Fonctions de deux variables

Dans tout ce chapitre, $\mathbb{R}^2$ et $\mathbb{R}^3$ sont munis de leur structure euclidienne canonique. On identifiera $\mathbb{R}^2$ au plan d'équation $z = 0$ de $\mathbb{R}^3$, ce qui revient à identifier tout couple (x, y) de $\mathbb{R}^2$ au triplet $(x, y, 0)$ de $\mathbb{R}^3$. On notera $\left(\vec{\imath}, \vec{\jmath}\right)$ la base canonique de $\mathbb{R}^2$ et $\left(\vec{\imath}, \vec{\jmath}, \vec{k}\right)$ celle de $\mathbb{R}^3$, et enfin O le point $(0, 0)$ ou $(0, 0, 0)$.

1 Graphe d'une fonction de deux variables

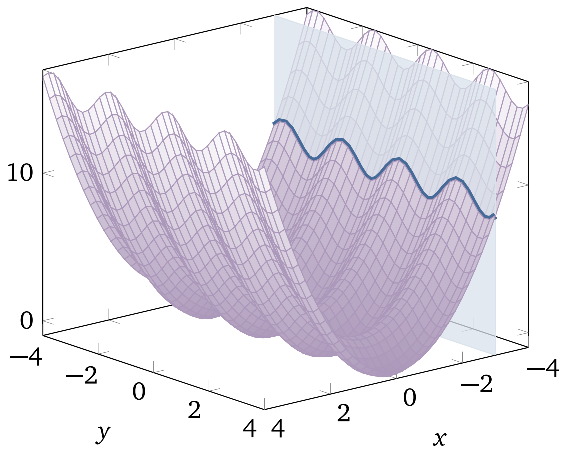
Nous avons l'habitude de représenter toute fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$ comme une courbe dans le plan $\mathbb{R}^2$, précisément la courbe d'équation $y = f(x)$. Une fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ sera quant à elle représentée comme la surface d'équation $z = f(x, y)$. Intéressons-nous par exemple à la fonction $(x, y) \overset{f}{\longmapsto} x^2 + \sin(3y)$. Pour construire son graphe $\mathscr{S}$, on étudie souvent son intersection avec une collection de plans parallèles qui balaient l'espace $\mathbb{R}^3$ tout entier. Ainsi, pour tout $\lambda \in \mathbb{R}$:

— l'intersection de $\mathscr{S}$ et du plan d'équation $x = \lambda$ est la courbe d'équation $z = \lambda^2 + \sin(3y)$ dans ce plan,

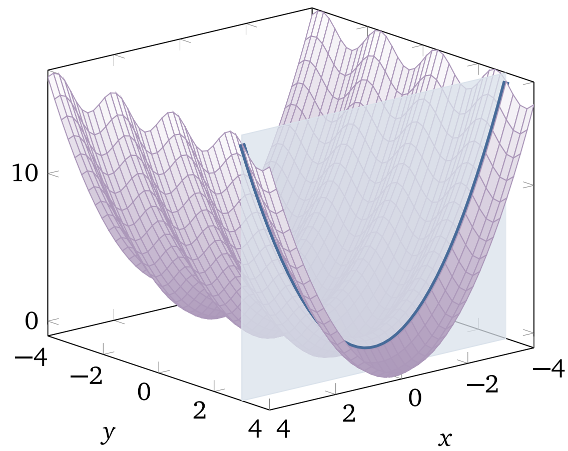
— l'intersection de $\mathscr{S}$ et du plan d'équation $y = \lambda$ est la courbe d'équation $z = x^2 + \sin(3\lambda)$ dans ce plan,

— l'intersection de $\mathscr{S}$ et du plan d'équation $z = \lambda$ est la courbe d'équation $x^2 + \sin(3y) = \lambda$ dans ce plan et on l'appelle la *ligne de niveau λ de f*.

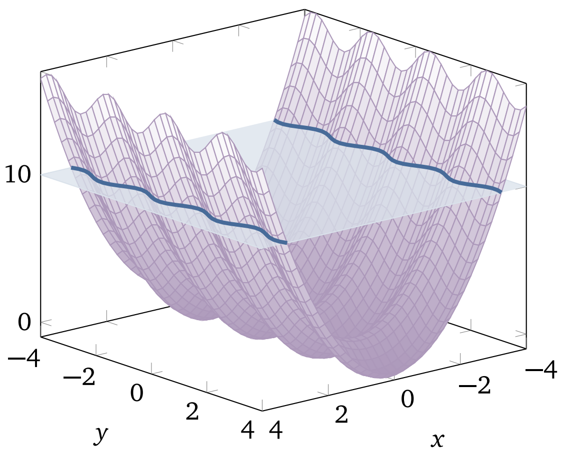
Intersection de $\mathscr{S}$ avec $x = -3$



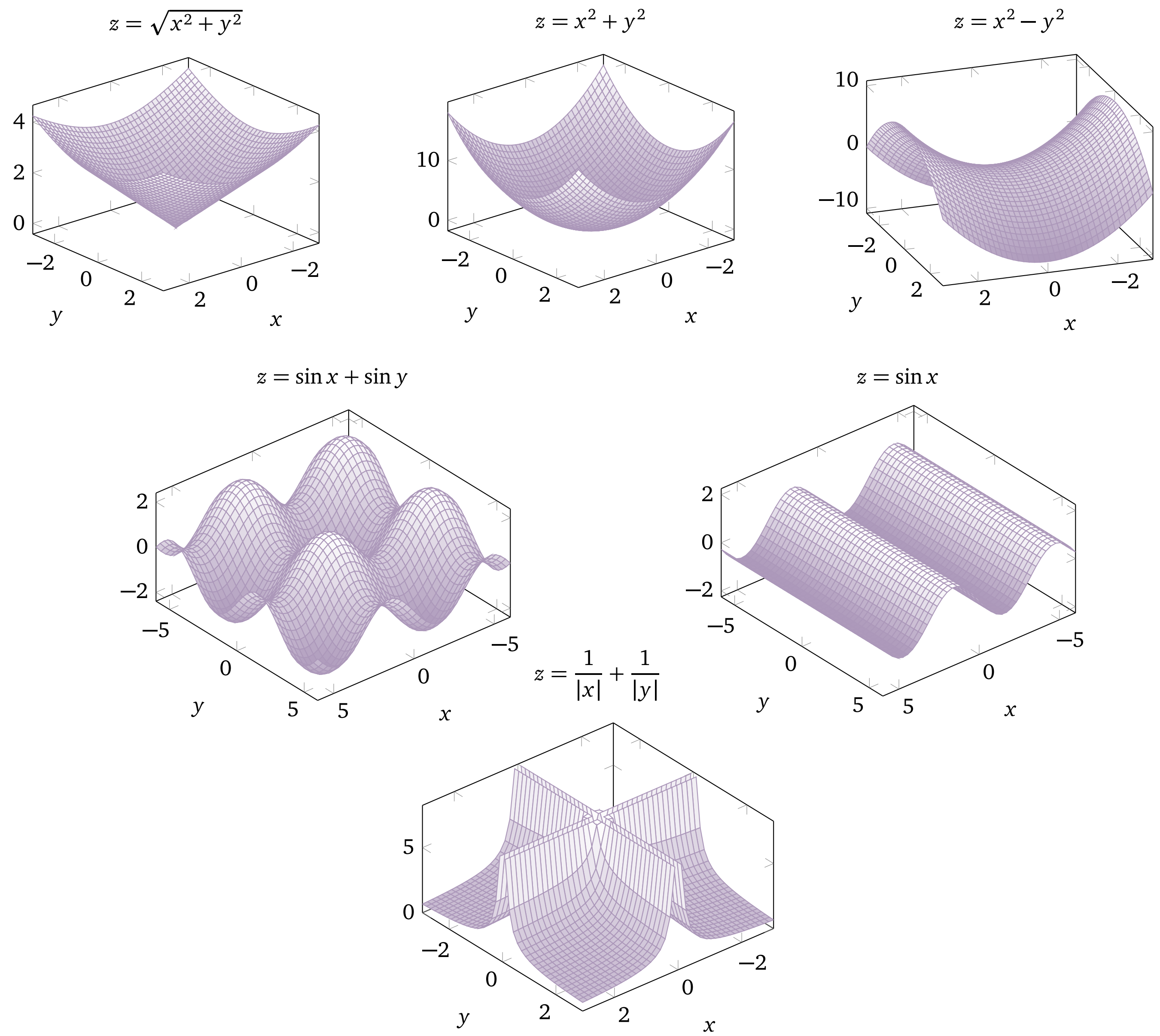
Intersection de $\mathscr{S}$ avec $y = 3$



Intersection de $\mathscr{S}$ avec $z = 10$



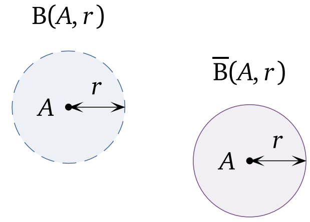
Faites l'effort de réfléchir en les mêmes termes aux exemples qui suivent et assurez-vous que vous les comprenez bien.



2 Rudiments de topologie dans $\mathbb{R}^2$

On introduit brièvement dans ce paragraphe, et sans s'apesantir sur les preuves, quelques notions topologiques simples dont nous aurons besoin ensuite.

2.1 Voisinages et ouverts



Pour tous $A \in \mathbb{R}^2$ et $r > 0$, on appelle :

— *boule ouverte de centre A et de rayon r* l'ensemble $\mathrm{B}(A,r) = \left\{M \in \mathbb{R}^2 \mid \quad AM < r\right\}$,

— *boule fermée de centre A et de rayon r* l'ensemble $\overline{\mathrm{B}}(A,r) = \left\{M \in \mathbb{R}^2 \mid \quad AM \leqslant r\right\}$.

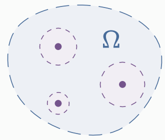
Dans le plan $\mathbb{R}^2$, on gagnerait bien sûr à parler de disques et de cercles, mais la théorie développée dans ce chapitre s'étend sans difficulté majeure à $\mathbb{R}^n$ pour tout $n \geqslant 3$, alors autant parler tout de suite de boules et de sphères.

La définition qui suit ne devrait pas trop vous dépayser.

Définition (Voisinage d'un point) Soit $A \in \mathbb{R}^2$. On appelle *voisinage de A* toute partie de $\mathbb{R}^2$ contenant une boule ouverte de centre A. On notera dans ce cours $\mathscr{V}_A\left(\mathbb{R}^2\right)$ l'ensemble des voisinages de A.

Définition-théorème (Ouvert) Soit Ω une partie de $\mathbb{R}^2$. On dit que Ω est (un) *ouvert* si :

$$\forall A \in \Omega, \quad \exists\, r > 0, \quad \mathrm{B}(A, r) \subset \Omega,$$



autrement dit si Ω est un voisinage de chacun de ses points.

Exemples d'ouverts :

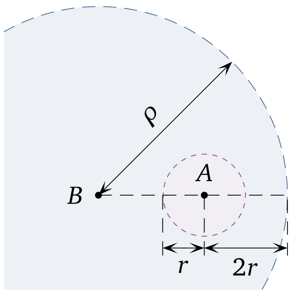
(i) Toute boule ouverte est un ouvert.

(ii) Le complémentaire de toute boule fermée est un ouvert. En particulier, $\mathbb{R}^2 \setminus \{A\}$ est un ouvert pour tout $A \in \mathbb{R}^2$.

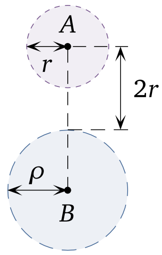
(iii) Tout produit de deux intervalles ouverts est un ouvert.

Démonstration Les assertions (i), (ii) et (iii) se prouvent de la même manière. On se donne un point A dans l'ouvert candidat Ω étudié et on cherche un réel $r > 0$ pour lequel $\mathrm{B}(A, r) \subset \Omega$. Je me contenterai ici d'une preuve graphique, mais au fond tout y est.

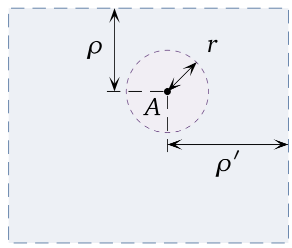
Ici $r = \dfrac{\rho - AB}{2}$.



Ici $r = \dfrac{AB - \rho}{2}$.



Ici $r = \min\{\rho, \rho'\}$.



2.2 Limites et continuité

Définition (Limite (finie) en un point) Soient D une partie de $\mathbb{R}^2$, $f : D \longrightarrow \mathbb{R}$ une fonction, $A \in D$ et $\ell \in \mathbb{R}$. On dit que f *admet ℓ pour limite en A* si :

$$\forall V_\ell \in \mathscr{V}_\ell(\mathbb{R}), \quad \exists\, V_A \in \mathscr{V}_A(\mathbb{R}^2), \quad \forall M \in D \cap V_A, \quad f(M) \in V_\ell,$$

i.e. si : $\quad \forall \varepsilon > 0, \quad \exists\, \alpha > 0, \quad \forall M \in D, \qquad AM < \alpha \quad \Longrightarrow \quad \big|f(M) - \ell\big| < \varepsilon.$

La plupart des résultats qu'on a démontrés pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ sont conservés : unicité et notation $\lim\limits_{M \to A} f(M)$ ou $\lim\limits_{A} f$, caractère localement borné, opérations (combinaison linéaire, produit, inverse, composition avec une fonction de $\mathbb{R}$ dans $\mathbb{R}$), interaction avec les inégalités larges/strictes, théorème d'encadrement... Les preuves sont faciles à faire, il suffit de remplacer les valeurs absolues dans l'ensemble de départ par des distances dans $\mathbb{R}^2$.

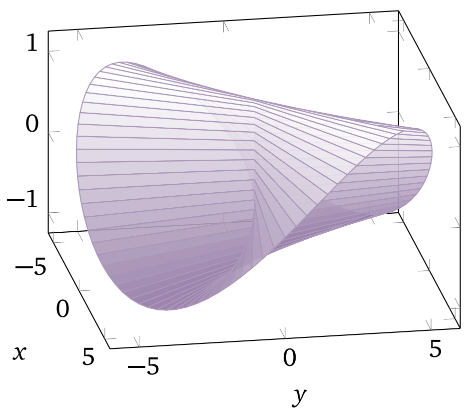
On aurait pu définir aussi des limites de valeur $+\infty$ ou $-\infty$, mais nous n'en aurons pas dans ce chapitre.

Attention ! Faire tendre $M = (x, y)$ vers $A = (x_A, y_A)$ ne revient pas à faire tendre x vers x_A puis y vers y_A ou le contraire. Le point M peut se rapprocher de A de bien des manières, il n'est pas obligé d'y aller en ligne droite selon $\vec{\imath}$ ou $\vec{\jmath}$.

Intéressons-nous par exemple à la fonction $(x, y) \overset{f}{\longmapsto} \dfrac{2xy}{x^2 + y^2}$ sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. Clairement :

$$\lim_{x \to 0} \lim_{y \to 0} \frac{2xy}{x^2 + y^2} = \lim_{x \to 0} 0 = 0 \qquad \text{et} \qquad \lim_{y \to 0} \lim_{x \to 0} \frac{2xy}{x^2 + y^2} = \lim_{y \to 0} 0 = 0,$$

mais nous allons voir que f N'a PAS de limite en $(0, 0)$. Cela se visualise assez bien, mais cela se montre aussi assez bien en *coordonnées polaires* (r, θ).



Tout point $M = (x, y)$ de $\mathbb{R}^2$ peut être écrit sous la forme $(r \cos\theta, r \sin\theta)$ pour un certain $(r, \theta) \in \mathbb{R}_+ \times \mathbb{R}$ dit *couple de coordonnées polaires de M*. Par 2π-périodicité du cosinus et du sinus, un tel couple n'est pas jamais unique, attention ! Trois relations méritent d'être parfaitement comprises : $\quad x = r\cos\theta, \quad y = r \sin\theta \quad$ et $r = \sqrt{x^2 + y^2}$. Le point $(0, 0)$ a la particularité d'admettre le couple $(0, \theta)$ pour coordonnées polaires pour tout $\theta \in \mathbb{R}$.

Mais revenons à notre exemple. Pour tout $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$ de coordonnées polaires (r,θ) :
$$f(x,y) = \frac{2xy}{x^2+y^2} = \frac{2 \times r\cos\theta \times r\sin\theta}{r^2} = \sin(2\theta), \qquad \text{donc } f \text{ NE dépend QUE de la variable } \theta.$$
Par conséquent, si (x,y) tend vers $(0,0)$ en maintenant constant l'angle θ, $f(x,y)$ tend vers le réel $\sin(2\theta)$ pour ce θ fixé. On obtient ainsi des limites différentes selon la manière dont on fait tendre (x,y) vers $(0,0)$, donc en effet, f n'a pas de limite en $(0,0)$. La figure ci-dessus illustre bien l'indépendance de f vis-à-vis de r. Son graphe peut être construit comme un « faisceau tournant » de demi-droites orthogonales à l'axe $\mathrm{Vect}\big(\overrightarrow{k}\big)$.

Définition (Continuité en un point ou sur une partie) Soient D une partie de $\mathbb{R}^2$, $f : D \longrightarrow \mathbb{R}$ une fonction et $A \in D$. On dit que f est *continue en A* si $\lim_{M \to A} f(M) = f(A)$, i.e. si :
$$\forall \varepsilon > 0, \quad \exists\, \alpha > 0, \quad \forall M \in D, \qquad AM < \alpha \quad \Longrightarrow \quad \big|f(M) - f(A)\big| < \varepsilon.$$
L'ensemble des fonctions sur D et à valeurs dans $\mathbb{R}$, i.e. continues en tout point de D, est noté $\mathscr{C}(D,\mathbb{R})$.

Toute combinaison linéaire de fonctions continues est une fonction continue. Même chose avec le produit, l'inverse et la composition avec une fonction de $\mathbb{R}$ dans $\mathbb{R}$.

✖ **Attention !** Pour montrer qu'une fonction de DEUX variables $(x,y) \longmapsto f(x,y)$ est continue en $A = (x_A, y_A)$, il ne suffit pas de montrer que les fonctions d'UNE variable $x \longmapsto f(x, y_A)$ et $y \longmapsto f(x_A, y)$ sont continues respectivement en x_A et y_A. De nouveau, (x,y) peut s'approcher de A de bien des manières, il n'est pas obligé de le faire « à x fixé » ou « à y fixé ». La fonction $(x,y) \overset{f}{\longmapsto} \dfrac{2xy}{x^2+y^2}$ de l'exemple précédent N'est PAS prolongeable par continuité en $(0,0)$, mais pourtant les fonctions $x \longmapsto f(x,0)$ et $y \longmapsto f(0,y)$, identiquement nulles, le sont en 0.

Exemple Les fonctions $(x,y) \longmapsto x$ et $(x,y) \longmapsto y$ sont continues sur $\mathbb{R}^2$. A fortiori, par produit et combinaison linéaire, toute fonction polynomiale de deux variables est continue sur $\mathbb{R}^2$.

Démonstration Démontrons-le pour $(x,y) \overset{f}{\longmapsto} x$. Soient $A = (x_A, y_A) \in \mathbb{R}^2$ et $\varepsilon > 0$.
Pour tout $M = (x,y) \in \mathbb{R}^2$: $\quad \big|f(M) - f(A)\big| = |x - x_A| \leqslant \sqrt{(x-x_A)^2 + (y-y_A)^2} = AM$.
Posons donc $\alpha = \varepsilon$. Pour tout point M : $\quad AM < \alpha \quad \Longrightarrow \quad \big|f(M) - f(A)\big| < \varepsilon$.

Exemple La norme $\overrightarrow{v} \longmapsto \|\overrightarrow{v}\|$ est continue sur $\mathbb{R}^2$.

Démonstration Simple composition de la fonction polynomiale $(x,y) \longmapsto x^2 + y^2$ continue sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}_+$ et de la fonction $t \longmapsto \sqrt{t}$ continue sur $\mathbb{R}_+$.

Exemple Soient I et J deux intervalles, $\varphi \in \mathscr{C}(I,\mathbb{R})$ et $\psi \in \mathscr{C}(J,\mathbb{R})$. Les deux fonctions $(x,y) \longmapsto \varphi(x) + \psi(y)$ et $(x,y) \longmapsto \varphi(x)\psi(y)$ sont continues sur $I \times J$.

Démonstration $(x,y) \longmapsto x$ est continue sur $I \times J$ à valeurs dans I et φ l'est sur I, donc $(x,y) \longmapsto \varphi(x)$ est continue sur $I \times J$ par composition. Même raisonnement pour $(x,y) \longmapsto \psi(y)$. On conclut par somme et produit.

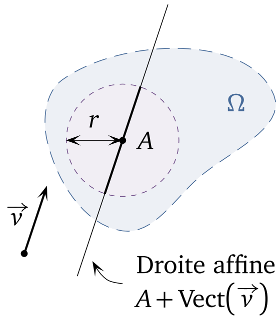
✖ **Attention !** Dans les exemples qui précèdent, on a soigneusement fait appel aux fonctions élémentaires $(x,y) \longmapsto x$ et $(x,y) \longmapsto y$, même pour justifier la continuité d'une fonction aussi simple que $(x,y) \longmapsto \varphi(x)$.

3 Comment « dériver » une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$?

La théorie de la « dérivation » des fonctions en général, pas forcément de $\mathbb{R}$ dans $\mathbb{R}$, est appelée *calcul différentiel*. Je mets des guillemets parce que, comme nous allons le voir, il est beaucoup plus compliqué et subtil de « dériver » une fonction de plusieurs variables qu'une fonction d'une seule variable.

Définition (Dérivée directionnelle) Soient Ω un ouvert de $\mathbb{R}^2$, $f : \Omega \longrightarrow \mathbb{R}$ une fonction, $A \in \Omega$ et $\overrightarrow{v} \in \mathbb{R}^2$. On dit que f est *dérivable en A dans la direction v* si la limite $\lim_{t \to 0} \dfrac{f\big(A + t\overrightarrow{v}\big) - f(A)}{t}$ existe et est finie, i.e. si la fonction $t \overset{F}{\longmapsto} f\big(A + t\overrightarrow{v}\big)$ est dérivable en 0.
Le cas échéant, on appelle *dérivée de f en A dans la direction $\overrightarrow{v}$* le réel $\mathrm{D}_{\overrightarrow{v}} f(A) = F'(0) = \lim_{t \to 0} \dfrac{f\big(A + t\overrightarrow{v}\big) - f(A)}{t}$.

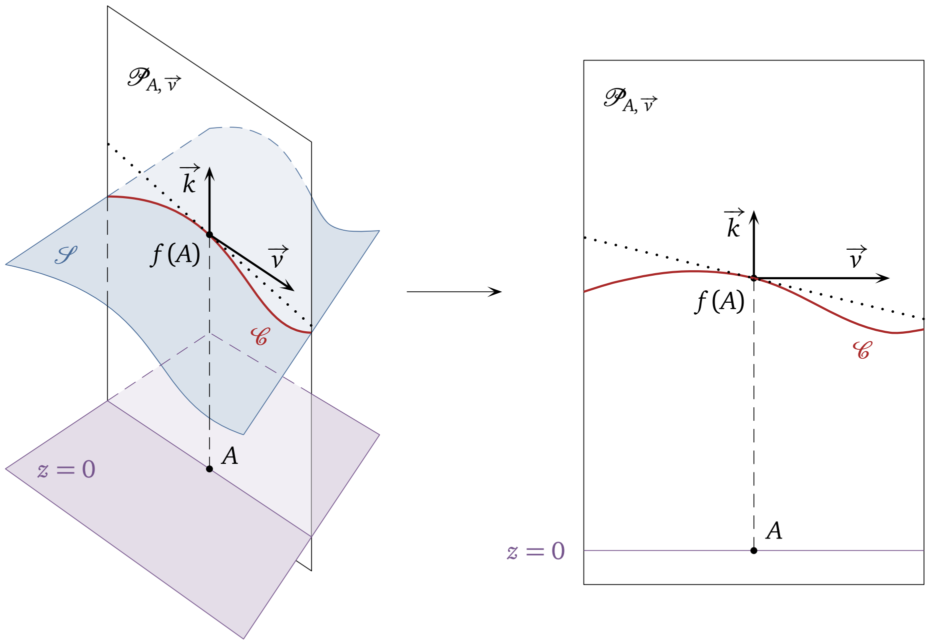
Si $\overrightarrow{v} = \overrightarrow{0} = (0,0)$, la fonction $t \longmapsto f\big(A + t\overrightarrow{v}\big) = f(A)$ est définie et constante sur $\mathbb{R}$ tout entier, donc dérivable en 0 avec $D_{\overrightarrow{0}} f(A) = 0$.



Pour $\overrightarrow{v}$ non nul, la définition précédente n'a de sens que parce que la fonction $t \longmapsto f\big(A + t\overrightarrow{v}\big)$ est **DÉFINIE AU VOISINAGE DE** 0. Cela découle du caractère **OUVERT** de Ω, qui contient $\mathrm{B}(A,r)$ pour un certain $r > 0$. Posons en effet $\alpha = \dfrac{r}{\|\overrightarrow{v}\|}$. Pour tout $t \in \left]-\alpha, \alpha\right[$: $\quad A + tv \in \mathrm{B}(A,r) \subset \Omega \quad$ car $d\big(A, A + t\overrightarrow{v}\big) = \|t\overrightarrow{v}\| = |t|.\|\overrightarrow{v}\| < \alpha\,\|\overrightarrow{v}\| = r$. La fonction $t \longmapsto f\big(A + t\overrightarrow{v}\big)$ est ainsi définie sur $\left]-\alpha, \alpha\right[$.

Et maintenant, que représente géométriquement la dérivée de f en A dans la direction $\overrightarrow{v}$?

1) Considérez dans $\mathbb{R}^3$ le graphe $\mathscr{S}$ de f : $\quad \mathscr{S} = \Big\{(x,y,z) \in \mathbb{R}^3 \;\big|\; \; (x,y) \in \Omega \quad \text{et} \quad z = f(x,y)\Big\}$.
2) Coupez-le sans ménagement par le plan « vertical » $\mathscr{P}_{A,\overrightarrow{v}}$ passant par A dirigé par les deux vecteurs orthogonaux $\overrightarrow{v}$ et $\overrightarrow{k}$. En toute rigueur, $\overrightarrow{v}$ appartient à $\mathbb{R}^2$, mais nous avons identifié $\mathbb{R}^2$ au plan d'équation $z = 0$ de $\mathbb{R}^3$. L'intersection $\mathscr{S} \cap \mathscr{P}_{A,\overrightarrow{v}}$ est une courbe $\mathscr{C}$ du plan $\mathscr{P}_{A,\overrightarrow{v}}$.
3) Oubliez désormais le reste de l'espace et concentrez-vous sur $\mathscr{P}_{A,\overrightarrow{v}}$ et $\mathscr{C}$. Munissez $\mathscr{P}_{A,\overrightarrow{v}}$ du repère $\big(A, \overrightarrow{v}, \overrightarrow{k}\big)$. La courbe $\mathscr{C}$ n'est alors rien de plus que le graphe de la fonction $t \longmapsto f\big(A + t\overrightarrow{v}\big)$ et, s'il existe, $\mathrm{D}_{\overrightarrow{v}} f(A)$ est son nombre dérivé en 0 dans la base $\big(\overrightarrow{v}, \overrightarrow{k}\big)$, i.e. la pente de sa tangente en 0. Notez bien ici que $\overrightarrow{v}$ n'est pas forcément unitaire, il faut en tenir compte quand on se représente la pente.



La définition suivante n'est qu'un cas particulier de la précédente.

Définition (Dérivées partielles) Soient Ω un ouvert de $\mathbb{R}^2$, $f : \Omega \longrightarrow \mathbb{R}$ une fonction et $A \in \Omega$. S'il existe, le réel $\mathrm{D}_{\overrightarrow{i}} f(A)$ est appelé *première dérivée partielle de f en A* et noté $\partial_1 f(A)$.

Même chose pour la notion de *deuxième dérivée partielle*. Le cas échéant, le réel $\mathrm{D}_{\overrightarrow{j}} f(A)$ est noté $\partial_2 f(A)$.

Une fonction f d'**UNE SEULE** variable possède **UNE SEULE** dérivée f' quand elle est dérivable. Il arrive qu'on note $\dfrac{\mathrm{d}f}{\mathrm{d}x}$ cette dérivée, mais cela signifie qu'on a choisi **UNE FOIS POUR TOUTES** de voir f comme une fonction de la variable x. Après un tel choix, il n'est pas possible de noter $\dfrac{\mathrm{d}f}{\mathrm{d}t}$ la dérivée de f.

La situation est la même avec les fonctions de deux variables. Une fonction f de **DEUX** variables possède **DEUX** dérivées partielles quand elle en possède. On peut les noter $\partial_1 f$ et $\partial_2 f$, mais il est courant qu'on choisisse de voir f comme une fonction du couple (x,y). Le cas échéant, $\partial_1 f$ est notée $\dfrac{\partial f}{\partial x}$ et $\partial_2 f$ est notée $\dfrac{\partial f}{\partial y}$. Dans certains contextes, on préférera voir f comme une fonction du couple (u,v), auquel cas ses dérivées partielles sont notées $\dfrac{\partial f}{\partial u}$ et $\dfrac{\partial f}{\partial v}$. Tous les choix sont permis, **MAIS** une fois qu'on a choisi une convention de notation, on s'y tient et on n'accorde pas plus de deux dérivées partielles à f.

Attention ! Pour une fonction d'une seule variable, la notation $\frac{\mathrm{d}f}{\mathrm{d}x}$ s'écrit avec des « d » droits. Pour une fonction de deux variables, les notations $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ s'écrivent au contraire avec des « d » ronds. C'est comme ça...
Par ailleurs, quand on évalue $\frac{\partial f}{\partial x}$ en (x,y), le résultat se note $\frac{\partial f}{\partial x}(x,y)$ ET NON PAS $\frac{\partial f(x,y)}{\partial x}$.

Le calcul concret des dérivées partielles est très facile. En effet, avec des notations évidentes :

$$\frac{\partial f}{\partial x}(A) = \lim_{t\to 0} \frac{f\big(A+t\vec{\imath}\big)-f(A)}{t} = \lim_{t\to 0} \frac{f(x_A+t,y_A)-f(x_A,y_A)}{t},$$

donc calculer $\frac{\partial f}{\partial x}(A)$ revient à fixer y à la valeur y_A dans $f(x,y)$ et à dériver par rapport à x au sens usuel. Pour la même raison, on obtient $\frac{\partial f}{\partial y}$ en dérivant $f(x,y)$ par rapport à y à x fixé.

Exemple La fonction $(x,y) \overset{f}{\longmapsto} e^{xy^2}$ possède des dérivées partielles sur $\mathbb{R}^2$ et pour tout $(x,y)\in\mathbb{R}^2$:

$$\frac{\partial f}{\partial x}(x,y) = y^2\,e^{xy^2} \qquad \text{et} \qquad \frac{\partial f}{\partial y}(x,y) = 2xy\,e^{xy^2}.$$

Le résultat qui suit n'est qu'une adaptation au cas des fonctions de deux variables de résultats bien connus pour les fonctions d'une seule variable. Nous n'avons rien à démontrer car les dérivées partielles ne sont jamais que des dérivées par rapport à une seule variable, l'autre étant gelée.

Théorème (Opérations sur les dérivées partielles)

- **Combinaison linéaire, produit, inverse :** Soient Ω un ouvert de $\mathbb{R}^2$ et $f:\Omega\longrightarrow\mathbb{R}$ et $g:\Omega\longrightarrow\mathbb{R}$ deux fonctions. Si f et g possèdent des dérivées partielles sur Ω, il en va de même de toute combinaison linéaire de f et g et du produit fg, mais aussi de l'inverse $\frac{1}{f}$ si f ne s'annule pas. En outre, pour tous $i\in\{1,2\}$ et $\lambda,\mu\in\mathbb{R}$:
$$\partial_i(\lambda f+\mu g) = \lambda\partial_i f+\mu\partial_i g, \qquad \partial_i(fg) = g\,\partial_i f + f\,\partial_i g \qquad \text{et} \qquad \partial_i\left(\frac{1}{f}\right) = -\frac{\partial_i f}{f^2}.$$
- **Composition :** Soient Ω un ouvert de $\mathbb{R}^2$, I un intervalle, $f:\Omega\longrightarrow I$ et $\varphi:I\longrightarrow\mathbb{R}$ deux fonctions avec $f(\Omega)\subset I$. Si f possède des dérivées partielles sur Ω et si φ est dérivable sur I, $\varphi\circ f$ possède des dérivées partielles sur Ω et pour tout $i\in\{1,2\}$: $\partial_i(\varphi\circ f) = \partial_i f\times\varphi'\circ f$.

Alors que les dérivées partielles sont... justement partielles, le *gradient* d'une fonction f collecte des informations sur les variations de f dans les deux directions $\vec{\imath}$ et $\vec{\jmath}$ et constitue de ce point de vue notre première définition de ce qu'on pourrait avoir envie d'appeler « la dérivée de f ». Nous en donnerons bientôt une interprétation géométrique.

Définition (Gradient) Soient Ω un ouvert de $\mathbb{R}^2$, $f:\Omega\longrightarrow\mathbb{R}$ une fonction et $A\in\Omega$. Si f possède des dérivées partielles en A, on appelle *gradient de f en A* le VECTEUR de $\mathbb{R}^2$:

$$\nabla f(A) = \mathrm{grad} f(A) = \big(\partial_1 f(A)\,,\,\partial_2 f(A)\big) = \partial_1 f(A)\,\vec{\imath} + \partial_2 f(A)\,\vec{\jmath}.$$
Le symbole ∇ se lit « nabla » !

Opérations sur les gradients : $\nabla(\lambda f+\mu g) = \lambda\nabla f+\mu\nabla g$, $\qquad \nabla(fg) = g\nabla f + f\nabla g$,

$$\nabla\left(\frac{1}{f}\right) = -\frac{1}{f^2}\,\nabla f \qquad \text{et} \qquad \nabla(\varphi\circ f) = \varphi'\circ f\times\nabla f.$$

Par exemple, si les variables de f sont notées x et y : $\nabla f(A) = \frac{\partial f}{\partial x}(A)\,\vec{\imath} + \frac{\partial f}{\partial y}(A)\,\vec{\jmath}$.

4 Fonctions de classe $\mathscr{C}^1$

4.1 Définition et premières propriétés

Définition (Fonction de classe $\mathscr{C}^1$) Soient Ω un ouvert de $\mathbb{R}^2$ et $f : \Omega \longrightarrow \mathbb{R}$ une fonction. On dit que f est *de classe $\mathscr{C}^1$ sur Ω* si f possède des dérivées partielles sur Ω et si celles-ci sont continues sur Ω.

L'ensemble des fonctions de classe $\mathscr{C}^1$ sur Ω à valeurs dans $\mathbb{R}$ est noté $\mathscr{C}^1(\Omega, \mathbb{R})$.

Toute combinaison linéaire et tout produit de fonctions de classe $\mathscr{C}^1$ est bien sûr de classe $\mathscr{C}^1$. L'inverse d'une fonction de classe $\mathscr{C}^1$ qui ne s'annule pas aussi. Enfin, pour tout ouvert Ω de $\mathbb{R}^2$ et tout intervalle I, la composée $\varphi \circ f$ d'une fonction $f \in \mathscr{C}^1(\Omega, \mathbb{R})$ avec $f(\Omega) \subset I$ et d'une fonction $\varphi \in \mathscr{C}^1(I, \mathbb{R})$ est une fonction de classe $\mathscr{C}^1$ sur Ω.

Exemple Les fonctions $(x, y) \longmapsto x$ et $(x, y) \longmapsto y$ sont de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$. A fortiori, par produit et combinaison linéaire, toute fonction polynomiale de deux variables est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$.

Démonstration Montrons-le pour $(x, y) \overset{f}{\longmapsto} x$. Pour tout $y \in \mathbb{R}$ (variable gelée), la fonction $x \longmapsto f(x, y) = x$ est dérivable sur $\mathbb{R}$, donc f possède une première dérivée partielle sur $\mathbb{R}^2$ et $\dfrac{\partial f}{\partial x}(x, y) = 1$ pour tout $(x, y) \in \mathbb{R}^2$. De même, pour tout $x \in \mathbb{R}$ (variable gelée), la fonction $y \longmapsto f(x, y) = x$ est dérivable sur $\mathbb{R}$, donc f possède une deuxième dérivée partielle sur $\mathbb{R}^2$ et $\dfrac{\partial f}{\partial y}(x, y) = 0$ pour tout $(x, y) \in \mathbb{R}^2$. Finalement, $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ étant continues sur $\mathbb{R}^2$ car constantes, f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$.

Exemple La norme $\vec{v} \longmapsto \|\vec{v}\|$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2 \setminus \{\vec{0}\}$ — attention, pas en $\vec{0}$!

Démonstration D'abord, $\mathbb{R}^2 \setminus \{\vec{0}\}$ est bien ouvert. Ensuite, la fonction polynomiale $(x, y) \longmapsto x^2 + y^2$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2 \setminus \{\vec{0}\}$ à valeurs dans $\mathbb{R}_+^*$ et la fonction $t \longmapsto \sqrt{t}$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$, donc par composition, $\vec{v} \longmapsto \|\vec{v}\|$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2 \setminus \{\vec{0}\}$.

Et en $\vec{0}$? Pour tout $t \in \mathbb{R}^*$: $\dfrac{\left\|\vec{0} + t\,\vec{\imath}\right\| - \left\|\vec{0}\right\|}{t} = \dfrac{|t|}{t}$ et bien sûr, cette quantité n'a pas de limite quand t tend vers 0, donc $\vec{v} \longmapsto \|\vec{v}\|$ n'a pas de première dérivée partielle en $\vec{0}$.

Exemple La fonction $(x, y) \longmapsto e^{-x} \ln(y)$ est de classe $\mathscr{C}^1$ sur $\mathbb{R} \times \mathbb{R}_+^*$.

Démonstration D'abord, $\mathbb{R} \times \mathbb{R}_+^*$ est ouvert en tant que produit de deux intervalles ouverts. Ensuite, $(x, y) \longmapsto x$ est de classe $\mathscr{C}^1$ sur $\mathbb{R} \times \mathbb{R}_+^*$ et $x \longmapsto e^{-x}$ de classe $\mathscr{C}^1$ sur $\mathbb{R}$, donc $(x, y) \longmapsto e^{-x}$ est de classe $\mathscr{C}^1$ sur $\mathbb{R} \times \mathbb{R}_+^*$ par composition. De même, $(x, y) \longmapsto \ln y$ est de classe $\mathscr{C}^1$ sur $\mathbb{R} \times \mathbb{R}_+^*$ par composition via les fonctions $(x, y) \longmapsto y$ et $y \longmapsto \ln y$. On conclut par produit.

✖ **Attention !** Dans les exemples qui précèdent, on a soigneusement fait appel aux fonctions élémentaires $(x, y) \longmapsto x$ et $(x, y) \longmapsto y$, même pour justifier qu'une fonction aussi simple que $(x, y) \longmapsto e^{-x}$ est de classe $\mathscr{C}^1$.

4.2 Existence de développements limités à l'ordre 1

La suite du chapitre repose en grande partie sur le résultat que voici dont la preuve est explicitement hors programme.

Théorème (Existence de développements limités à l'ordre 1 pour les fonctions de classe $\mathscr{C}^1$) Soient Ω un ouvert de $\mathbb{R}^2$, $f \in \mathscr{C}^1(\Omega, \mathbb{R})$ et $A \in \Omega$. Alors f possède un *développement limité à l'ordre 1 au voisinage de A* :

$$f\left(A + \vec{v}\right) \underset{\vec{v} \to \vec{0}}{=} f(A) + \left\langle \nabla f(A), \vec{v} \right\rangle + o\left(\|\vec{v}\|\right), \qquad \text{ou encore :} \qquad f(M) \underset{M \to A}{=} f(A) + \left\langle \nabla f(A), \overrightarrow{AM} \right\rangle + o\left(\left\|\overrightarrow{AM}\right\|\right).$$

Si les variables de f sont notées x et y et si $A = (x_A, y_A)$, il est équivalent d'affirmer que :

$$f(x_A + h, y_A + k) \underset{(h,k) \to (0,0)}{=} f(A) + \frac{\partial f}{\partial x}(A)\, h + \frac{\partial f}{\partial y}(A)\, k + o\left(\left\|(h, k)\right\|\right).$$

Démonstration On peut supposer sans perte de généralité que A est le point $(0, 0)$ quitte à remplacer f par $(x, y) \longmapsto f(x_A + x, y_A + y)$.

Soit $\varepsilon > 0$. Comme Ω est ouvert, nous pouvons nous donner un réel $r > 0$ pour lequel $\mathrm{B}(A, r) \subset \Omega$. En outre, les fonctions $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ étant continues en $(0,0)$ par hypothèse, il existe des réels $\alpha_x > 0$ et $\alpha_y > 0$, que nous pouvons choisir inférieurs à r, pour lesquels pour tout $\vec{v} = (h,k) \in \mathbb{R}^2$:

$$\|\vec{v}\| < \alpha_x \implies \left|\frac{\partial f}{\partial x}(h,k) - \frac{\partial f}{\partial x}(0,0)\right| < \frac{\varepsilon}{2} \qquad \text{et} \qquad \|\vec{v}\| < \alpha_y \implies \left|\frac{\partial f}{\partial y}(h,k) - \frac{\partial f}{\partial y}(0,0)\right| < \frac{\varepsilon}{2}.$$

Posons $\alpha = \min\{\alpha_x, \alpha_y\}$. Soit ensuite $\vec{v} = (h,k) \in \mathbb{R}^2$ un vecteur pour lequel $\|\vec{v}\| < \alpha$.

$$\begin{aligned}
\left|f(\vec{v}) - f(0,0) - \langle \nabla f(0,0), \vec{v}\rangle\right| &= \left|f(h,k) - f(0,0) - \frac{\partial f}{\partial x}(0,0)\, h - \frac{\partial f}{\partial y}(0,0)\, k\right| \\
&= \left|\big(f(h,k) - f(0,k)\big) - \frac{\partial f}{\partial x}(0,0)\, h + \big(f(0,k) - f(0,0)\big) - \frac{\partial f}{\partial y}(0,0)\, k\right| \\
&= \left|\int_0^h \frac{\partial f}{\partial x}(s,k)\, \mathrm{d}s - \int_0^h \frac{\partial f}{\partial x}(0,0)\, \mathrm{d}s + \int_0^k \frac{\partial f}{\partial y}(0,t)\, \mathrm{d}t - \int_0^k \frac{\partial f}{\partial y}(0,0)\, \mathrm{d}t\right| \qquad \text{d'après le théorème fondamental de l'analyse} \\
&\leqslant \left|\int_0^h \left(\frac{\partial f}{\partial x}(s,k) - \frac{\partial f}{\partial x}(0,0)\right) \mathrm{d}s\right| + \left|\int_0^k \left(\frac{\partial f}{\partial y}(0,t) - \frac{\partial f}{\partial y}(0,0)\right) \mathrm{d}t\right|.
\end{aligned}$$

Or pour tout s compris entre 0 et h : $\big\|(s,k)\big\| = \sqrt{s^2+k^2} \leqslant \sqrt{h^2+k^2} = \|\vec{v}\| < \alpha \leqslant \alpha_x$ et pour tout t compris entre 0 et k : $\big\|(0,t)\big\| = |t| \leqslant |k| \leqslant \sqrt{h^2+k^2} = \|\vec{v}\| < \alpha \leqslant \alpha_y$, donc :

$$\left|f(\vec{v}) - f(0,0) - \langle \nabla f(0,0), \vec{v}\rangle\right| \leqslant \frac{\varepsilon}{2}|h| + \frac{\varepsilon}{2}|k| \leqslant \frac{\varepsilon}{2}\|\vec{v}\| + \frac{\varepsilon}{2}\|\vec{v}\| = \varepsilon\|\vec{v}\|.$$

En résumé, nous avons trouvé un réel $\alpha > 0$ pour lequel pour tout $\vec{v} \in \mathbb{R}^2 \setminus \{\vec{0}\}$:

$$\|\vec{v}\| < \alpha \implies \left|f(\vec{v}) - f(0,0) - \langle \nabla f(0,0), \vec{v}\rangle\right| \leqslant \varepsilon\|\vec{v}\|.$$

Comme voulu : $f(\vec{v}) - f(0,0) - \langle \nabla f(0,0), \vec{v}\rangle \underset{\vec{v}\to\vec{0}}{=} \mathrm{o}(\|\vec{v}\|)$.

Théorème (Continuité d'une fonction de classe $\mathscr{C}^1$) Soient Ω un ouvert de $\mathbb{R}^2$ et $f \in \mathscr{C}^1(\Omega, \mathbb{R})$. Alors $f \in \mathscr{C}(\Omega, \mathbb{R})$.

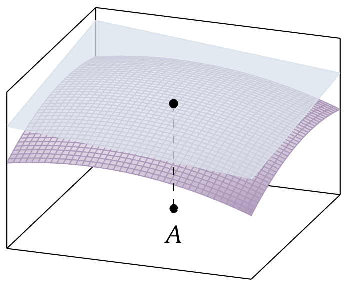
Démonstration Pour tout $A \in \Omega$, par simple troncature du développement limité à l'ordre 1 de f en A : $\lim\limits_{M\to A} f(M) = f(A)$.

Il peut sembler curieux qu'on dispose d'une notion de continuité et d'une notion de classe $\mathscr{C}^1$, mais pas d'une notion intermédiaire de dérivabilité. Le chaînon manquant s'appelle la *différentiabilité* et ne figure pas au programme de MPSI, mais je lui consacrerai un petit paragraphe hors programme en fin de chapitre.

Définition (Plan tangent) Soient Ω un ouvert de $\mathbb{R}^2$, $f \in \mathscr{C}^1(\Omega, \mathbb{R})$ et $A = (x_A, y_A) \in \Omega$. Le plan de $\mathbb{R}^3$ d'équation :

$$z = f(A) + \frac{\partial f}{\partial x}(A)(x - x_A) + \frac{\partial f}{\partial y}(A)(y - y_A) \qquad \text{est appelé le } \textit{plan tangent de } f \textit{ en } A.$$

Nous savons depuis longtemps qu'une courbe suffisamment régulière ressemble localement à une droite. Sans surprise, toute surface suffisamment régulière ressemble localement à un plan. Pour une fonction φ d'une seule variable, l'équation de la tangente en a s'exprimait naturellement en termes de développement limité à l'ordre 1 : $\varphi(x) \underset{x\to a}{=} \varphi(a) + \varphi'(a)(x-a) + \mathrm{o}(x-a)$. La situation est finalement la même pour les fonctions de deux variables.



Exemple La fonction $(x,y) \overset{f}{\longmapsto} \sin(x+2y)$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ et son plan tangent en $(0,0)$ a pour équation $z = x + 2y$.

Démonstration Polynomiale, la fonction $(x,y) \longmapsto x + 2y$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ et la fonction $t \longmapsto \sin t$ l'est sur $\mathbb{R}$, donc f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ par composition.

Ensuite, pour tout $(x,y) \in \mathbb{R}^2$: $\dfrac{\partial f}{\partial x}(x,y) = \cos(x+2y)$ et $\dfrac{\partial f}{\partial y}(x,y) = 2\cos(x+y)$, donc le plan tangent de f en $(0,0)$ a donc pour équation $z = f(0,0) + \dfrac{\partial f}{\partial x}(0,0)\, x + \dfrac{\partial f}{\partial y}(0,0)\, y = x + 2y$.

4.3 Règle de la chaîne

La maîtrise pratique du résultat qui suit est l'un des objectifs majeurs de ce chapitre.

Théorème (Règle de la chaîne) Soient Ω un ouvert de $\mathbb{R}^2$, I un intervalle, $f \in \mathscr{C}^1(\Omega, \mathbb{R})$ et enfin $x, y \in \mathscr{C}^1(I, \mathbb{R})$ deux fonctions pour lesquelles pour tout $t \in I$: $\big(x(t), y(t)\big) \in \Omega$.

La fonction $t \overset{F}{\longmapsto} f\big(x(t), y(t)\big)$ est de classe $\mathscr{C}^1$ sur I et pour tout $t \in I$:

$$F'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\Big(f\big(x(t), y(t)\big)\Big) = x'(t)\,\frac{\partial f}{\partial x}\big(x(t), y(t)\big) + y'(t)\,\frac{\partial f}{\partial y}\big(x(t), y(t)\big).$$

Dans cet énoncé, f est perçue comme une fonction des variables x et y, mais attention, ce x et ce y ne sont pas les fonctions $t \longmapsto x(t)$ et $t \longmapsto y(t)$ de l'énoncé, ce sont juste des noms de variables qui fixent la notation $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ des dérivées partielles.

Démonstration Notons γ la fonction $t \longmapsto \big(x(t), y(t)\big)$, de sorte que $F = f \circ \gamma$. Soit $t \in I$. Composons le développement limité de f au voisinage de $\gamma(t)$ à l'ordre 1 avec la limite $\lim\limits_{h \to 0} \big(\gamma(t+h) - \gamma(t)\big) = (0,0)$.

$$\begin{aligned} F(t+h) &= f\big(\gamma(t+h)\big) = f\Big(\gamma(t) + \big(\gamma(t+h) - \gamma(t)\big)\Big) \\ &\underset{h \to 0}{=} f\big(\gamma(t)\big) + \big\langle \nabla f\big(\gamma(t)\big)\,,\, \big(\gamma(t+h) - \gamma(t)\big)\big\rangle + \mathrm{o}\Big(\big\|\gamma(t+h) - \gamma(t)\big\|\Big). \end{aligned}$$

Or : $\big\|\gamma(t+h) - \gamma(t)\big\| = \sqrt{\big(x(t+h) - x(t)\big)^2 + \big(y(t+h) - y(t)\big)^2} \underset{h \to 0}{=} |h|\sqrt{x'(t)^2 + y'(t)^2} + \mathrm{o}(h)$, donc on peut remplacer le $\mathrm{o}\Big(\big\|\gamma(t+h) - \gamma(t)\big\|\Big)$ par un $\mathrm{o}(h)$:

$$\begin{aligned} F(t+h) &\underset{h \to 0}{=} F(t) + \frac{\partial f}{\partial x}\big(\gamma(t)\big)\,\big(x(t+h) - x(t)\big) + \frac{\partial f}{\partial y}\big(\gamma(t)\big)\,\big(y(t+h) - y(t)\big) + \mathrm{o}(h) \\ &\underset{h \to 0}{=} F(t) + \frac{\partial f}{\partial x}\big(\gamma(t)\big)\,\big(hx'(t) + \mathrm{o}(h)\big) + \frac{\partial f}{\partial y}\big(\gamma(t)\big)\,\big(hy'(t) + \mathrm{o}(h)\big) + \mathrm{o}(h) \\ &\underset{h \to 0}{=} F(t) + \left(x'(t)\,\frac{\partial f}{\partial x}\big(\gamma(t)\big) + y'(t)\,\frac{\partial f}{\partial y}\big(\gamma(t)\big)\right) h + \mathrm{o}(h). \end{aligned}$$

Ce développement limité prouve que F est dérivable en t et que $F'(t) = x'(t)\,\dfrac{\partial f}{\partial x}\big(\gamma(t)\big) + y'(t)\,\dfrac{\partial f}{\partial y}\big(\gamma(t)\big)$. Pour finir, l'expression de F' ainsi obtenue est continue sur I, donc F y est de classe $\mathscr{C}^1$. La petite subtilité de l'affaire, c'est que γ est à valeurs dans $\mathbb{R}^2$ et non pas $\mathbb{R}$, mais ses composantes x et y sont continues toutes les deux. ■

Exemple Soit $f \in \mathscr{C}^1\big(\mathbb{R}^2, \mathbb{R}\big)$. La fonction $t \overset{F}{\longmapsto} f\big(t^2, \sin t\big)$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}$ car $t \longmapsto t^2$ et $t \longmapsto \sin t$ le sont sur $\mathbb{R}$, et pour tout $t \in \mathbb{R}$: $F'(t) = 2t\,\dfrac{\partial f}{\partial x}\big(t^2, \sin t\big) + \cos t\,\dfrac{\partial f}{\partial y}\big(t^2, \sin t\big)$.

L'exemple qui suit mérite une lecture très attentive.

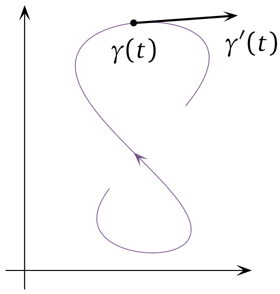
Exemple Pour tout $\alpha \in \mathbb{R}$ et toute fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, on dit que f est *homogène de degré* α si pour tous $(x, y) \in \mathbb{R}^2$ et $\lambda > 0$: $f(\lambda x, \lambda y) = \lambda^\alpha f(x, y)$ $\bigstar$.

Fixons une fois pour toutes $\alpha \in \mathbb{R}$ et $f \in \mathscr{C}^1\big(\mathbb{R}^2, \mathbb{R}\big)$ homogène de degré α. Nous noterons par convention x et y les variables de f. Les dérivées partielles de f sont donc les deux fonctions $\partial_1 f = \dfrac{\partial f}{\partial x}$ et $\partial_2 f = \dfrac{\partial f}{\partial y}$ et f n'a pas d'autre dérivée partielle. La relation $\bigstar$ fait pourtant appel à trois variables x, y et λ et on peut la dériver par rapport à chacune de ces variables. Attention, on va dériver par rapport à λ, mais la notation $\dfrac{\partial f}{\partial \lambda}$ n'a aucun sens !

- **Dérivation par rapport à x :** Pour dériver $f(\lambda x, \lambda y)$ par rapport à x, on adapte ce qu'on sait faire pour dériver les expressions de la forme $f\big(x(t), y(t)\big)$ par rapport à t. Cela donne :
$$\lambda \times \frac{\partial f}{\partial x}(\lambda x, \lambda y) + 0 \times \frac{\partial f}{\partial y}(\lambda x, \lambda y) = \lambda^\alpha\,\frac{\partial f}{\partial x}(x, y), \qquad \text{i.e. } \frac{\partial f}{\partial x}(\lambda x, \lambda y) = \lambda^{\alpha-1}\,\frac{\partial f}{\partial x}(x, y).$$
Cette égalité vraie pour tous $(x, y) \in \mathbb{R}^2$ et $\lambda > 0$ montre que $\dfrac{\partial f}{\partial x}$ est homogène de degré $\alpha - 1$.

- **Dérivation par rapport à y :** Même calcul qu'à l'instant, x et y jouant des rôles symétriques dans la relation $\bigstar$: $\dfrac{\partial f}{\partial y}(\lambda x, \lambda y) = \lambda^{\alpha-1}\,\dfrac{\partial f}{\partial y}(x, y)$, donc $\dfrac{\partial f}{\partial y}$ est homogène de degré $\alpha - 1$.

- **Dérivation par rapport à λ :** Pour dériver $f(\lambda x, \lambda y)$ par rapport à λ, on adapte ce qu'on sait faire pour dériver les expressions de la forme $f\big(x(t), y(t)\big)$ par rapport à t. Cela donne :
$$x\,\frac{\partial f}{\partial x}(\lambda x, \lambda y) + y\,\frac{\partial f}{\partial y}(\lambda x, \lambda y) = \alpha\lambda^{\alpha-1} f(x,y),$$
mais $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont homogènes de degré $\alpha-1$, donc on peut simplifier un peu : $\quad x\,\dfrac{\partial f}{\partial x} + y\,\dfrac{\partial f}{\partial y} = \alpha f.\quad$ Une telle relation qui lie f à ses dérivées partielles est appelée une *équation aux dérivées partielles*. Les équations aux dérivées partielles sont aux fonctions de plusieurs variables ce que les équations différentielles sont aux fonctions d'une seule variable et la physique en est truffée.



La règle de la chaîne est maintenant acquise, mais que raconte-t-elle géométriquement ? C'est l'objet du théorème que voici, qui éclaire de plus les notions de dérivée directionnelle et gradient.

Mais d'abord, un détail technique. Nous avons travaillé avec des fonctions à valeurs dans $\mathbb{R}$ jusqu'ici, mais rien n'empêche qu'on utilise à l'occasion des fonctions à valeurs dans $\mathbb{R}^2$. Pour tout intervalle I et toutes fonctions $x, y \in \mathscr{C}^1(I,\mathbb{R})$, on considèrera par définition que la fonction $t \overset{\gamma}{\longmapsto} \big(x(t), y(t)\big)$ est *de classe $\mathscr{C}^1$ sur I* et on notera γ' sa *dérivée*, i.e. la fonction $t \longmapsto \big(x'(t), y'(t)\big)$. Géométriquement, γ est ce qu'on appelle une *courbe paramétrée* ou un *arc paramétré* tracé dans le plan $\mathbb{R}^2$ et γ' est en tout point le vecteur vitesse du mobile γ.

Théorème (Règle de la chaîne, dérivées directionnelles et gradient pour les fonctions de classe $\mathscr{C}^1$) Soient Ω un ouvert de $\mathbb{R}^2$ et $f \in \mathscr{C}^1(\Omega,\mathbb{R})$.

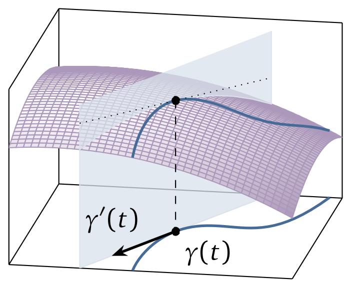
(i) **Règle de la chaîne deuxième version :** Soient I un intervalle et $\gamma \in \mathscr{C}^1\big(I,\mathbb{R}^2\big)$ une fonction pour laquelle $\gamma(I) \subset \Omega$. La fonction $f \circ \gamma$ est de classe $\mathscr{C}^1$ sur I et pour tout $t \in I$: $\quad (f\circ\gamma)'(t) = \big\langle \nabla f\big(\gamma(t)\big)\,,\, \gamma'(t)\big\rangle$.

(ii) **Dérivées directionnelles :** La fonction f possède une dérivée en tout point de Ω dans toutes les directions. Plus précisément, pour tous $A \in \Omega$ et $\vec{v} = (h,k) \in \mathbb{R}^2$:
$$\mathrm{D}_{\vec{v}} f(A) = \langle \nabla f(A)\,,\, \vec{v}\rangle = \frac{\partial f}{\partial x}(A)\,h + \frac{\partial f}{\partial y}(A)\,k.$$

(iii) **Interprétation géométrique du gradient :** Le gradient de f est orthogonal aux lignes de niveau de f et dirigé dans le sens des pentes croissantes.

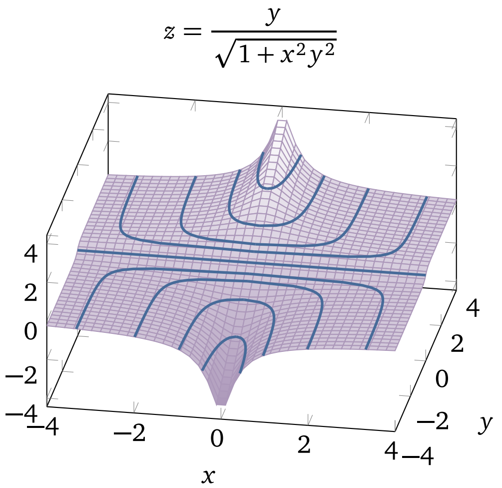
Par définition, une fonction de deux variables est de classe $\mathscr{C}^1$ quand elle possède des dérivées CONTINUES dans les deux directions privilégiées $\vec{\imath}$ et $\vec{\jmath}$. D'après l'assertion (ii), cette hypothèse garantit en fait l'existence de dérivées dans TOUTES les directions que l'on peut calculer facilement à partir des dérivées partielles ou du gradient.

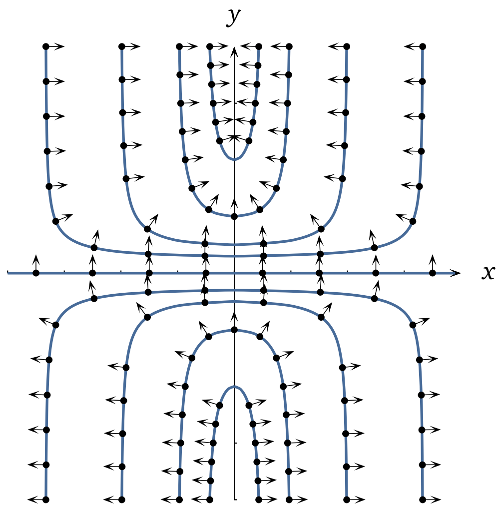
Pour une fonction f d'une seule variable, vous savez depuis longtemps que le réel $f'(a)$ est la pente de la tangente de f en a. Pour vous le faire comprendre, on vous a sans doute dit que « quand on avance de 1 vers la droite en abscisse, on monte de $f'(a)$ en ordonnée sur la tangente ». La situation est la même pour une fonction f de deux variables. Au voisinage de A, le graphe de f a l'allure de son plan tangent et $\mathrm{D}_{\vec{v}} f(A)$ n'est jamais que la pente de ce plan dans la direction $\vec{v}$. L'assertion (ii) énoncé que quand on avance de h dans la direction $\vec{\imath}$ et de k dans la direction $\vec{\jmath}$, on monte de $\dfrac{\partial f}{\partial x}(A)\,h + \dfrac{\partial f}{\partial y}(A)\,k$ dans la direction $\vec{k}$ sur le plan tangent.



L'assertion (ii) nous permet en retour d'interpréter géométriquement la règle de la chaîne. La relation : $\quad (f\circ\gamma)'(t) = \mathrm{D}_{\gamma'(t)} f\big(\gamma(t)\big) \quad$ montre que $(f\circ\gamma)'(t)$ est une dérivée directionnelle de f, en l'occurrence la dérivée de f en $\gamma(t)$ dans la direction $\gamma'(t)$. La figure ci-contre illustre le phénomène. Le mobile γ évolue dans le plan $\mathbb{R}^2$, mais on peut s'intéresser à sa projection verticale sur le graphe de f. Le résultat est un nouvel arc paramétré, à trois dimensions cette fois et entièrement contenu dans le graphe de f. Sur la figure proposée, la pente de la droite en pointillé dans le repère $\big(\gamma(t), \gamma'(t), \vec{k}\big)$ vaut à la fois $(f\circ\gamma)'(t)$ et $\mathrm{D}_{\gamma'(t)} f\big(\gamma(t)\big)$.

Pour finir, l'assertion (iii) se visualise très bien et les figures ci-dessous vaudront mieux qu'un long discours. On y a représenté les lignes de niveau λ pour $\lambda \in \left\{2, 1, \dfrac{1}{2}, \dfrac{3}{10}, 0, -\dfrac{3}{10}, -\dfrac{1}{2}, -1, -2\right\}$. ATTENTION CEPENDANT, les gradients ont été ramenés de force à la même norme sur la figure de droite pour une raison esthétique.





Démonstration

(i) Simple réécriture de la règle de la chaîne.

(ii) Comme Ω est ouvert, la fonction $t \overset{F}{\longmapsto} f\big(A+t\vec{v}\big)$ est définie sur $]-\alpha,\alpha[$ pour un certain $\alpha>0$ et elle y est de classe $\mathscr{C}^1$ d'après la règle de la chaîne avec $F'(0) = \dfrac{\partial f}{\partial x}(A)\,h + \dfrac{\partial f}{\partial y}(A)\,k$. Le résultat en découle car $F'(0) = \mathrm{D}_{\vec{v}}f(A)$ par définition des dérivées directionnelles.

(iii) Fixons $\lambda \in \mathbb{R}$ et notons $\mathscr{C}_\lambda$ la courbe de niveau λ de f : $\mathscr{C}_\lambda = \Big\{(x,y)\in\mathbb{R}^2 \;|\; f(x,y)=\lambda\Big\}$. Pour montrer que ∇f est orthogonal à $\mathscr{C}_\lambda$, nous pouvons supposer $\mathscr{C}_\lambda$ non vide. Faisons aussi l'hypothèse — pas vraie en toute généralité sous cette forme, mais passons — que $\mathscr{C}_\lambda$ peut être paramétrée par une courbe paramétrée $\gamma=(x,y)$ de classe $\mathscr{C}^1$ sur un intervalle I dont la dérivée ne s'annule pas. Dans ces conditions, pour tout $M\in\mathbb{R}^2$: $M\in\mathscr{C}_\lambda \iff \exists\, t\in I,\quad M=\gamma(t)$, donc $f\circ\gamma(t)=\lambda$ pour tout $t\in I$. En retour, d'après la règle de la chaîne : $\Big\langle \nabla f\big(\gamma(t)\big)\,,\,\gamma'(t)\Big\rangle = 0$. Cette égalité signifie que $\nabla f\big(\gamma(t)\big)$ est orthogonal au vecteur $\gamma'(t)$ qui dirige la tangente de γ en t. C'est cela qu'on veut dire quand on affirme que le gradient est orthogonal aux lignes de niveau.

Le théorème qui suit n'est rien de plus qu'une nouvelle application de la règle de la chaîne. On y manipule cette fois des fonctions de $\mathbb{R}^2$ dans $\mathbb{R}^2$ et, de nouveau, la dérivée d'une telle fonction se calcule composante par composante.

Théorème (Composition d'une fonction de $\mathbb{R}^2$ dans $\mathbb{R}^2$ suivie d'une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$) Soient U et Ω deux ouverts de $\mathbb{R}^2$, $x,y\in\mathscr{C}^1(U,\mathbb{R})$ et $f\in\mathscr{C}^1(\Omega,\mathbb{R})$. On suppose que $\big(x(u,v),y(u,v)\big)\in\Omega$ pour tout $(u,v)\in U$ et on pose $F(u,v)=f\big(x(u,v),y(u,v)\big)$. La fonction F est de classe $\mathscr{C}^1$ sur U et pour tout $(u,v)\in U$:

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

et : $$\frac{\partial F}{\partial v}(u,v) = \frac{\partial}{\partial v}\Big(f\big(x(u,v),y(u,v)\big)\Big) = \frac{\partial x}{\partial v}(u,v)\,\frac{\partial f}{\partial x}\big(x(u,v),y(u,v)\big) + \frac{\partial y}{\partial v}(u,v)\,\frac{\partial f}{\partial y}\big(x(u,v),y(u,v)\big).$$

En résumé : $$\frac{\partial F}{\partial u} = \frac{\partial x}{\partial u}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial v}\frac{\partial f}{\partial y} \qquad \text{et} \qquad \frac{\partial F}{\partial v} = \frac{\partial x}{\partial v}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial v}\frac{\partial f}{\partial y}.$$

Dans cet énoncé, les fonctions x et y sont perçues comme des fonctions des variables u et v et f comme une fonction de x et y, mais attention, ce x et ce y sont juste des noms de variables et non pas les fonctions $(u,v)\longmapsto x(u,v)$ et $(u,v)\longmapsto y(u,v)$.

Démonstration La règle de la chaîne nous permet de dériver la fonction $u\longmapsto F(u,v)$ à v fixé et la fonction $v\longmapsto F(u,v)$ à u fixé. Il se trouve enfin que les expressions obtenues de $\dfrac{\partial F}{\partial u}$ et $\dfrac{\partial F}{\partial v}$ sont continues sur U, donc F y est de classe $\mathscr{C}^1$.

Exemple Soit $f\in\mathscr{C}^1\big(\mathbb{R}^2,\mathbb{R}\big)$. On pose pour tout $(u,v)\in\mathbb{R}^2$: $F(u,v)=f\big(u^2+v,uv-v^2\big)$. La fonction F est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ car les fonctions $(u,v)\longmapsto u^2+v$ et $(u,v)\longmapsto uv-v^2$ le sont et pour tout $(u,v)\in\mathbb{R}^2$:

$$\frac{\partial F}{\partial u}(u,v) = 2u\,\frac{\partial f}{\partial x}\big(u^2+v,uv-v^2\big)+v\,\frac{\partial f}{\partial y}\big(u^2+v,uv-v^2\big) \qquad \text{et} \qquad \frac{\partial F}{\partial v}(u,v) = \frac{\partial f}{\partial x}\big(u^2+v,uv-v^2\big)+(u-2v)\,\frac{\partial f}{\partial y}\big(u^2+v,uv-v^2\big).$$

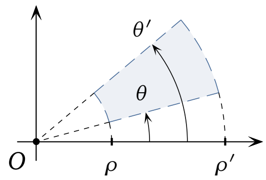
Le passage d'un problème posé en termes de coordonnées cartésiennes à un problème posé en termes de coordonnées polaires est particulièrement courant. Rappelons donc que tout point (x,y) de $\mathbb{R}^2$ peut être écrit sous la forme $(r\cos\theta, r\sin\theta)$ pour un certain couple $(r,\theta)\in\mathbb{R}_+\times\mathbb{R}$. Trois relations à connaître : $x=r\cos\theta$, $y=r\sin\theta$ et $r=\sqrt{x^2+y^2}$. On sait aussi exprimer θ en fonction de x et y à l'aide des arctangente, arcsinus et arccosinus, mais cela demande plus de doigté.

La plupart du temps, on autorise en réalité r à être négatif quand on parle de coordonnées polaires. Ainsi, si (r,θ) est un couple de coordonnées polaires de (x,y) avec $r\geqslant 0$: $x=r\cos\theta=(-r)\cos(\theta+\pi)$ et $y=r\sin\theta=(-r)\sin(\theta+\pi)$, donc $(-r,\theta+\pi)$ est un autre couple de coordonnées polaires de (x,y). Faites un dessin !

Notons à présent x la fonction $(r,\theta)\longmapsto r\cos\theta$ et y la fonction $(r,\theta)\longmapsto r\sin\theta$, toutes deux de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$. La fonction $(r,\theta)\overset{\varphi}{\longmapsto}\big(x(r,\theta),y(r,\theta)\big)=(r\cos\theta,r\sin\theta)$ est quant à elle de classe $\mathscr{C}^1$ de $\mathbb{R}^2$ dans $\mathbb{R}^2$ et c'est grâce à elle qu'on passe d'une expression en coordonnées cartésiennes à une expression en coordonnées polaires. Par exemple, en notant O le point $(0,0)$, φ transforme :

— la bande horizontale $\Omega_{\text{pol}}=\,]-\rho,\rho[\,\times\mathbb{R}$ en la boule ouverte $\Omega_{\text{cart}}=\mathrm{B}(O,\rho)$ pour tout $\rho>0$,

— le rectangle $\Omega_{\text{pol}}=\,]\rho,\rho'[\,\times\,]\theta,\theta'[$ en la portion de disque Ω_{cart} représentée ci-contre pour tous $\rho,\rho',\theta,\theta'\in\mathbb{R}$ pour lesquels $0<\rho<\rho'$ et $\theta<\theta'$.



Donnons-nous à présent une fonction $f\in\mathscr{C}^1$ de variables x et y sur un ouvert Ω_{cart} de $\mathbb{R}^2$. Notons ensuite Ω_{pol} l'image réciproque de Ω_{cart} par φ — on peut montrer que Ω_{pol} est un ouvert de $\mathbb{R}^2$. L'ouvert Ω_{pol} est la version polaire de l'ouvert Ω_{cart}. Posons enfin pour tout $(r,\theta)\in\Omega_{\text{pol}}$: $F(r,\theta)=f(r\cos\theta,r\sin\theta)$. La fonction F ainsi définie est la version polaire de f. Un(e) physicien(ne) donnerait le même nom aux deux fonctions car elles représentent pour lui le même phénomène physique — force, température, champ magnétique… — exprimé simplement de deux façons différentes. Pour nous autres matheux, f et F sont bel et bien deux fonctions différentes définies sur des ensembles différents — mais de même image.

La fonction F est de classe $\mathscr{C}^1$ sur Ω_{pol} et pour tous $(x,y)\in\Omega_{\text{cart}}$ et $(r,\theta)\in\Omega_{\text{pol}}$ liés par la relation $(x,y)=\varphi(r,\theta)$:

$$\frac{\partial F}{\partial r}(r,\theta)=\cos\theta\,\frac{\partial f}{\partial x}(r\cos\theta,r\sin\theta)+\sin\theta\,\frac{\partial f}{\partial y}(r\cos\theta,r\sin\theta)=\cos\theta\,\frac{\partial f}{\partial x}(x,y)+\sin\theta\,\frac{\partial f}{\partial y}(x,y)$$

donc : $\boxed{r\,\frac{\partial F}{\partial r}(r,\theta)=x\,\frac{\partial f}{\partial x}(x,y)+y\,\frac{\partial f}{\partial y}(x,y)}$ après multiplication par r. De même :

$$\frac{\partial F}{\partial \theta}(r,\theta)=-r\sin\theta\,\frac{\partial f}{\partial x}(r\cos\theta,r\sin\theta)+r\cos\theta\,\frac{\partial f}{\partial y}(r\cos\theta,r\sin\theta),\quad\text{donc :}\quad\boxed{\frac{\partial F}{\partial \theta}(r,\theta)=-y\,\frac{\partial f}{\partial x}(x,y)+x\,\frac{\partial f}{\partial y}(x,y).}$$

Exemple On s'intéresse à l'équation aux dérivées partielles $y\,\frac{\partial f}{\partial x}=x\,\frac{\partial f}{\partial y}$ d'inconnue $f\in\mathscr{C}^1(\Omega,\mathbb{R})$ où $\Omega=\mathbb{R}^2\setminus\big\{(0,0)\big\}$, avec la *condition aux limites* $f(x,0)=\mathrm{e}^{-x}$ pour tout $x>0$. Ce type de condition joue le même rôle que les conditions initiales d'une équation différentielle.

- **Analyse :** Soit $f\in\mathscr{C}^1(\Omega,\mathbb{R})$. Posons pour tout $(r,\theta)\in\mathbb{R}_+^*\times\mathbb{R}$: $F(r,\theta)=f(r\cos\theta,r\sin\theta)$. La fonction F ainsi définie est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*\times\mathbb{R}$ et pour tous $(x,y)\in\Omega$ et $(r,\theta)\in\mathbb{R}_+^*\times\mathbb{R}$ liés par la relation $(x,y)=(r\cos\theta,r\sin\theta)$, d'après les calculs qui précèdent : $\frac{\partial F}{\partial\theta}(r,\theta)=-y\,\frac{\partial f}{\partial x}(x,y)+x\,\frac{\partial f}{\partial y}(x,y)=0$. Ainsi, POUR TOUT $r>0$, la fonction $\theta\longmapsto F(r,\theta)$ est constante sur $\mathbb{R}$, disons de valeur $\varphi(r)$ DÉPENDANTE DE r. Cela dit, F est de classe $\mathscr{C}^1$ en tant que fonction de deux variables, elle l'est donc aussi par rapport à r quand on gèle la variable θ — c'est faux dans l'autre sens, ATTENTION. Conclusion : φ est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$. Et pour tout $(x,y)\in\Omega$: $f(x,y)=\varphi\big(\sqrt{x^2+y^2}\big)$, donc pour tout $x>0$: $\varphi(x)=\varphi\big(\sqrt{x^2+0^2}\big)=f(x,0)=\mathrm{e}^{-x}$.

- **Synthèse :** Notons f la fonction $(x,y)\longmapsto\mathrm{e}^{-\sqrt{x^2+y^2}}$ sur Ω. La fonction polynomiale $(x,y)\longmapsto x^2+y^2$ est de classe $\mathscr{C}^1$ sur Ω et à valeurs dans $\mathbb{R}_+^*$ et la fonction $t\longmapsto\mathrm{e}^{-\sqrt{t}}$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$, donc f est de classe $\mathscr{C}^1$ sur Ω et pour tout $(x,y)\in\Omega$: $y\,\frac{\partial f}{\partial x}(x,y)=y\times\frac{-x}{\sqrt{x^2+y^2}}\,\mathrm{e}^{-\sqrt{x^2+y^2}}=x\times\frac{-y}{\sqrt{x^2+y^2}}\,\mathrm{e}^{-\sqrt{x^2+y^2}}=x\,\frac{\partial f}{\partial y}(x,y)$. Enfin, pour tout $x>0$: $f(x,0)=\mathrm{e}^{-\sqrt{x^2+0^2}}=\mathrm{e}^{-x}$.

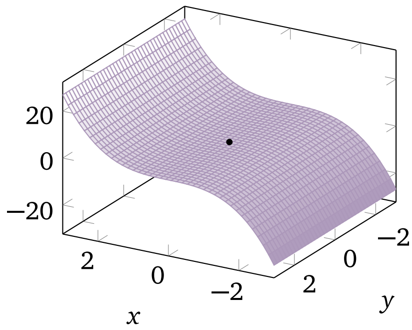
La fonction $(x,y)\longmapsto\mathrm{e}^{-\sqrt{x^2+y^2}}$ est donc la seule solution du problème étudié, mais notez bien que sans condition aux limites, le problème aurait eu ÉNORMÉMENT de solutions — toutes les fonctions $(x,y)\longmapsto\varphi\big(\sqrt{x^2+y^2}\big)$, φ décrivant $\mathscr{C}^1(\mathbb{R}_+^*,\mathbb{R})$.

4.4 EXTREMA LOCAUX

Définition **(Extremum local)** Soient D une partie de $\mathbb{R}^2$, $f : D \longrightarrow \mathbb{R}$ une fonction et $A \in D$. On dit que f possède un *maximum local en* A si f est majorée par $f(A)$ au voisinage de A, et un *minimum local en* A si elle est minorée par $f(A)$ au voisinage de A.

Théorème **(Extrema locaux d'une fonction de classe $\mathscr{C}^1$)** Soient Ω un ouvert de $\mathbb{R}^2$, $f \in \mathscr{C}^1(\Omega, \mathbb{R})$ et $A \in \Omega$. Si f possède en A un extremum local en A, alors A est un *point critique de* f, autrement dit $\nabla f(A) = \overrightarrow{0}$. A fortiori, le plan tangent de f en A est parallèle au plan abscisse et les dérivées directionnelles de f en A sont toutes nulles.

Attention ! **RÉCIPROQUE FAUSSE !** La nullité du gradient ou de toutes les dérivées directionnelles **N'**est **PAS** le signe d'un extremum local. Vous n'en serez pas surpris, la réciproque était déjà fausse pour les fonctions d'une seule variable. Par exemple, en notant f la fonction $(x, y) \longmapsto x^3$: $\nabla f(0,0) = \overrightarrow{0}$, donc pour tout $\overrightarrow{v} = (h,k) \in \mathbb{R}^2$: $\mathrm{D}_{(0,0)} f\left(\overrightarrow{v}\right) = \left\langle \nabla f(0,0)\, ,\, \overrightarrow{v} \right\rangle = 0$, donc les dérivées directionnelles de f en $(0,0)$ sont toutes nulles, mais pourtant f n'a pas d'extremum local en $(0,0)$.



Démonstration Faisons l'hypothèse que f a un maximum local en A. Pour tout $\overrightarrow{v} \in \mathbb{R}^2$, la fonction d'une variable $t \overset{F}{\longmapsto} f\left(A + t\overrightarrow{v}\right)$ est de classe $\mathscr{C}^1$ et majorée par $F(0) = f(A)$ au voisinage de 0, donc possède un maximum local en 0. Comme voulu : $\mathrm{D}_{\overrightarrow{v}} f(A) = F'(0) = 0$.

Mais comment détermine-t-on concrètement les extrema locaux d'une fonction de deux variables ? Nous nous contenterons d'exploiter nos connaissances sur les polynômes du second degré et de raisonner par analyse-synthèse.

Pour $a, b, c \in \mathbb{R}$ fixés avec $(a, c) \neq (0,0)$, le signe de la fonction $(h, k) \overset{\varphi}{\longmapsto} ah^2 + bhk + ck^2$ est facile à déterminer. En notant P le polynôme $aX^2 + bX + c$: $\varphi(h,k) = k^2 P\left(\dfrac{h}{k}\right)$ pour tous $h \in \mathbb{R}$ et $k \in \mathbb{R}^*$.

— **Cas où $b^2 - 4ac \leqslant 0$:** Les réels a et c sont de même signe et P est de signe constant, celui de a ou c selon que a ou c est non nul, donc φ est elle aussi de signe constant.

— **Cas où $b^2 - 4ac > 0$:** Le polynôme P prend à la fois des valeurs strictement positives et des valeurs strictement négatives, donc la fonction φ aussi. Le signe de $\varphi(h,k)$ dépend du rapport $\dfrac{h}{k}$, i.e. de l'angle que le vecteur (h,k) forme avec le vecteur $\overrightarrow{\imath}$.

Exemple La fonction $(x,y) \overset{f}{\longmapsto} x^2 - 3x + xy + y^2$ possède un et un seul extremum local, en l'occurrence un minimum local en $(2, -1)$.

Démonstration Tout d'abord, f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$.

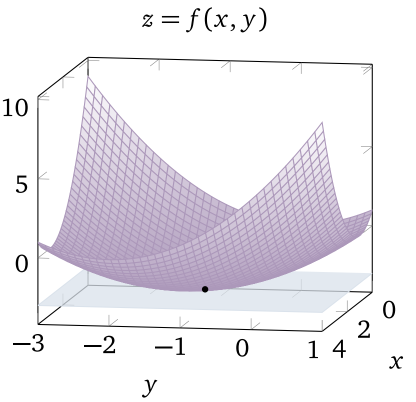
- **Analyse :** Soit $(x,y) \in \mathbb{R}^2$. Si f possède un extremum local en (x,y) : $\dfrac{\partial f}{\partial x}(x,y) = \dfrac{\partial f}{\partial y}(x,y) = 0$, donc $2x + y = 3$ et $x + 2y = 0$, donc $x = 2$ et $y = -1$.
- **Synthèse :** Le point $(2,-1)$ est-il bel et bien un extremum local de f ? Pour trancher, nous devons comparer $f(2+h, -1+k)$ et $f(2,-1)$ pour (h,k) choisi dans un voisinage de $(0,0)$. — Il vaut mieux se ramener au point $(0,0)$, on y voit plus clair. — Or pour tout $(h,k) \in \mathbb{R}^2$:
$$f(2+h,-1+k) - f(2,-1) = (2+h)^2 - 3(2+h) + (2+h)(-1+k) + (-1+k)^2 + 3 = h^2 + hk + k^2 = \left(h + \frac{k}{2}\right)^2 + \frac{3k^2}{4} \geqslant 0,$$
donc f possède un minimum local en $(2,-1)$. J'ai privilégié ici une mise sous forme canonique, bien pratique, mais on aurait pu s'intéresser au discriminant du polynôme $X^2 + X + 1$.

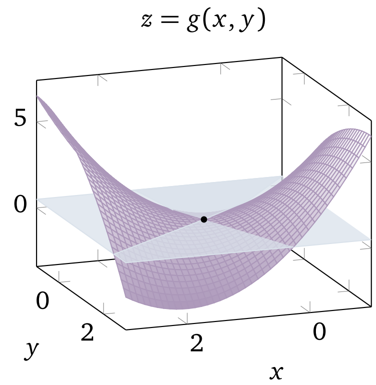
Exemple La fonction $(x,y) \overset{g}{\longmapsto} x^2 - x - xy - \dfrac{y^2}{2} + 2y$ ne possède aucun extremum local sur $\mathbb{R}^2$.

Démonstration Tout d'abord, g est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$.

- **Analyse :** Soit $(x,y) \in \mathbb{R}^2$. Si g possède un extremum local en (x,y) : $\dfrac{\partial g}{\partial x}(x,y) = \dfrac{\partial g}{\partial y}(x,y) = 0$, donc $2x - y = 1$ et $x + y = 2$, donc $x = 1$ et $y = 1$.

- **Synthèse :** Le point $(1,1)$ est-il bel et bien un extremum local de g ? Pour tout $(h,k)\in\mathbb{R}^2$:
$$g(1+h,1+k)-g(1,1)=(1+h)^2-(1+h)-(1+h)(1+k)-\frac{(1+k)^2}{2}+2(1+k)-\frac{1}{2}=h^2-hk-\frac{k^2}{2}.$$
Le discriminant du polynôme $X^2-X-\frac{1}{2}$ étant strictement positif, cette expression n'a pas un signe constant au voisinage de $(0,0)$, mais soyons précis. Pour tout $h\in\mathbb{R}^*$: $\quad g(1+h,1)-g(1,1)=h^2>0, \quad$ donc g n'est majorée par $g(1,1)$ sur aucun voisinage de $(1,1)$, et pour tout $k\in\mathbb{R}^*$: $\quad g(1,1+k)-g(1,1)=-\frac{k^2}{2}<0$, donc f n'est minorée par $g(1,1)$ sur aucun voisinage de $(1,1)$. Ainsi, g n'a pas d'extremum local en $(1,1)$.





5 Fonction différentiable et différentielle

Vous avez rencontré plusieurs *infiniment petits* en physique, notés $\mathrm{d}x$, $\mathrm{d}f$, $\mathrm{d}T$, $\mathrm{d}t$... mais la nature exacte de ces objets vous a sans doute paru mystérieuse. Comment le mathématicien les définit-il formellement ? La notion de *différentielle* est hors programme en MPSI, vous la découvrirez l'année prochaine et je me contenterai de lever un coin du voile.

La *différentielle* est aux fonctions de deux variables ce que la dérivée est aux fonctions d'une seule variable. Et de même qu'on dispose d'un concept de dérivabilité pour les fonctions d'une seule variable, on définit d'un concept de *différentiabilité* pour les fonctions de deux variables. Je n'irai cependant pas plus loin que la définition. Mais rappelons d'abord qu'une fonction d'une seule variable est dérivable en un point si et seulement si elle y possède un développement limité à l'ordre 1.

Définition (Fonction différentiable et différentielle) Soient Ω un ouvert de $\mathbb{R}^2$ et $f:\Omega\longrightarrow\mathbb{R}$ une fonction.

- **En un point :** Soit $A\in\Omega$. On dit que f est *différentiable en* A s'il existe une forme linéaire $\varphi\in\mathscr{L}\big(\mathbb{R}^2,\mathbb{R}\big)$ pour laquelle : $\quad f\big(A+\vec{v}\big)\underset{\vec{v}\to\vec{0}}{=}f(A)+\varphi\big(\vec{v}\big)+\mathrm{o}\big(\|\vec{v}\|\big), \quad$ ou encore : $\quad f(M)\underset{M\to A}{=}f(A)+\varphi\big(\overrightarrow{AM}\big)+\mathrm{o}\Big(\big\|\overrightarrow{AM}\big\|\Big)$.
Dans ces conditions, l'application φ est unique, appelée la *différentielle de f en A* et notée $\mathrm{d}f_A$. Ainsi :
$$f\big(A+\vec{v}\big)\underset{\vec{v}\to\vec{0}}{=}f(A)+\mathrm{d}f_A\big(\vec{v}\big)+\mathrm{o}\big(\|\vec{v}\|\big).$$
- **Sur un ouvert :** On dit que f est *différentiable sur* Ω si elle l'est en tout point de Ω.
L'application $\left\{\begin{array}{ccc}\Omega & \longrightarrow & \mathscr{L}\big(\mathbb{R}^2,\mathbb{R}\big)\\ A & \longmapsto & \mathrm{d}f_A\end{array}\right.$ est appelée dans ce cas la *différentielle de f sur* Ω et notée $\mathrm{d}f$.

Le réel $\mathrm{d}f_A\big(\vec{v}\big)$ est une approximation à l'ordre 1 de la variation de f au point A en ordonnée induite par une variation de vecteur $\vec{v}$ en abscisse. En d'autres termes, quand on passe de A à $A+\vec{v}$, f varie approximativement de $\mathrm{d}f_A\big(\vec{v}\big)$.

La différentielle $\mathrm{d}f$ elle-même est un objet relativement abstrait quand on le découvre. L'application $\mathrm{d}f$ associe à tout point de Ω une forme linéaire de $\mathbb{R}^2$.

Exemple Les fonctions $(x,y)\overset{f}{\longmapsto}x$ et $(x,y)\longmapsto y$ sont différentiables sur $\mathbb{R}^2$ et coïncident chacune avec sa différentielle. On les note respectivement $\mathrm{d}x$ et $\mathrm{d}y$.

Ce résultat n'est guère étonnant. Quand on passe de $A=(x_A,y_A)$ à $A+\vec{v}=(x_A+h,y_A+k)$, la fonction f varie de $h=f\big(\vec{v}\big)$.

Démonstration La fonction f a une propriété importante, elle est **linéaire** ! Soient $A=(x_A,y_A)\in\mathbb{R}^2$. Pour tout $\vec{v}=(h,k)\in\mathbb{R}^2$: $\quad f\big(A+\vec{v}\big)=x_A+h=f(A)+f\big(\vec{v}\big)\underset{\vec{v}\to\vec{0}}{=}f(A)+f\big(\vec{v}\big)+\mathrm{o}\big(\|\vec{v}\|\big), \quad$ donc comme annoncé, f est différentiable en A et $\mathrm{d}f_A=f$.

Définition **(Différentielle d'une fonction de classe $\mathscr{C}^1$)** Soient Ω un ouvert de $\mathbb{R}^2$ et $f \in \mathscr{C}^1(\Omega,\mathbb{R})$. Alors f est différentiable sur Ω et pour tous $A\in\Omega$ et $\vec{v}\in\mathbb{R}^2$: $\quad \mathrm{d}f_A\big(\vec{v}\big) = \dfrac{\partial f}{\partial x}(A)\,\mathrm{d}x_A\big(\vec{v}\big) + \dfrac{\partial f}{\partial y}(A)\,\mathrm{d}y_A\big(\vec{v}\big).$

En résumé : $\quad \mathrm{d}f = \dfrac{\partial f}{\partial x}\,\mathrm{d}x + \dfrac{\partial f}{\partial y}\,\mathrm{d}y.$

Démonstration De classe $\mathscr{C}^1$, f possède un développement limité à l'ordre 1 en tout point $A\in\Omega$:

$$f\big(A+\vec{v}\big) \underset{\vec{v}\to\vec{0}}{=} f(A) + \langle \nabla f(A)\,,\vec{v}\rangle + \mathrm{o}\big(\|\vec{v}\|\big).$$

La différentiabilité de f en A découle ainsi de la linéarité du produit scalaire $\vec{v}\longmapsto\langle\nabla f(A)\,,\vec{v}\rangle$. Enfin, pour tout $\vec{v}=(h,k)\in\mathbb{R}^2$: $\quad \mathrm{d}f_A\big(\vec{v}\big) = \langle\nabla f(A)\,,\vec{v}\rangle = \dfrac{\partial f}{\partial x}(A)\,h + \dfrac{\partial f}{\partial y}(A)\,k = \dfrac{\partial f}{\partial x}(A)\,\mathrm{d}x_A\big(\vec{v}\big) + \dfrac{\partial f}{\partial y}(A)\,\mathrm{d}y_A\big(\vec{v}\big).$

Et les physiciens dans tout ça ? Ils confondent souvent les fonctions qu'ils manipulent et leurs valeurs, et quand ils écrivent $\mathrm{d}f$, c'est de $\mathrm{d}f_A\big(\vec{v}\big)$ qu'ils parlent. Quand on passe du point A au point $A+\vec{v}$, f varie à l'ordre 1 de $\mathrm{d}f_A\big(\vec{v}\big)$, mais les physiciens auront tendance à considérer que f a varié de $\mathrm{d}f$.