset D(x, r)$, ergo le disque ouvert $D(x, r)$ est...ouvert.

Concernant la fermeture du disque fermé, il nous faut démontrer que l'ensemble $A = \mathbb{R}^2 \setminus \overline{D}(x, r)$ est ouvert ; fixons donc $y \in A$ et remarquons qu'alors $\|x - y\| > r$. De fait, si l'on pose $\mu = \|x - y\| - r$ on a, pour tout $z \in D(y, \mu)$:

$$\begin{aligned} \|x - y\| &= \|x - z + z - y\| \\ &\leq \|x - z\| + \|z - y\| \\ &< \|x - z\| + \mu \end{aligned}$$

et donc :

$$\|x - z\| > \|x - y\| - \mu = r$$

d'où l'inclusion $D(y, \mu) \subset A$ et le résultat escompté. □

✎ **Remarque XXVII.3.** On peut démontrer que toute réunion d'ouverts est un ouvert (un dessin peut aider). De même, toute intersection **finie** d'ouverts reste ouverte ; une intersection infinie d'ouverts pouvant toutefois ne pas l'être. Par exemple :

$$\forall a \in \mathbb{R}^2, \quad \bigcap_{n \in \mathbb{N}^*} D\left(a, \frac{1}{n}\right) = \{a\}.$$

Proposition XXVII.2 (Séparation de \mathbb{R}^2). Soient $u, v \in \mathbb{R}^2$ tels que $u \neq v$. Alors il existe $r > 0$ tel que :

$$D(u, r) \cap D(v, r) = \emptyset.$$

Démonstration. Poser $r = \frac{\|u - v\|}{3}$. □

b) Applications continues de \mathbb{R}^2 dans \mathbb{R}

Définition XXVII.3. Soit $\Omega \subset \mathbb{R}^2$ un **ouvert** et soit $f : \Omega \rightarrow \mathbb{R}$. On dit que f est **continue** en un point $a \in \Omega$ lorsque :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in \Omega, \quad (\|x - a\| \leq \delta) \Rightarrow (|f(x) - f(a)| \leq \varepsilon).$$

Une fonction continue en tout point de l'ouvert Ω sera dite continue sur Ω .

✂ **Remarque XXVII.4.** Cette définition peut être reformulée de la façon suivante : f est continue en $a \in \Omega$ si et seulement si :

$$\forall V \in \mathcal{V}(f(a)), \exists \delta > 0, f(\overline{D}(a, \delta) \cap \Omega) \subset V.$$

▣ **Exemple XXVII.2.**

— La norme euclidienne est continue sur \mathbb{R}^2 . En effet, si $a, x \in \mathbb{R}^2$, on a :

$$| \|x\| - \|a\| | \leq \|x - a\|$$

et donc, pour tout $\varepsilon > 0$, " $\delta = \varepsilon$ " convient.

— Les fonctions $(x, y) \mapsto x$ et $(x, y) \mapsto y$ sont continues.

— Si $f : \Omega \rightarrow \mathbb{R}$ est telle qu'il existe $M > 0$ vérifiant la condition

$$\forall (x, y) \in \Omega, \quad |f(x) - f(y)| \leq M\|x - y\|$$

alors f est continue (on parle, sans surprise, de fonction *lipschitzienne*).

✂ **Remarque XXVII.5.** On peut étendre la définition de limite vue au chapitre IX assez aisément aux fonctions de deux variables à valeurs réelles. Plus précisément, on dira que $f : \Omega \rightarrow \mathbb{R}$ tend vers $\ell \in \mathbb{R}$ lorsque x tends vers $a \in \Omega$ si :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in \Omega, \quad (\|x - a\| \leq \delta) \Rightarrow (|f(x) - \ell| \leq \varepsilon).$$

Pour les limites infinies **en un point** $a \in \Omega$, on procède similairement : on dira que f tend vers $+\infty$ (resp. $-\infty$) lorsque x tends vers a si :

$$\forall A \in \mathbb{R}, \exists \delta > 0, \forall x \in \Omega, \quad (\|x - a\| \leq \delta) \Rightarrow (f(x) \geq A) \text{ (resp. } f(x) \leq A).$$

On remarquera que ces définitions sont liées à la norme euclidienne ; il sera donc assez efficace dans les calculs de passer en **coordonnées** polaires. Plus précisément, si $x \in \Omega$ et que l'on étudie une limite en $a \in \Omega$, en posant $r = \|x - a\|$ on sait (cf. chapitre XXV) qu'il existe un réel θ tel que :

$$x - a = (r \cos(\theta), r \sin(\theta))$$

et donc on peut souvent se ramener à étudier la limite de la fonction f lorsque $r \rightarrow 0$.

✎ **Exercice XXVII.2.** Démontrer que l'application définie par $f((0,0)) = 0$ et $f : (x, y) \mapsto \frac{xy}{x^2 + y^2}$ sur $\mathbb{R}^2 \setminus \{(0,0)\}$ n'est pas continue en $(0,0)$.

➔ **Correction :** Soit $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$. Alors, si l'on pose $r = \|(x, y)\|$, nous avons vu au chapitre **XXV** qu'il existe $\theta \in \mathbb{R}$ tel que $x = r \cos(\theta)$ et $y = r \sin(\theta)$. On a ensuite :

$$f(x, y) = \frac{r^2 \cos(\theta) \sin(\theta)}{r^2} = \frac{1}{2} \sin(2\theta)$$

ce qui entraîne que :

$$|f(x, y) - f(0, 0)| = \frac{1}{2} |\sin(2\theta)|.$$

Fixons $\alpha > 0$; si l'on choisit $\theta \notin \pi\mathbb{Z}$ et $r \leq \alpha$, on a :

$$\|(x, y) - (0, 0)\| = r \leq \alpha$$

et pourtant

$$|f(x, y) - f(0, 0)| = \frac{1}{2} |\sin(2\theta)|$$

ce qui contredit la définition de limite avec tout $\varepsilon < \frac{1}{2} |\sin(2\theta)|$.

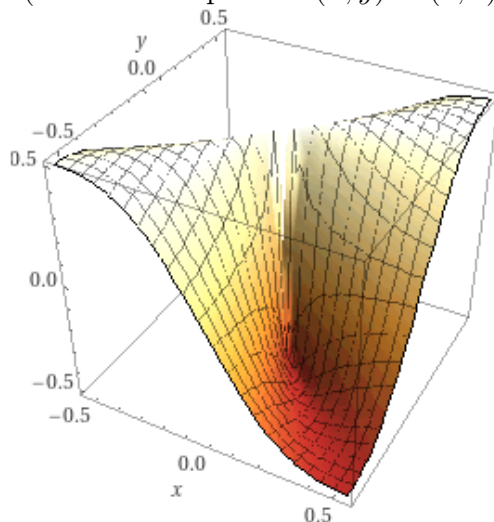
✂ Remarque XXVII.6.

- Le passage en coordonnées polaire (r, θ) est très commode au voisinage de l'origine car il permet de contrôler la "convergence" du couple (x, y) "vers l'origine" en faisant tendre (de façon rigoureuse) r vers 0. Pour se convaincre de la nécessité de ce type de démarche, remarquer que

$$\lim_{x \rightarrow 0} \left(\lim_{y \rightarrow 0} \frac{y}{x^2 + y} \right) \neq \lim_{y \rightarrow 0} \left(\lim_{x \rightarrow 0} \frac{y}{x^2 + y} \right).$$

Oui je sais, c'est glauque. Mais c'est la vie.

- Une fonction $f : \Omega \rightarrow \mathbb{R}$ peut être représentée graphiquement par la surface d'équation $z = f(x, y)$ dans \mathbb{R}^3 . Par exemple, la fonction *supra* correspond à la surface suivante (observer le "pic" en $(x, y) = (0, 0)$).



Définition XXVII.4. Soit $f : \Omega \rightarrow \mathbb{R}$ et soit $c \in \mathbb{R}$. On appelle **ligne de niveau** c de f l'ensemble (éventuellement vide) :

$$\mathcal{L}_c(f) = \{(x, y) \in \Omega \mid f(x, y) = c\}.$$

▮▮▮ **Exemple XXVII.3.** Sur une carte IGN, on peut lire les lignes de niveau de la fonction qui à un point (x, y) du plan associe l'altitude (relative au niveau de la mer) au sol en ce point.



✎ **Exercice XXVII.3.** Déterminer les lignes de niveau des fonction $(x, y) \mapsto x^2 + y^2$ et $g : (x, y) \mapsto x^2 - y$.

Les propositions qui suivent s'obtiennent de façon analogue au cas d'une variable réelle vu au chapitre IX. Leur démonstration ne constitue pas un attendu du programme de MPSI.

Proposition XXVII.3. Soient Ω un ouvert de \mathbb{R}^2 , $a \in \Omega$, $f, g : \Omega \rightarrow \mathbb{R}$ deux applications continues en a et $\lambda \in \mathbb{R}$. Alors :

- (i) $f + \lambda g$ est continue en a ;
- (ii) $f \times g$ est continue en a ;
- (iii) si $g(a) \neq 0$, il existe un $r > 0$ tel que g ne s'annule pas sur $D(a, r) \cap \Omega$ et $\frac{f}{g}$ est continue en a .

▮▮▮ **Exemple XXVII.4.**

— Les fonctions polynomiales, *i.e* celles de la forme :

$$(x, y) \mapsto \sum_{i=0}^p \sum_{j=0}^q \lambda_{i,j} x^i y^j$$

avec les $\lambda_{i,j} \in \mathbb{R}$ sont continues.

— La fonction $(x, y) \mapsto \frac{xy}{x^2 + y^2}$ est continue sur $\mathbb{R}^2 \setminus \{(0, 0)\}$

Proposition XXVII.4. Soient Ω un ouvert de \mathbb{R}^2 , I un intervalle de \mathbb{R} , $a \in \Omega$, $f : \Omega \rightarrow I$ continue en a et $g : I \rightarrow \mathbb{R}$ continue en $f(a)$. Alors $g \circ f$ est continue en a .

2. Dérivées partielles

On fixe dans ce paragraphe un ouvert Ω de \mathbb{R}^2 .

a) C'est quoi ?

Définition XXVII.5. Soit $u \in \mathbb{R}^2$, $a \in \Omega$ et $f : \Omega \rightarrow \mathbb{R}$. On dit que f est **dérivable selon le vecteur** u si il existe $\ell \in \mathbb{R}$ tel que :

$$\frac{f(a + tu) - f(a)}{t} \xrightarrow{t \rightarrow 0} \ell.$$

Le nombre réel ℓ est alors appelé **nombre dérivé** de f en a selon u et noté $df(a)(u)$.

☞ Remarque XXVII.7.

- Pour que cette définition ait un sens, il nous faut remarquer que, comme Ω est ouvert, il existe $r > 0$ tel que $D(a, r) \subset \Omega$. De fait, si $|t||u| < r$, on a bien $a + tu \in \Omega$ et donc la quantité $f(a + tu)$ a un sens quel que soit u pour t au voisinage de 0.
- f est dérivable en a selon le vecteur u si et seulement si la fonction d'une variable réelle $g : t \mapsto f(a + tu)$ est dérivable en 0. Par développement limité on a alors :

$$g(t) \underset{t \rightarrow 0}{=} g(0) + tg'(0) + o(t)$$

i.e

$$f(a + tu) \underset{t \rightarrow 0}{=} f(a) + t df(a)(u) + o(t).$$

☛ **Exemple XXVII.5.** Soit $f : (x, y) \mapsto x^2 + y$, définie sur \mathbb{R}^2 . Alors, pour tous $a = (x, y)$, $u = (\alpha, \beta) \in \mathbb{R}^2$ et $t \in \mathbb{R}^*$ on a :

$$\begin{aligned} \frac{f(a + tu) - f(a)}{t} &= \frac{(x + t\alpha)^2 + y + t\beta - x^2 - y}{t} \\ &= \frac{t(2x\alpha + t\alpha^2 + \beta)}{t} \\ &= 2x\alpha + t\alpha^2 + \beta \\ &\xrightarrow{t \rightarrow 0} 2x\alpha + \beta. \end{aligned}$$

On en déduit que l'application $df(a)$ est la forme linéaire

$$\begin{aligned} df(a) : \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (\alpha, \beta) &\mapsto 2x\alpha + \beta. \end{aligned}$$

Ceci préfigure un résultat plus général qui sera démontré en deuxième année.

Définition XXVII.6. Posons $e_1 = (1, 0)$ et $e_2 = (0, 1)$ les deux vecteurs de la base canonique de \mathbb{R}^2 . Si $f : \Omega \rightarrow \mathbb{R}$ est une application et que $a \in \Omega$, on appelle **dérivées partielles** de f en a les quantités $df(a)(e_1)$ et $df(a)(e_2)$, lorsqu'elles existent.

Notation. On pose dans ce cas $\frac{\partial f}{\partial x}(a) = df(a)(e_1)$ et $\frac{\partial f}{\partial y}(a) = df(a)(e_2)$. On trouve aussi dans la littérature les notations $\partial_x f(a)$ et $\partial_y f(a)$.

☞ **Remarque XXVII.8.** Si $a = (x, y)$ est un point en lequel f admet des dérivées partielles, on a :

$$\frac{\partial f}{\partial x} f(a) = \lim_{t \rightarrow 0} \frac{f(x+t, y) - f(x, y)}{t}$$

et

$$\frac{\partial f}{\partial y} f(a) = \lim_{t \rightarrow 0} \frac{f(x, y+t) - f(x, y)}{t}.$$

On reconnaît ici deux taux d'accroissements : le premier étant celui de $x \mapsto f(x, y)$ à y fixé, et le second celui de $y \mapsto f(x, y)$ à x fixé. On en déduit que les dérivées partielles peuvent être obtenues en dérivant "par rapport à la variable considérée" sans crainte de se planter violemment.

☛ **Exemple XXVII.6.** Les dérivées partielles de $f : (x, y) \mapsto x^2 - 3 \cos(y) + xy$ sont :

$$\frac{\partial f}{\partial x} : (x, y) \mapsto 2x + y \quad \text{et} \quad \frac{\partial f}{\partial y} : (x, y) \mapsto 3 \sin(y) + x,$$

ces deux fonctions étant définies sur \mathbb{R}^2 .

✖ **ATTENTION :** l'existence des dérivées partielles n'implique **PAS** la continuité de la fonction au point considéré. En effet, si l'on reprend la fonction définie par $f((0,0)) = 0$ et $f : (x, y) \mapsto \frac{xy}{x^2 + y^2}$ sur $\mathbb{R}^2 \setminus \{(0,0)\}$ (que nous savons discontinue en 0), nous avons, pour tout $t \in \mathbb{R}^*$:

$$\frac{f(t,0) - f(0,0)}{t} = 0 \quad \text{et} \quad \frac{f(0,t) - f(0,0)}{t} = 0$$

ce qui entraîne que les deux dérivées partielles $\frac{\partial f}{\partial x}(0,0)$ et $\frac{\partial f}{\partial y}(0,0)$ existent et sont nulles.

Proposition XXVII.5. Soient $f, g : \Omega \rightarrow \mathbb{R}$ deux applications admettant une dérivée partielle "selon x " (i.e selon e_1) en un point $a \in \Omega$ et soit $\lambda \in \mathbb{R}$. Alors :

(i) $f + \lambda g$ admet une dérivée partielle "selon x " en a et

$$\frac{\partial}{\partial x}(f + \lambda g)(a) = \frac{\partial f}{\partial x}(a) + \lambda \frac{\partial g}{\partial x}(a);$$

(ii) $f \times g$ admet une dérivée partielle "selon x " en a et

$$\frac{\partial}{\partial x}(f \times g)(a) = \frac{\partial f}{\partial x}(a) \times g(a) + f(a) \times \frac{\partial g}{\partial x}(a);$$

(iii) si $g(a) \neq 0$, il existe un $r > 0$ tel que g ne s'annule pas sur $D(a, r) \cap \Omega$, les fonction $\frac{1}{g}$ et $\frac{f}{g}$ admettent une dérivée partielle "selon x " en a vérifiant :

$$\frac{\partial}{\partial x} \left(\frac{1}{g} \right) (a) = -\frac{1}{g^2(a)} \frac{\partial g}{\partial x}(a) \quad \text{et} \quad \frac{\partial}{\partial x} \left(\frac{f}{g} \right) (a) = \frac{\frac{\partial f}{\partial x}(a) \times g(a) - f(a) \times \frac{\partial g}{\partial x}(a)}{g^2(a)}.$$

Démonstration. Passer par les fonctions $t \mapsto f(a + tu)$ et $t \mapsto g(a + tu)$ et leur appliquer les résultats démontrés au chapitre XI. \square

✂ **Remarque XXVII.9.** On peut (et on doit !) évidemment obtenir un résultat similaire pour les dérivées partielles "selon y " (et même sur les dérivées selon un vecteur quelconque).

b) Fonctions de classe \mathcal{C}^1

Définition XXVII.7. Une fonction $f : \Omega \rightarrow \mathbb{R}$ est dite **de classe \mathcal{C}^1** si elle admet des dérivées partielles **continues** en tout point de Ω .

Notation. $f \in \mathcal{C}^1(\Omega)$

✂ **Remarque XXVII.10.** La proposition XXVII.5 nous permet d'affirmer que l'ensemble $(\mathcal{C}^1(\Omega), +, \times, \cdot)$ est une \mathbb{R} -algèbre commutative (*i.e* un anneau commutatif et un \mathbb{R} -e.v).

▣ **Exemple XXVII.7.**

- Les fonctions polynomiales sont de classe \mathcal{C}^1 sur \mathbb{R}^2 .
- La fonction $(x, y) \mapsto e^x + x \ln(y)$ est de classe \mathcal{C}^1 sur $\mathbb{R} \times \mathbb{R}_+^*$.

Proposition XXVII.6. Soit $f \in \mathcal{C}^1(\Omega)$ et soit $a \in \Omega$. Alors, pour tout $u = (h, k) \in \mathbb{R}^2$ tel que $a + u \in \Omega$ on a :

$$f(a + u) = f(a) + h \frac{\partial f}{\partial x}(a) + k \frac{\partial f}{\partial y}(a) + o_0(\|u\|),$$

où $o_0(\|u\|)$ représente un produit de la forme $\varphi(u)\|u\|$, avec une fonction φ définie sur un disque ouvert centré en $(0, 0)$ et vérifiant que :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall u \in \mathbb{R}^2, (\|u\| \leq \delta) \Rightarrow (|\varphi(u)| \leq \varepsilon).$$

Démonstration. Admis. \square

✂ **Remarque XXVII.11.**

- Plus généralement, on peut démontrer que si $f \in \mathcal{C}^1(\Omega)$, alors elle admet une dérivée selon tout vecteur $u \in \mathbb{R}^2$ et qu'alors, pour tout $a \in \Omega$:

$$df(a)(u) = h \frac{\partial f}{\partial x}(a) + k \frac{\partial f}{\partial y}(a)$$

et donc :

$$f(a + u) = f(a) + df(a)(u) + o_0(\|u\|).$$

Ceci est remarquable : l'existence et la continuité des dérivées partielles entraîne l'existence de dérivée selon tous les vecteurs du plan.

- La fonction f pouvant être représentée graphiquement par la surface d'équation $z = f(x, y)$ dans \mathbb{R}^3 , il est raisonnable de se demander si le développement limité *supra* possède une interprétation géométrique. Si $a = (x_0, y_0)$ et $u = (h, k)$ on a, en posant $b = (x, y) = a + u$, $\lambda = \frac{\partial f}{\partial x}(a)$ et $\mu = \frac{\partial f}{\partial y}(a)$:

$$\begin{aligned} f(b) &= f(a) + \lambda h + \mu k + o_0(\|u\|) \\ &= f(a) + \lambda(x - x_0) + \mu(y - y_0) + o_0(\|u\|). \end{aligned}$$

Nous savons (*cf.* chapitre **XXI**) que l'ensemble d'équation :

$$z = f(a) + \lambda(x - x_0) + \mu(y - y_0)$$

est un plan affine de \mathbb{R}^3 . Celui-ci se retrouve "confondu" à la surface représentative de f "au voisinage" du point a ; on l'appellera **plan tangent** en a à cette dernière.

Corollaire XXVII.6.a. Toute fonction de classe \mathcal{C}^1 est continue.

Démonstration. Soit $f \in \mathcal{C}^1(\Omega)$ et soit $a \in \Omega$. Alors, pour tout $\varepsilon > 0$ et tout $x \in \Omega$ on a, en posant $(h, k) = x - a$ et en utilisant la proposition **XXVII.6** :

$$\begin{aligned} |f(x) - f(a)| &= \left| h \frac{\partial f}{\partial x}(a) + k \frac{\partial f}{\partial y}(a) + o_0(\|x - a\|) \right| \\ &\leq \left| h \frac{\partial f}{\partial x}(a) + k \frac{\partial f}{\partial y}(a) \right| + |o_0(\|x - a\|)| \\ &\leq 2 \max \left(\left| \frac{\partial f}{\partial x}(a) \right|, \left| \frac{\partial f}{\partial y}(a) \right| \right) \|x - a\| + o_0(\|x - a\|). \end{aligned}$$

Ainsi, on a par définition de continuité et de o le résultat voulu. □

c) Gradient, extrema

Définition XXVII.8. Soit $f \in \mathcal{C}^1(\Omega)$ et soit $a \in \Omega$. On appelle **gradient** de f en a le vecteur

$$\nabla f(a) = \left(\frac{\partial f}{\partial x}(a), \frac{\partial f}{\partial y}(a) \right) \in \mathbb{R}^2.$$

▮ **Exemple XXVII.8.** Soit $f : (x, y) \mapsto x^3 + e^x y - 7$; alors :

$$\forall (x, y) \in \mathbb{R}^2, \quad \nabla f(x, y) = (3x^2 + e^x y, e^x).$$

☞ **Remarque XXVII.12.** La proposition **XXVII.6** se réécrit de la façon suivante : pour tout $a \in \Omega$ et tout $u \in \mathbb{R}^2$ tel que $a + u \in \Omega$:

$$f(a + u) = f(a) + \langle \nabla f(a), u \rangle + o_0(\|u\|),$$

en remarquant au passage que :

$$df(a)(u) = \langle \nabla f(a), u \rangle .$$

Si l'on suppose que $\nabla f(a) \neq 0$ et que l'on prend $u \neq (0, 0)$, alors on a :

$$f(a+u) - f(a) = \|\nabla f(a)\| \|u\| \cos(\theta) + o_0(\|u\|),$$

où θ est l'écart angulaire (cf. **XXV**) entre $\nabla f(a)$ et u . À la louche, en fixant $\|u\|$, on obtient que la valeur absolue $|f(a+u) - f(a)|$ est maximale lorsque $\theta \equiv 0 [2\pi]$ ou $\theta \equiv \pi [2\pi]$, i.e lorsque $\nabla f(a)$ et u sont colinéaires dans \mathbb{R}^2 . On en déduit que **le gradient de f en a définit la direction dans laquelle f croît le plus vite** au voisinage de a .

Vocabulaire. Un point $a \in \Omega$ tel que $\nabla f(a) = (0, 0)$ est appelé **point critique** de la fonction f .

Le lecteur perspicace se souviendra que lorsqu'il est question de points critiques, les extrema ne sont généralement pas loin. Ce lecteur a raison.

Définition XXVII.9. Soit $f \in \mathcal{C}^1(\Omega)$ et soit $a \in \Omega$. Alors :

— on dit que f admet un **maximum** (resp. minimum) **local** en a si :

$$\exists \varepsilon > 0, \forall x \in D(a, \varepsilon) \cap \Omega, \quad f(x) \leq f(a) \quad (\text{resp. } f(x) \geq f(a));$$

— on dit que f admet un **maximum** (resp. minimum) **global** en a si :

$$\forall x \in \Omega, \quad f(x) \leq f(a) \quad (\text{resp. } f(x) \geq f(a)).$$

L'étude des extrema locaux de fonctions de plusieurs variables est sensiblement plus subtile que son analogue dans le cas univarié. En effet, une surface possède une infinité de directions dans lesquelles elle peut "monter" ou "descendre" au voisinage d'un point. On note toutefois avec soulagement que l'on dispose d'une condition nécessaire analogue à celle vue au chapitre **XI** (proposition **XI.9**).

Proposition XXVII.7. Soit $f \in \mathcal{C}^1(\Omega)$ et soit $a \in \Omega$. Si f admet un extremum local en a , alors $\nabla f(a) = 0$.

Démonstration. Supposons, pour fixer les idées, que f admet un maximum local en a et soit $u \in \mathbb{R}^2$. Alors, pour tout $t \in \mathbb{R}_+^*$ assez proche de zéro, on a $a + tu \in \Omega$ et $f(a + tu) \leq f(a)$, ce qui entraîne que

$$\frac{f(a + tu) - f(a)}{t} \leq 0$$

et donc en passant à la limite quand $t \rightarrow 0$, on obtient que $df(a)(u) \leq 0$. Le même raisonnement mené avec $t < 0$ nous livre l'inégalité inverse et donc $df(a)(u) = 0$; en particulier on a bien $\nabla f(a) = 0$. \square

Remarque XXVII.13. La démonstration *supra* nous indique qu'en plus du gradient, toutes les dérivées selon un vecteur donnée sont nulles en un extremum local. Le plan tangent est alors "horizontal", i.e de direction " $z = 0$ ".

▣▣▣ **Exemple XXVII.9.**

- Il est aisé de vérifier que $(x, y) \mapsto x^2 + y^2$ admet un minimum local et un point critique en $(0, 0)$.
- La fonction $f : (x, y) \mapsto x^2 - y^2$ admet un point critique en $(0, 0)$. Cependant, si $t \in \mathbb{R}_+^*$ on a :

$$f(t, 0) = t^2 > 0 \quad \text{et} \quad f(0, t) = -t^2 < 0$$

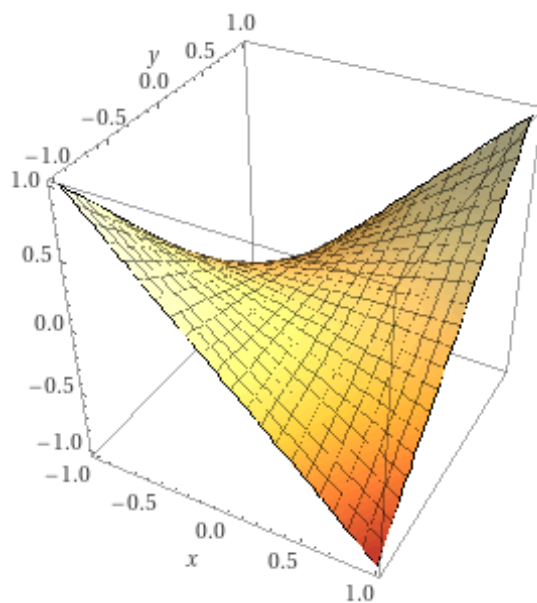
ce qui prouve que f n'admet ni un maximum local ni un minimum local en $(0, 0)$. La réciproque de la proposition **XXVII.7** est donc, une fois n'est pas coutume, **fausse**.

✎ **Exercice XXVII.4.** Déterminer les éventuels extrema locaux de la fonction $f : (x, y) \mapsto xy^3$.

➔ **Correction :** La fonction f est de classe \mathcal{C}^1 sur \mathbb{R}^2 car polynomiale. De plus, pour tout $(x, y) \in \mathbb{R}^2$:

$$\nabla f(x, y) = (y^3, 3xy^2)$$

et donc les seuls points critiques de f sont les $(t, 0)$, pour $t \in \mathbb{R}$, et $f(t, 0) = 0$. Si x et y sont de même signe, on a $f(x, y) > 0$; à l'inverse, si ils sont de signes opposés, on a $f(x, y) < 0$. De fait, f n'admet aucun extremum local.



d) Composition

Proposition XXVII.8 (Règle de la chaîne). Soit I un intervalle de \mathbb{R} , $f \in \mathcal{C}^1(\Omega)$, et $x, y \in \mathcal{C}^1(I)$ telles que :

$$\forall t \in I, (x(t), y(t)) \in \Omega.$$

Alors :

- (i) la fonction $g : t \mapsto f(x(t), y(t))$ est de classe $\mathcal{C}^1(I)$;
- (ii) pour tout $t \in I$, on a :

$$g'(t) = x'(t) \frac{\partial f}{\partial x}(x(t), y(t)) + y'(t) \frac{\partial f}{\partial y}(x(t), y(t)).$$

Démonstration. Soit $t \in I$ et $h \in \mathbb{R}^*$ tel que $t+h \in I$; si on pose $T = (x(t), y(t))$ et $U = (x(t+h), y(t+h))$ on a :

$$\begin{aligned} g(t+h) - g(t) &= f(U) - f(T) \\ &= \langle \nabla f(T), U - T \rangle + o_0(\|U - T\|) \end{aligned}$$

d'après la proposition **XXVII.6**. De plus, comme $x, y \in \mathcal{C}^1(I)$ on a (cf. proposition **XI.1**) :

$$\begin{aligned} U - T &= (x(t+h) - x(t), y(t+h) - y(t)) \\ &= (hx'(t) + o_0(h), hy'(t) + o_0(h)) \end{aligned}$$

ce qui entraîne que $o_0(\|U - T\|) = o_0(h)$ et donc :

$$g(t+h) - g(t) = (hx'(t) + o_0(h)) \frac{\partial f}{\partial x}(x(t), y(t)) + (hy'(t) + o_0(h)) \frac{\partial f}{\partial y}(x(t), y(t)) + o_0(h)$$

ergo :

$$\begin{aligned} \frac{g(t+h) - g(t)}{h} &= (x'(t) + o_0(1)) \frac{\partial f}{\partial x}(x(t), y(t)) + (y'(t) + o_0(1)) \frac{\partial f}{\partial y}(x(t), y(t)) + o_0(1) \\ &\xrightarrow{h \rightarrow 0} x'(t) \frac{\partial f}{\partial x}(x(t), y(t)) + y'(t) \frac{\partial f}{\partial y}(x(t), y(t)) \end{aligned}$$

d'où le résultat. □

✂ **Remarque XXVII.14.**

— Cette formule peut se résumer, "à la physicienne", de la façon suivante :

$$\frac{dg}{dt} = \frac{dx}{dt} \frac{\partial f}{\partial x} + \frac{dy}{dt} \frac{\partial f}{\partial y}$$

pour peu que l'on sache en quels points ces dérivées sont calculées, ce qui se déduit aisément du contexte.

- Cinématiquement / géométriquement, la fonction g correspond à la "restriction" de f le long d'un arc paramétré par $\gamma : t \mapsto (x(t), y(t))$. La règle de la chaîne se réécrit alors :

$$\forall t \in I, \quad (f \circ \gamma)'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle .$$

Cette expression correspond à la dérivée de f le long de l'arc γ . Ce type de considération est très courant (par exemple) en mécanique. Par exemple, pour déterminer la distance maximale ou minimal d'un mobile $t \mapsto (x(t), y(t))$ à l'origine, on pourra étudier les extrema de $t \mapsto x(t)^2 + y(t)^2$.

Proposition XXVII.9. Soit $f \in \mathcal{C}^1(\Omega)$ et soit $c \in \mathbb{R}$. Alors, pour tout arc paramétré de classe \mathcal{C}^1 (coordonnée par coordonnée) $\gamma : I \rightarrow \mathcal{L}_c(f)$ on a :

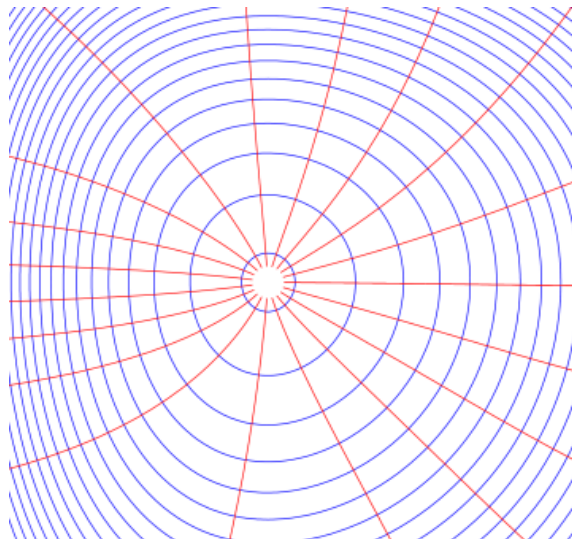
$$\forall t \in I, \quad \langle \nabla f(\gamma(t)), \gamma'(t) \rangle = 0$$

Démonstration. Si γ vérifie les conditions de l'énoncé, d'après la remarque **XXVII.14** :

$$\forall t \in I, \quad (f \circ \gamma)'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle .$$

Or, par définition de $\mathcal{L}_c(f)$, la fonction $f \circ \gamma$ est constante égale à c , d'où le résultat. \square

Remarque XXVII.15. Cela signifie que le gradient est "orthogonal" en tout point à toute ligne de niveau passant par ce point, au sens où il est orthogonal à la tangente (dirigée par $\gamma'(t)$ en tout t) de celle-ci.



Proposition XXVII.10 (Règle de la chaîne généralisée). Soit Λ un ouvert de \mathbb{R}^2 , $f \in \mathcal{C}^1(\Omega)$ et $u, v : \Lambda \rightarrow \mathbb{R}$ deux fonctions de classes \mathcal{C}^1 telles que :

$$\forall x, y \in \Lambda, (u(x, y), v(x, y)) \in \Omega.$$

Alors :

- (i) la fonction $g : (x, y) \mapsto f(u(x, y), v(x, y))$ est de classe \mathcal{C}^1 sur Λ ;
- (ii) pour tout $(x, y) \in \Lambda$ on a :

$$\frac{\partial g}{\partial x}(x, y) = \frac{\partial u}{\partial x}(x, y) \frac{\partial f}{\partial x}(u(x, y), v(x, y)) + \frac{\partial v}{\partial x}(x, y) \frac{\partial f}{\partial y}(u(x, y), v(x, y))$$

et

$$\frac{\partial g}{\partial y}(x, y) = \frac{\partial u}{\partial y}(x, y) \frac{\partial f}{\partial x}(u(x, y), v(x, y)) + \frac{\partial v}{\partial y}(x, y) \frac{\partial f}{\partial y}(u(x, y), v(x, y)).$$

Démonstration. Beurk. □

☞ **Remarque XXVII.16.** Si on prend le parti de noter $\frac{\partial f}{\partial u}$ et $\frac{\partial f}{\partial v}$ les dérivées partielles de f , on peut "résumer" (comme pour la règle de la chaîne "basique") la proposition *supra* par les "formules" suivante :

$$\frac{\partial g}{\partial x} = \frac{\partial u}{\partial x} \frac{\partial f}{\partial u} + \frac{\partial v}{\partial x} \frac{\partial f}{\partial v}$$

et

$$\frac{\partial g}{\partial y} = \frac{\partial u}{\partial y} \frac{\partial f}{\partial u} + \frac{\partial v}{\partial y} \frac{\partial f}{\partial v}.$$

🔗 **Exercice XXVII.5.** Déterminer toutes les fonctions $f \in \mathcal{C}^1(\mathbb{R}^2)$ vérifiant l'équation :

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} = 0. \quad (\text{E :XXVII.1})$$

On pourra faire usage du changement de variable $u = x + y, v = x - y$.

➡ **Correction :** *Procédons par analyse-synthèse et supposons trouvée f solution de (E :XXVII.1). Alors, comme les fonctions $u : (x, y) \mapsto x + y$ et $v : (x, y) \mapsto x - y$ sont de classe \mathcal{C}^1 , la fonction $\varphi : (x, y) \mapsto (x + y, x - y)$ est une bijection \mathcal{C}^1 de \mathbb{R}^2 . En posant $g = f \circ \varphi^{-1}$ on, par proposition XXVII.10 et comme $f = g \circ \varphi$, pour tout $(x, y) \in \mathbb{R}^2$:*

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial g}{\partial u}(x + y, x - y) + \frac{\partial g}{\partial v}(x + y, x - y)$$

et

$$\frac{\partial f}{\partial y}(x, y) = \frac{\partial g}{\partial u}(x + y, x - y) - \frac{\partial g}{\partial v}(x + y, x - y).$$

De fait, en réinjectant dans l'équation, on obtient :

$$\frac{\partial g}{\partial u}(x + y, x - y) = 0$$

ce qui signifie qu'il existe G de classe \mathcal{C}^1 sur \mathbb{R} telle que :

$$\forall x, y \in \mathbb{R}, g(x + y, x - y) = G(x - y)$$

i.e

$$\forall (x, y) \in \mathbb{R}^2, f(x, y) = G(x - y).$$

On vérifie ensuite rapidement que les fonctions de cette forme sont bien solution de (**E :XXVII.1**).

Addendum : méthode des moindres carrés

Ce qui suit n'est pas *stricto sensu* au programme mais nous semble un approfondissement adapté aux sujets déjà étudiés.

Nous nous intéressons dans ce paragraphe à une méthode de **régression linéaire** : l'idée est que l'on dispose de $n \geq 3$ points (x_i, y_i) du plan d'abscisses deux à deux distinctes, que l'on cherche à "approcher" par une droite affine d'équation $y = ax + b$ "raisonnablement proche" (pour parler comme un chimiste) de ces derniers. Deux (ou trois) questions nous viennent alors (en tous cas je l'espère) à l'esprit :

- quel sens donner à "raisonnablement proche" ?
- Cela est-il toujours réalisable ? Et si oui, comment ?

Pour formaliser ceci, introduisons la fonction suivante :

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

$$(a, b) \mapsto \sum_{i=1}^n (y_i - (ax_i + b))^2.$$

Cette fonction calcule la somme la somme des carrés des distances des points (x_i, y_i) à une droite d'équation donnée. Minimiser J semble donc une bonne idée pour trouver des paramètres a et b définissant une droite "raisonnablement proche" de notre nuage de points.

Existence du minimum

Commençons par remarquer que J est une fonction de classe \mathcal{C}^1 de \mathbb{R}^2 dans \mathbb{R} à valeurs positives. De fait, l'ensemble (non vide) $J(\mathbb{R}^2)$ admet une borne inférieure positive. Le problème devient donc de savoir si cette dernière est atteinte.

On se place dans \mathbb{R}^n muni de son produit scalaire canonique. Alors, en posant $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ et $e = (1, \dots, 1)$, on a :

$$\forall a, b \in \mathbb{R}, \quad J(a, b) = \|y - (ax + be)\|^2$$

et donc, si on pose $F = \text{Vect}(x, e)$, on a :

$$\inf(J) = d(y, F)^2.$$

F est un s-e.v de \mathbb{R}^n de dimension 2 (car les x_i sont deux à deux distincts), donc cette distance est atteinte en le projeté orthogonal de y sur F : $\min(J)$ existe donc.

Calcul explicite

Pour le calcul effectif du minimum, il serait possible de déterminer le projeté orthogonal mentionné *supra* par les méthodes vues au chapitre **XXV**. Cependant, il est bien plus rapide ici de s'intéresser aux point(s) critique(s) de J . En effet, on a pour tous $a, b \in \mathbb{R}$:

$$\frac{\partial J}{\partial a}(a, b) = -2 \sum_{i=1}^n x_i (y_i - ax_i - b)$$

et

$$\frac{\partial J}{\partial b}(a, b) = -2 \sum_{i=1}^n (y_i - ax_i - b).$$

Posons :

$$\mu_x = \sum_{i=1}^n x_i^2, \quad \lambda_x = \sum_{i=1}^n x_i, \quad \lambda_y = \sum_{i=1}^n y_i \quad \text{et} \quad \mu_{xy} = \sum_{i=1}^n x_i y_i;$$

on a alors que (a, b) est un point critique de J si et seulement si :

$$\begin{cases} \mu_x a + \lambda_x b = \mu_{xy} \\ \lambda_x a + nb = \lambda_y \end{cases}$$

Les x_i étant deux à deux distincts, on a $n\mu_x - \lambda_x^2 \neq 0$ (appliquer l'inégalité de Cauchy-Schwarz aux vecteurs x et e) et donc ce système admet une unique solution donnée par :

$$a = \frac{n\mu_{xy} - \lambda_x \lambda_y}{n\mu_x - \lambda_x^2} \quad \text{et} \quad b = \frac{\lambda_y \mu_x - \lambda_x \mu_{xy}}{n\mu_x - \lambda_x^2}.$$

Comme J admet un minimum et un unique point critique, ces derniers sont confondus : nous venons de déterminer les paramètres optimaux de régression linéaire (au sens de notre fonction J), qui sont *in fine* donnés par la magnifique expression suivante¹ :

$$a = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad \text{et} \quad b = \frac{\left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n x_i^2 \right) - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n x_i y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}.$$

Vision probabiliste

Nous donnons dans ce paragraphe une vision probabiliste² du résultat précédent. Pour ce faire, plaçons nous dans un espace probabilisé (Ω, \mathbb{P}) et considérons que nos points (x_i, y_i) sont en réalité une série de n réalisations d'un couple de variables aléatoires réelles (X, Y) . Plus précisément, supposons que :

$$\forall i \in \llbracket 1, n \rrbracket, \quad (x_i, y_i) \in X \times Y(\Omega)$$

et rajoutons l'hypothèse (raisonnable!) que $V(X) \neq 0$.

1. Lunettes de protection recommandées.

2. En vérité, cela relève plutôt d'une approche statistique mais préservons notre santé mentale, voulez-vous.

Remarquons que, par loi faible des grands nombres, on peut "raisonnablement" approximer l'espérance de X par sa moyenne empirique, *i.e.*, à la statisticienne :

$$\mathbb{E}(X) \cong \frac{1}{n} \sum_{i=1}^n x_i$$

et, similairement :

$$V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \cong \frac{1}{n} \sum_{i=1}^n x_i^2 - \frac{1}{n^2} \left(\sum_{i=1}^n x_i \right)^2.$$

En persévérant dans cette approche, on arrive à l'approximation :

$$\text{Cov}(X, Y) \cong \frac{1}{n} \left(\sum_{i=1}^n x_i y_i \right) - \frac{1}{n^2} \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right).$$

Ainsi, les coefficients optimaux déterminés dans la section précédente deviennent :

$$a = \frac{\text{Cov}(X, Y)}{V(X)} \quad \text{et} \quad b = \mathbb{E}(Y) - \frac{\text{Cov}(X, Y)}{V(X)} \mathbb{E}(X).$$

Cette approche nous permet de réaliser rapidement une régression linéaire de réalisations de variables aléatoires dont les lois sont connues ou supposées (pour des raisons physiques, par exemple).

Les valeurs classiques de ces (co)variances étant tabulées, ceci nous³ permet de vérifier la cohérence de résultats expérimentaux.

Généralisations diverses

Il est possible de généraliser l'approche développée ici à l'approximation de nuages de points par des fonctions non affines. On sera amené à considérer des applications du type :

$$J : f \mapsto \sum_{i=1}^n (y_i - f(x_i))^2$$

définies sur un espace de fonctions raisonnable (idéalement décrit par un nombre fini de paramètres). Par exemple, on pourra s'intéresser aux fonctions polynomiales de degré borné ou à des familles d'exponentielles à paramètres.

Il est important de noter que ces espaces fonctionnels seront rarement des \mathbb{R} -e.v de dimension finie : l'existence du minimum de J ne sera donc pas toujours aisée à établir, et son calcul explicite peut être difficile.

3. Non contractuel.