Chapitre 17

Fonctions de plusieurs variables sur $\mathbb{R}^n$.

Compétences attendues.

✓ Montrer qu'une fonction f est continue ou $\mathscr{C}^1$.

✓ Calculer les dérivées partielles ou le gradient d'une fonction de classe $\mathscr{C}^1$.

✓ Déterminer les points critiques d'une fonction f.

✓ Déterminer la nature (extremum ou non) des points critiques de f.

Mathieu Mansuy

Professeur en ECG deuxième année spécialité mathématiques approfondies au Lycée Louis Pergaud (Besançon)

Page personnelle : `mathieu-mansuy.fr/`

E-mail : `mathieu.mansuy@ac-besancon.fr`

1 Introduction aux fonctions définies sur $\mathbb{R}^n$

1.1 Définition, exemples

Définition.

On appelle *fonction numérique à n variables* toute fonction f définie sur un sous-ensemble $\mathscr{D}$ de $\mathbb{R}^n$ à valeur dans $\mathbb{R}$:

$$f : (x_1, \dots, x_n) \in \mathscr{D} \mapsto f(x_1, \dots, x_n) \in \mathbb{R}.$$

Exemples.

- $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = x^3 - 4xy^2$.
- $g : \mathbb{R}^2 \to \mathbb{R}$ définie par $g(x, y) = \ln(1 + x^2 + y^2) + 2e^x$.
- $h : \mathscr{D} \to \mathbb{R}$ définie par $h(x, y) = \dfrac{xy}{x - y}$ sur le domaine $\mathscr{D} = \{(x, y) \in \mathbb{R}^2,\ x \neq y\}$.
- $i : \mathbb{R}^n \to \mathbb{R}$ définie par : $i(x_1, \dots, x_n) = \left(\displaystyle\sum_{i=1}^{n} x_i\right)^2 - \displaystyle\sum_{i=1}^{n} x_i^2$.

Définition.

- On appelle *fonction polynomiale* toute fonction de $\mathbb{R}^n$ dans $\mathbb{R}$, combinaison linéaire de fonctions de la forme $(x_1, \dots, x_n) \mapsto x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$, où $k_1, k_2, \dots, k_n$ sont des entiers naturels.
- On appelle *fonction affine* toute fonction polynomiale de $\mathbb{R}^n$ dans $\mathbb{R}$ de la forme :

  $$f : (x_1, \dots, x_n) \in \mathbb{R}^n \mapsto u_1 x_1 + \cdots + u_n x_n + c$$

  où $u_1, \dots, u_n, c \in \mathbb{R}$.

Exemple. f et i sont des fonctions polynomiales.

1.2 Graphe

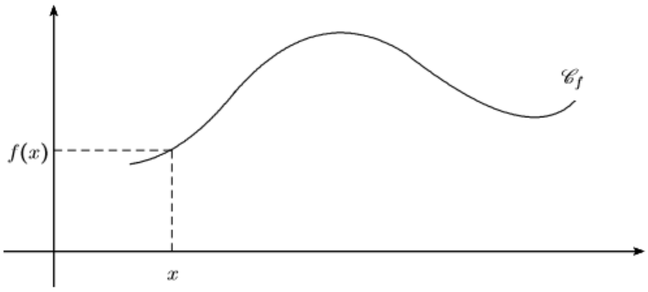
Définition.

Soit $f : \mathbb{R}^n \to \mathbb{R}$. On appelle *graphe de f* le sous-ensemble de $\mathbb{R}^{n+1}$ défini par :

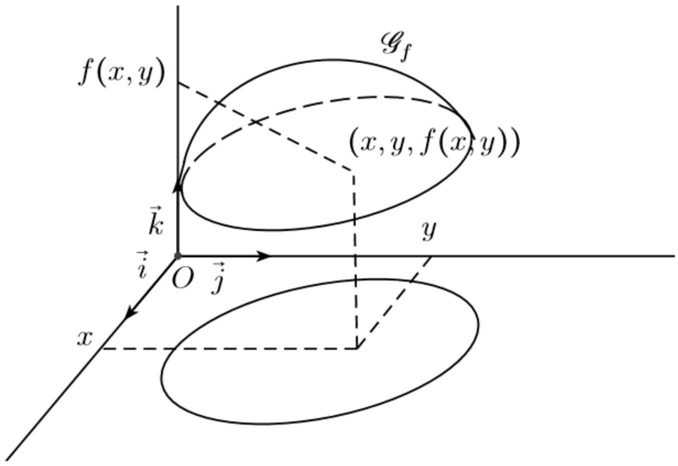
$$\mathscr{G}_f = \{(x_1, \dots, x_n, x_{n+1}) \in \mathbb{R}^{n+1},\ x_{n+1} = f(x_1, \dots, x_n)\}.$$

Remarques.

- Cas $n = 1$. Le graphe d'une fonction d'une variable réelle f est alors sa courbe représentative, qu'on représentera dans le plan $\mathbb{R}^2$.



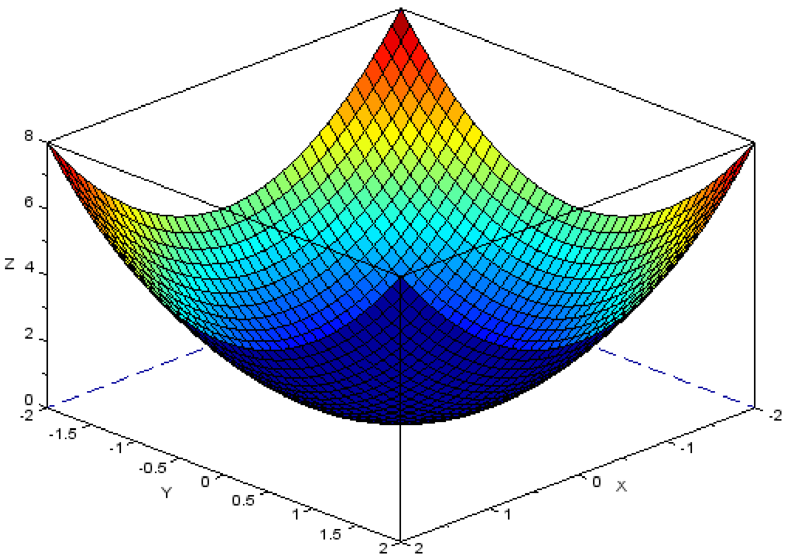
- Cas $n = 2$. Le graphe de f est une surface qu'on représentera dans l'espace $\mathbb{R}^3$.



- Cas $n \geq 3$. Le graphe de f est une *hypersurface* de $\mathbb{R}^{n+1}$, et il n'est alors pas possible de le représenter.

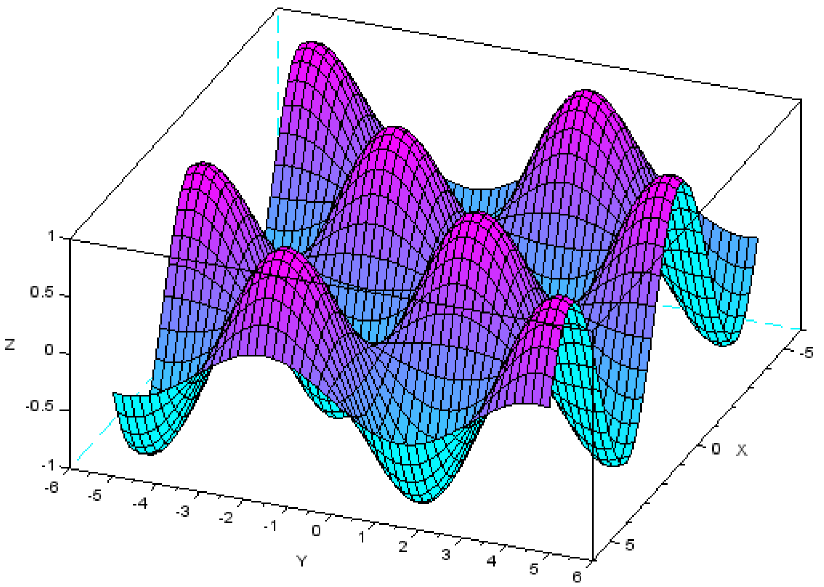
Exemples.

- Soit $f : (x, y) \in \mathbb{R}^2 \mapsto x^2 + y^2$. Le graphe de f est :



Paraboloïde elliptique.

- Soit $g : (x, y) \in \mathbb{R}^2 \mapsto \sin(x)\sin(y)$. Le graphe de g est :



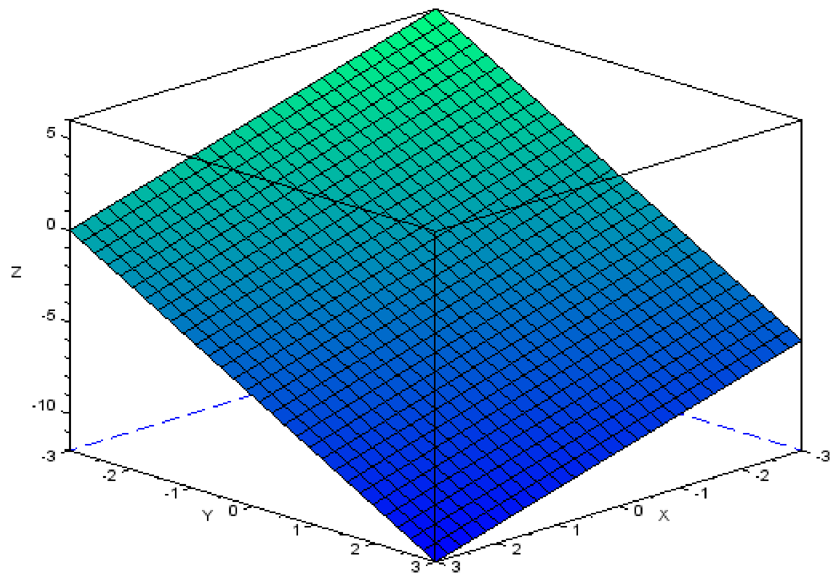
Remarque. Cas des fonctions affines.

- Supposons $n = 1$ et soit f une fonction affine, c'est-à-dire de la forme $f : x \mapsto ax + b$ avec $a, b \in \mathbb{R}$. Alors le graphe de f est la droite affine d'équation $y = ax + b$ (droite vectorielle si $b = 0$).

- Supposons $n = 2$ et considérons f une fonction affine de la forme $f : (x, y) \mapsto ax + by + c$ avec $a, b, c \in \mathbb{R}$. Alors on a :

$$(x, y, z) \in \mathscr{G}_f \quad \Leftrightarrow \quad z = f(x, y) \quad \Leftrightarrow \quad ax + by - z + c = 0.$$

  Le graphe de f est alors un plan de l'espace dit *affine* (ou *vectoriel* si $c = 0$).



Exemple d'un plan affine.

Plus généralement, si $f : \mathbb{R}^n \to \mathbb{R}$ est une fonction affine, son graphe est appelé *hyperplan affine* (*hyperplan vectoriel* s'il passe par $(0, \ldots, 0)$, soit si son terme constant est nul).

1.3 Lignes de niveau

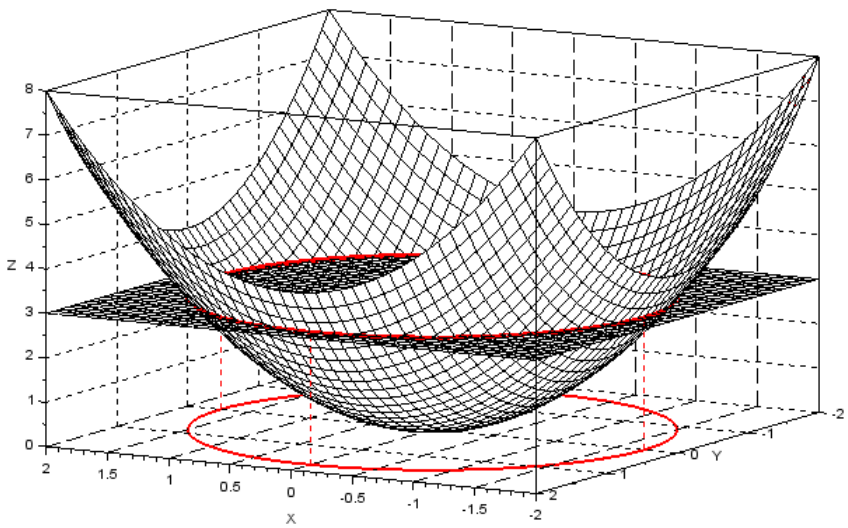
Définition.

Soit $f : \mathbb{R}^n \to \mathbb{R}$, et soit $\lambda \in \mathbb{R}$. On appelle *ligne de niveau* λ de f l'ensemble des points $(x_1, \ldots, x_n) \in \mathbb{R}^n$ qui vérifient l'équation :

$$f(x_1, \ldots, x_n) = \lambda.$$

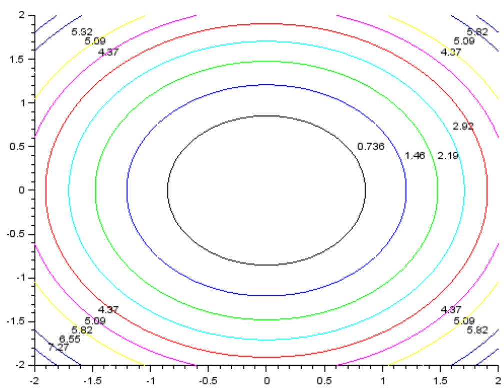
Remarques.

- La ligne de niveau λ de f s'obtient en faisant l'intersection du graphe $\mathscr{G}_f$ de f avec le plan « horizontal » d'équation $x_{n+1} = \lambda$.

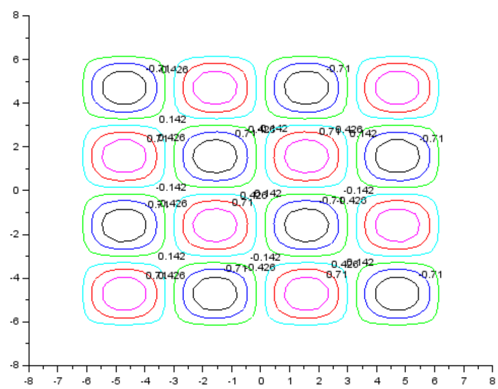


Ligne de niveau $\lambda = 3$ de $f : (x, y) \mapsto x^2 + y^2$.

- L'indication d'un nombre suffisant de lignes de niveau permet d'avoir une représentation assez fidèle du graphe de la fonction.



Lignes de niveau de $f : (x, y) \mapsto x^2 + y^2$.



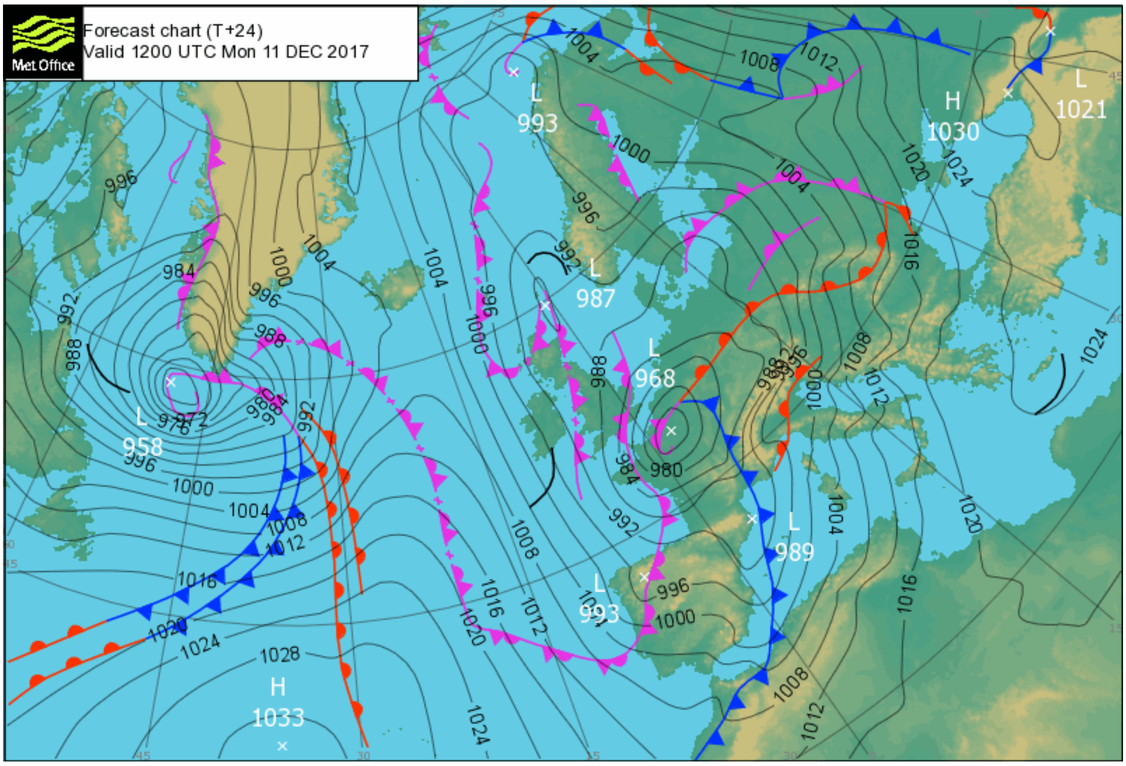
Lignes de niveau de $g : (x, y) \mapsto \sin(x)\sin(y)$.

Exemple. Si f est la fonction qui à la longitude et la latitude associe l'altitude, alors les lignes de niveau représentent les points qui sont à la même altitude : si on se promène sur une ligne de niveau, on ne monte ni ne descend. Ce sont ces lignes qui sont représentées sur les cartes topographiques.

Carte topographique de la citadelle.

Vue satellite 3D de la citadelle.

Exemple. Si f est la fonction qui à la longitude et à la latitude associe la pression (au niveau de la mer), alors les lignes de niveau relient des points d'égale pression. Ces lignes sont appelées en météorologie des *courbes isobares.*



Carte de surface.

Un ensemble d'isobares incurvées entourant une zone de basse (resp. haute) pression indique une dépression (resp. anticyclone), repéré par un L (resp. H) sur la carte de surface. La vitesse du vent est fonction de l'écartement des isobares : plus les isobares sont serrées, plus la pression varie rapidement, plus le vent souffle fort[1].

Exemple. Considérons une fonction $f : \mathbb{R}^3 \to \mathbb{R}$ qui en tout point de l'espace donne la température en ce point. Alors la ligne de niveau 0 est appelée en météorologie l'*isotherme* 0. C'est l'ensemble des points de l'espace en lesquels la température est égale à $0°C$. C'est une surface de l'espace.

[1] Pour apprendre à lire une carte de pression atmosphérique, on pourra consulter ce lien.

2 Continuité d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$

2.1 Rappels sur la norme euclidienne

Rappels.

- Soit $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

  On appelle *norme (euclidienne) de x* et on note $\|x\|$ le réel $\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$.

- Soit $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$ deux éléments de $\mathbb{R}^n$.

  On appelle *distance de x à y* le réel $d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$.

Propriété 1

La norme $\|\cdot\|$ sur $\mathbb{R}^n$ vérifie les trois propriétés suivantes :

- $\forall x \in \mathbb{R}^n, \quad \|x\| = 0 \Leftrightarrow x = 0_{\mathbb{R}^n}$.
- $\forall x \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}, \quad \|\lambda x\| = |\lambda| \cdot \|x\|$.
- *Inégalité triangulaire :* $\quad \forall x, y \in \mathbb{R}^n, \quad \|x + y\| \leq \|x\| + \|y\|$.

2.2 Définition de la continuité

Définition.

Soit $f : \mathbb{R}^n \to \mathbb{R}$.

- Soit $x_0 \in \mathbb{R}^n$. On dit que f est *continue en x_0* si :

$$\forall \varepsilon > 0, \quad \exists \eta > 0, \quad \forall x \in \mathbb{R}^n, \quad \|x - x_0\| \leq \eta \quad \Rightarrow \quad |f(x) - f(x_0)| \leq \varepsilon.$$

- On dit que f est *continue sur $\mathbb{R}^n$* lorsque f est continue en tout point de $\mathbb{R}^n$.

Remarques.

- Si $n = 1$, alors la norme euclidienne n'est autre que la valeur absolue (puisque $\|x\| = \sqrt{x^2} = |x|$), et la définition de la continuité est celle d'une fonction d'une variable aléatoire réelle donnée en première année.
- La définition de la continuité de f en x_0 est la traduction mathématique de la propriété suivante : la distance entre $f(x)$ et $f(x_0)$ (qui est $|f(x) - f(x_0)|$) est aussi petite qu'on veut dès que la distance entre x et x_0 (qui est $\|x - x_0\|$) est suffisamment petite.

Exercice. Montrer la continuité des applications coordonnées $p_i : \mathbb{R}^n \to \mathbb{R}$ (où $i \in [\![1, n]\!]$) définies par :

$$p_i : x = (x_1, \dots, x_n) \in \mathbb{R}^n \mapsto x_i.$$

2.3 Opérations sur les fonctions continues

Propriété 2

On suppose que f et g sont deux fonctions continues sur $\mathbb{R}^n$.

- Les fonctions λf (pour λ réel), $f+g$ et $f \times g$ sont continues sur $\mathbb{R}^n$.
- Si de plus f ne s'annule pas sur $\mathbb{R}^n$, alors les fonctions $\dfrac{1}{f}$ et $\dfrac{g}{f}$ sont continues sur $\mathbb{R}^n$.

Corollaire 3

Les fonctions polynomiales sont continues sur $\mathbb{R}^n$.

Preuve.

□

Propriété 4

Soient $I \subset \mathbb{R}$ un intervalle et $\varphi : I \to \mathbb{R}$ une fonction. On suppose que :

- f est continue sur $\mathbb{R}^n$,
- f est à valeurs dans I : $\forall (x_1, \dots, x_n) \in \mathbb{R}^n,\ f(x_1, \dots, x_n) \in I$,
- φ est continue sur I.

Alors $\varphi \circ f$ est continue sur $\mathbb{R}^n$.

Exercice. Montrer la continuité des fonctions suivantes :

- $f : (x, y) \in \mathbb{R}^2 \mapsto x \ln(1 + x^2 + y^2)$.
- $f : (x, y) \in \mathbb{R}^2 \mapsto \max(x, y)$.
- $f : x \in \mathbb{R}^n \mapsto \|x\|$.

Remarque. Ces énoncés sont également valables ponctuellement, c'est-à-dire si l'on remplace « sur $\mathbb{R}^n$ » par « en x_0 » avec $x_0 \in \mathbb{R}^n$.

3 Calcul différentiel

3.1 Dérivées partielles, gradient

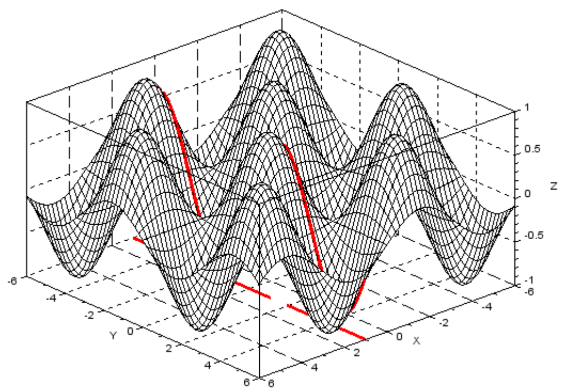
Définition.

Soient $i \in [\![1,n]\!]$ et $a = (a_1, \dots, a_n) \in \mathbb{R}^n$. On appelle *$i$-ème fonction partielle de f en a* la fonction d'une variable réelle :

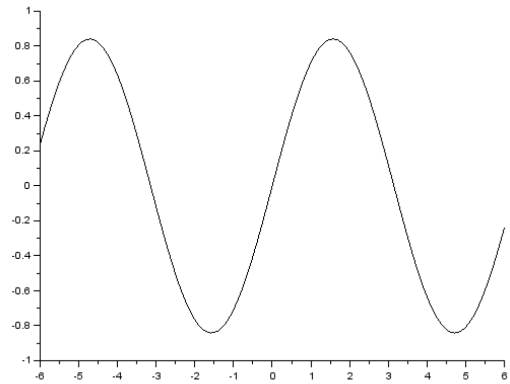
$$f_{a,i} \ : \ \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_n). \end{array} \right.$$

Remarque. Si f est continue, alors toutes ses fonctions partielles sont continues. Mais la réciproque est fausse : il existe des fonctions non continues dont les fonctions partielles sont toutes continues.

Remarque. Dans le cas où $n = 2$ et $a = (a_1, a_2)$, la courbe de la deuxième fonction partielle $f_{a,2} : y \to f(a_1, y)$ s'obtient en faisant l'intersection du graphe $\mathscr{G}_f$ de f avec le plan d'équation $x = a_1$. Représentons $f_{a,2}$ dans le cas où $f : (x,y) \in \mathbb{R}^2 \mapsto \sin(x)\sin(y)$ et $a = (1, -\pi)$.



Intersection de $\mathscr{G}_f$ et du plan d'équation $x = 1$.



Graphe de $f_{a,2} : y \mapsto \sin(1)\sin(y)$.

Définition.

Soient $i \in [\![1,n]\!]$ et $a = (a_1, \dots, a_n) \in \mathbb{R}^n$.

On dit que f *admet une dérivée partielle d'ordre 1 par rapport à la i-ème variable en a* lorsque la fonction $f_{a,i}$ est dérivable en a_i, soit si :

$$\lim_{h \to 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_n)}{h} \quad \text{existe.}$$

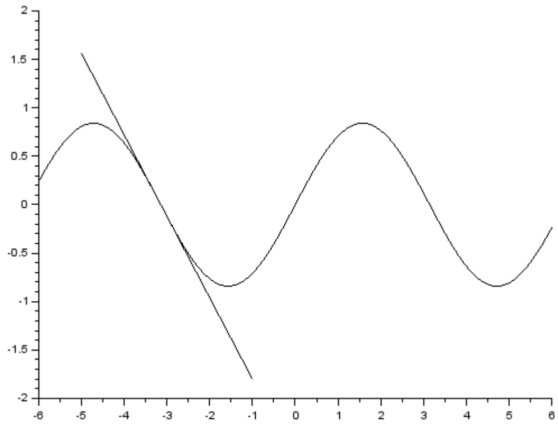
On appelle alors *dérivée partielle d'ordre 1 de f par rapport à la i-ème variable en a* et on note $\partial_i(f)(a)$ le réel :

$$\partial_i(f)(a) = f'_{a,i}(a_i).$$

Notation.

La notation $\dfrac{\partial}{\partial x_i}$ est aussi courante pour la dérivée partielle d'ordre 1 par rapport à la i-ème variable. Le programme officiel privilégie cependant la notation ∂_i.

Exemple. Reprenons $f : (x,y) \in \mathbb{R}^2 \mapsto \sin(x)\sin(y)$ et $a = (1, -\pi)$. Alors $\partial_2 f(1, -\pi)$ est le coefficient directeur de la tangente à la courbe de $f_{a,2}$ en $-\pi$.



Tangente en $-\pi$ à la courbe de $f_{a,2}$.

Pour tout $y \in \mathbb{R}$,

$$\partial_2 f(1, y) = \sin(1)\cos(y)$$

et donc :

$\partial_2 f(1, -\pi) = \sin(1)\cos(-\pi) = -\sin(1) \approx -0,84.$

Le coefficient directeur de la tangente à la courbe de $f_{a,2}$ en $-\pi$ est de $-0,84$.

Exercice. Déterminer les dérivées partielles des fonctions suivantes :

- $f : (x, y, z) \mapsto xyz + 2x^3 + yz.$

- $g : (x, y) \in \mathbb{R}^2 \mapsto xe^{x^2+y^2}.$

Définition.

Soit $a \in \mathbb{R}^n$ et $f : \mathbb{R}^n \to \mathbb{R}$ une fonction. On suppose que f admet une dérivée partielle d'ordre 1 en a par rapport à chaque variable.

On appelle *gradient de f en a*, et on note $\nabla(f)(a)$ (∇ se lit « nabla »), le vecteur de $\mathbb{R}^n$ défini par :

$$\nabla(f)(a) = (\partial_1(f)(a), \ldots, \partial_n(f)(a)).$$

Exercice. Déterminer le gradient de $f : (x, y) \in \mathbb{R}^2 \mapsto x^2(1 + \ln(1 + y^2))$ en tout point de $\mathbb{R}^2$.

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

Définition.

On dit que f est *de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$* lorsque pour tout $i \in [\![1, n]\!]$, f admet une i-ème dérivée partielle en tout point de $\mathbb{R}^n$, et si $\partial_i(f)$ est continue sur $\mathbb{R}^n$.

Propriété 5

On suppose que f et g sont deux fonctions de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.

- Les fonctions λf (pour λ réel), $f+g$ et $f \times g$ sont de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.
- Si de plus f ne s'annule pas sur $\mathbb{R}^n$, alors les fonctions $\dfrac{1}{f}$ et $\dfrac{g}{f}$ sont de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.

Propriété 6

Soit $I \subset \mathbb{R}$ un intervalle. Soit $\varphi : I \to \mathbb{R}$. On suppose que :

- f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$,
- f est à valeurs dans I : $\quad \forall (x_1, \ldots, x_n) \in \mathbb{R}^n,\ f(x_1, \ldots, x_n) \in I$,
- φ est de classe $\mathscr{C}^1$ sur I.

Alors $\varphi \circ f$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.

Propriété 7

Toute fonction polynômiale est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.

Preuve.

$\square$

Exercice. Déterminer si les applications suivantes sont $\mathscr{C}^1$ sur $\mathbb{R}^n$ ou non.

- $f : (x,y) \in \mathbb{R}^2 \mapsto xye^{-x+y}$.
- $g : x \in \mathbb{R}^n \mapsto \|x\| = \sqrt{\displaystyle\sum_{i=1}^n x_i^2}$.

3.3 Développement limité d'ordre 1

Rappel. Soit f une fonction de classe $\mathscr{C}^1$ sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$. D'après la formule de Taylor-Young à l'ordre 1, elle admet un développement limité d'ordre 1 en tout réel a, qui s'écrit :

$$f(a+h) = f(a) + f'(a)h + h \cdot \varepsilon(h) \quad \text{avec} \quad \lim_{h\to 0} \varepsilon(h) = 0$$

ou encore en posant $x = a + h$:

$$f(x) = f(a) + f'(a)(x-a) + (x-a) \cdot \varepsilon(x-a).$$

Localement (autour de a), on peut approcher la fonction f par la fonction affine $x \mapsto f(a)+f'(a)(x-a)$, dont la courbe n'est autre que la tangente T_a à la courbe représentative de f en a.

On généralise ici ces résultats aux fonctions définies sur $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$.

Théorème 8 (*Formule de Taylor* (1685 - 1731) *d'ordre 1*)

Soient f une fonction de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$, et a un élément de $\mathbb{R}^n$.
Alors il existe une fonction ε continue en $(0,\ldots,0)$, vérifiant $\varepsilon(0,\ldots,0) = 0$ et telle que pour tout $h = (h_1,\ldots,h_n) \in \mathbb{R}^n$:

$$\begin{aligned} f(a+h) &= \underbrace{f(a) + \langle \nabla(f)(a), h\rangle}_{\text{partie principale du DL}} + \underbrace{\|h\| \cdot \varepsilon(h)}_{\text{reste du DL}} \\ &= f(a) + \sum_{i=1}^{n} \partial_i(f)(a_1,\ldots,a_n)h_i + \sqrt{\sum_{i=1}^{n} h_i^2} \cdot \varepsilon(h_1,\ldots,h_n) \end{aligned}$$

Une telle écriture est unique, appelée le *développement limité de f d'ordre 1 au point a* (ou la *formule de Taylor à l'ordre 1 de f en a*).

Remarque. Si on effectue le changement de variables $x = a + h$, on obtient :

$$f(x) = f(a) + \langle \nabla(f)(a), x-a\rangle + \|x-a\| \cdot \varepsilon(x-a).$$

Notons $t_a : x \mapsto f(a) + \langle \nabla(f)(a), x-a\rangle$. Si on note $a = (a_1,\ldots,a_n)$ et $x = (x_1,\ldots,x_n)$, alors :

$$t_a(x) = f(a_1,\ldots,a_n) + \partial_1(f)(a_1,\ldots,a_n)(x_1-a_1) + \cdots + \partial_n(f)(a_1,\ldots,a_n)(x_n-a_n).$$

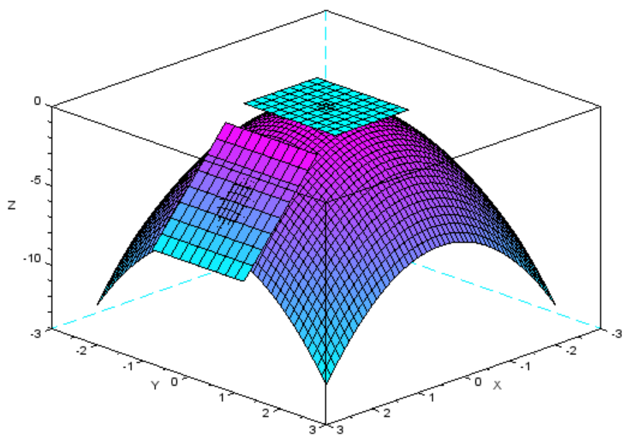
Donc t_a est une fonction affine. Et comme le reste $\|x-a\| \cdot \varepsilon(x-a)$ tend vers 0 lorsque x se rapproche de a, on peut approcher localement (autour du point a) la fonction f par t_a.

Définition.

On appelle *plan affine tangent au graphe de f en $a = (a_1,\ldots,a_n) \in \mathbb{R}^n$* le graphe de la fonction affine $x \mapsto t_a(x) = f(a) + \langle \nabla(f)(a), x-a\rangle$, c'est-à-dire l'hyperplan affine de $\mathbb{R}^{n+1}$ d'équation :

$$x_{n+1} = f(a) + \partial_1(f)(a)(x_1-a_1) + \cdots + \partial_n(f)(a)(x_n-a_n).$$

Exercice. Justifier que la fonction $f : (x,y) \in \mathbb{R}^2 \mapsto -x^2 - y^2$ est $\mathscr{C}^1$ sur $\mathbb{R}^2$, puis déterminer l'équation des plans tangents aux points $(a,b) = (0,0)$ et $(2,0)$.



Graphe de f et plans tangents en $(0,0)$ et $(2,0)$.

Corollaire 9

Si $f : \mathbb{R}^n \to \mathbb{R}$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$, alors elle est continue.

Preuve. Puisque f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$, elle admet un développement limité d'ordre 1 en tout point de $\mathbb{R}^n$. Pour tout $a \in \mathbb{R}^n$, il existe donc une fonction ε continue en $(0,\ldots,0)$, vérifiant $\varepsilon(0,\ldots,0) = 0$ et telle que pour tout $h = (h_1,\ldots,h_n) \in \mathbb{R}^n$:

$$f(a+h) = f(a) + \langle \nabla(f)(a), h\rangle + \|h\| \cdot \varepsilon(h)$$

Majorons :

$$\begin{aligned}
|f(a+h) - f(a)| &= |\langle \nabla(f)(a), h\rangle + \|h\| \cdot \varepsilon(h)| \\
&\leq |\langle \nabla(f)(a), h\rangle| + \|h\| \cdot |\varepsilon(h)| && \text{par l'inégalité triangulaire} \\
&\leq \|\nabla(f)(a)\|\|h\| + \|h\| \cdot |\varepsilon(h)| && \text{par Cauchy Schwartz} \\
&\leq \|h\| \cdot (\|\nabla(f)(a)\| + |\varepsilon(h)|)
\end{aligned}$$

Or la fonction $h \in \mathbb{R}^n \mapsto \|h\| \cdot (\|\nabla(f)(a)\| + |\varepsilon(h)|)$ est continue en $0_{\mathbb{R}^n}$ comme composée et produit de fonctions continues en $0_{\mathbb{R}^n}$, et vaut 0 en ce point. D'où :

$$\forall \varepsilon > 0, \quad \exists \nu > 0, \quad \forall h \in \mathbb{R}^n, \quad \|h\| \leq \nu \quad \Rightarrow \quad \|h\| \cdot (\|\nabla(f)(a)\| + |\varepsilon(h)|) \leq \varepsilon$$

Par conséquent :

$$|f(a+h) - f(a)| \leq \varepsilon.$$

Ainsi, f est continue en tout point $a \in \mathbb{R}^n$, et donc sur $\mathbb{R}^n$. $\square$

Mise en garde.

- L'application norme $x \in \mathbb{R}^n \mapsto \|x\|$ montre que la réciproque est fausse : cette fonction est continue sur $\mathbb{R}^n$, mais pas de classe $\mathscr{C}^1$.
- **Attention**, une fonction peut admettre des dérivées partielles par rapport à chaque variable en un point a, mais ne pas être continue en a.

3.4 Dérivées directionnelles

Propriété 10

Si f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$, et $a, u \in \mathbb{R}^n$, $u = (u_1, \dots, u_n) \neq (0, \dots, 0)$. Alors la fonction

$$g : t \mapsto f(a + tu)$$

est de classe $\mathscr{C}^1$ sur $\mathbb{R}$, et :

$$\forall t \in \mathbb{R}, \quad g'(t) = \langle \nabla(f)(a + tu), u \rangle = \partial_1 f(a + tu)u_1 + \cdots + \partial_n f(a + tu)u_n.$$

Preuve. Fixons $t \in \mathbb{R}$, et écrivons le taux d'accroissement de g en t. Pour tout $s \in \mathbb{R}$, $s \neq 0$:

$$\frac{g(t+s) - g(t)}{s} = \frac{f(a + (t+s)u) - f(a + tu)}{s} = \frac{f((a + tu) + su) - f(a + tu)}{s}.$$

On cherche la limite de ce taux d'accroissement lorsque $s \to 0$. Pour cela, on écrit le développement limité de f au point $a + tu$ (qui existe car f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$). Il existe une fonction ε continue en $(0, \dots, 0)$ vérifiant $\varepsilon(0, \dots, 0) = 0$ et telle que pour tout $h \in \mathbb{R}^n$:

$$f((a + tu) + h) = f(a + tu) + \langle \nabla(f)(a + tu), h \rangle + \|h\|\varepsilon(h)$$

En prenant $h = su$, on obtient :

$$f((a+tu)+su) = f(a+tu)+\langle \nabla(f)(a+tu), su\rangle+\|su\|\varepsilon(su) = f(a+tu)+s\langle \nabla(f)(a+tu), u\rangle+|s|\|u\|\varepsilon(su)$$

Ainsi :

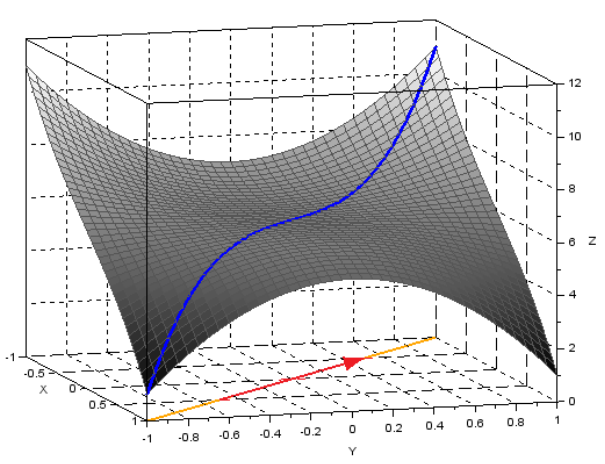
$$\frac{g(t+s) - g(t)}{s} = \frac{f(a + (t+s)u) - f(a + tu)}{s} = \langle \nabla(f)(a + tu), u \rangle \pm \|u\|\varepsilon(su).$$

Or $\lim\limits_{s \to 0} \|u\|\varepsilon(su)$ existe et vaut 0 car ε est continue en $(0, \dots, 0)$. On en déduit que g est bien dérivable en t et que :

$$g'(t) = \langle \nabla(f)(a + tu), u \rangle = \partial_1 f(a + tu)u_1 + \cdots + \partial_n f(a + tu)u_n.$$

Enfin f étant $\mathscr{C}^1$, toutes ses dérivées partielles sont continues. Par composition et somme, g' est donc continue et g est bien de classe $\mathscr{C}^1$. □

Remarque. Pour $n = 2$, le graphe de la fonction g est la section du graphe de f par le plan vertical contenant la droite $\mathscr{D}$ passant par a et dirigée par u. Par exemple, si $a = (0.5, -0.5)$, $u = (-1, 1)$, pour $f : (x, y) \in \mathbb{R} \mapsto x^3 - 4xy^2 + 6$, on obtient le graphe suivant :



En orange, la droite $\mathscr{D}$ passant par a et dirigée par le vecteur u en rouge, et en bleu la section du graphe de f par le plan vertical contenant $\mathscr{D}_{a,u}$.

Remarques.

- $g'(0) = \langle \nabla(f)(a), u \rangle$ indique la vitesse de variation de f si, partant du point a, on se déplace légèrement dans la direction u.
- Remarquons que pour $u = e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$:

$$g'(0) = \langle \nabla(f)(a), e_i \rangle = \partial_i f(a).$$

Interprétation géométrique du gradient.
Supposons que $\nabla(f)(a) \neq (0, \ldots, 0)$, et $\|u\| = 1$. D'après l'inégalité de Cauchy-Schwarz :

$$|g'(0)| = |\langle \nabla(f)(a), u \rangle| \leq \|\nabla(f)(a)\| \cdot \|u\| = \|\nabla(f)(a)\|$$

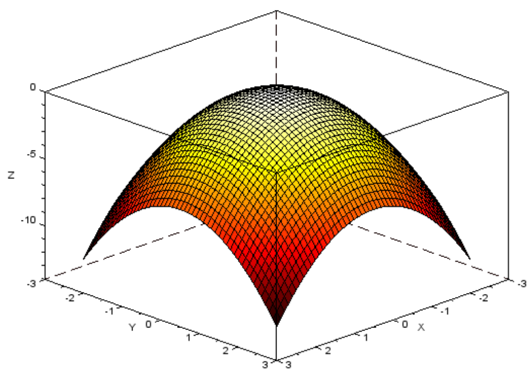
avec égalité lorsque les vecteurs $\nabla(f)(a)$ et u sont colinéaires.

Les variations de f au voisinage de a sont donc maximales en valeur absolue lorsqu'on se déplace dans la direction du vecteur $\nabla(f)(a)$. Ainsi :

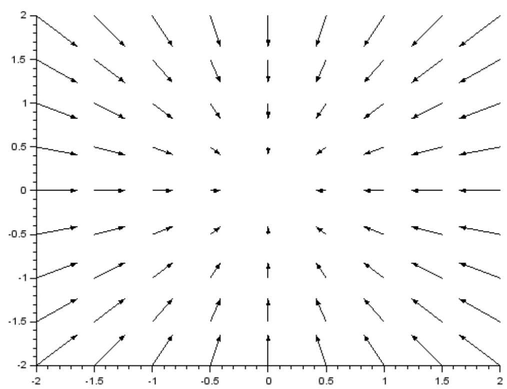
- $\nabla(f)(a)$ donne la direction de plus fort accroissement de f à partir du point a,
- $-\nabla(f)(a)$ donne la direction de plus forte diminution de f à partir du point a.

Physiquement, si on dépose une goutte d'eau en $(a, f(a))$ sur la surface représentant graphiquement f alors elle va glisser selon la direction et le sens $-\nabla(f)(a)$.

Exemple. Considérons toujours $f : (x, y) \in \mathbb{R}^2 \mapsto -x^2 - y^2$. Traçons le graphe de f ainsi que le champs de gradients de f (c'est-à-dire le gradient de f sur un réseau de $\mathbb{R}^2$).



Graphe de la fonction f.

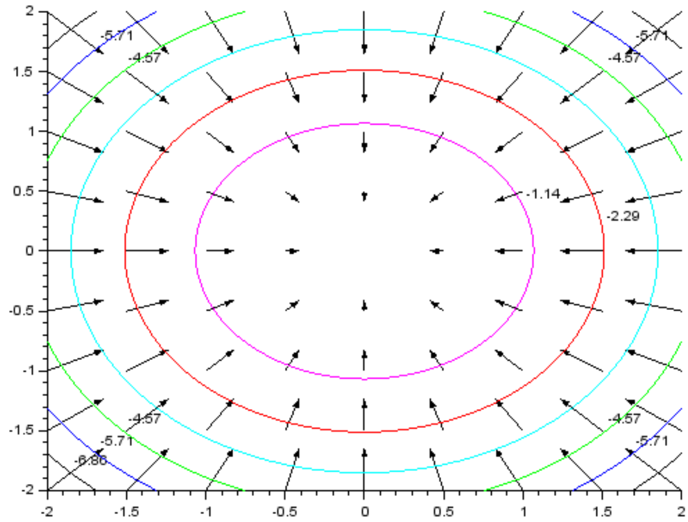


Champs de gradients associé à f.

On remarque bien sur cet exemple que le gradient indique la direction de plus grande pente.

Exemple. Si f est la fonction qui à la longitude et à la latitude associe la pression (au niveau de la mer), alors la norme du gradient de f correspond à la vitesse du vent. En pratique, un navigateur mesure la vitesse du vent en relevant l'écartement des isobares : plus les isobares sont serrées, plus le vent est important. La direction du vent, elle, n'est par contre pas donnée par le gradient du fait de l'action d'une autre force, la force de Coriolis.

Remarque. Gradient et lignes de niveau.
Reprenons la fonction $f : (x, y) \in \mathbb{R}^2 \mapsto -x^2 - y^2$, et représentons le *champs de gradient* de f ainsi que ses lignes de niveau sur un même graphe.



Champs de gradients et lignes de niveau de f.

On remarque que **les gradients sont orthogonaux aux** (tangentes aux) **lignes de niveau**. Cette propriété est en fait toujours vérifiée : une ligne de niveau est toujours orthogonale à la direction de la plus grande pente, et donc au gradient.

✍ À retenir. Interprétation géométrique du gradient.

- Le gradient $\nabla f(a)$ indique la direction « de plus grande pente », c'est-à-dire la direction dans laquelle f croit le plus fortement au voisinage de a.
- Les (tangentes aux) lignes de niveau sont orthogonales aux gradients.

4 Extremum globaux

4.1 Définition

Définition.

Soit $a \in \mathbb{R}^n$ et $f : \mathbb{R}^n \to \mathbb{R}$ une fonction.

- On dit que f admet un *minimum global en a* (resp. *maximum global en a*) lorsque :

$$\forall x \in \mathbb{R}^n, \quad f(x) \geq f(a) \quad (\text{resp. } f(x) \leq f(a)).$$

- On dit que f admet un *extremum global en a* lorsque f admet un minimum global ou un maximum global en a.

Remarque. f admet un minimum (resp. maximum) global en a lorsque $f(a)$ est la plus petite (resp. grande) valeur prise par f sur $\mathbb{R}^n$.

Exemples.

- La fonction $f : (x, y) \in \mathbb{R}^2 \mapsto -x^2 - y^2$ admet un maximum global en $a = (0, 0)$ puisque pour tout $(x, y) \in \mathbb{R}^2$,

$$f(x, y) - f(0, 0) = -x^2 - y^2 \leq 0.$$

Par contre, elle n'admet pas de minimum global. On peut le justifier en remarquant par exemple que :

$$f(x,0) = -x^2 \underset{x\to+\infty}{\longrightarrow} -\infty.$$

- La fonction $g : (x,y) \in \mathbb{R}^2 \mapsto \sin(x)\sin(y)$ a un maximum global égal à 1 et un minimum global égal à -1. De plus, ce maximum global est atteint si, et seulement si, $\sin(x)$ et $\sin(y)$ sont simultanément égaux à 1 ou à -1. Il est donc atteint une infinité de fois, en les points :

$$\left(\frac{\pi}{2} + k\pi, \frac{\pi}{2} + \ell\pi\right), \ (k,\ell) \in \mathbb{N}^2, \ k \text{ et } \ell \text{ de même parité.}$$

  De même, le minimum global est atteint une infinité de fois.

- La fonction $h : (x,y) \in \mathbb{R}^2 \mapsto x^3 - 4xy^2 + 6$ n'admet ni maximum global, ni minimum global, puisque :

$$h(x,0) = x^3 + 6 \underset{x\to\pm\infty}{\longrightarrow} \pm\infty.$$

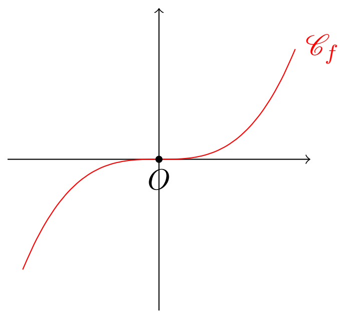
4.2 Condition nécessaire du premier ordre

Rappel. Soit $f :]a,b[\to \mathbb{R}$ une fonction de classe $\mathscr{C}^1$ sur $]a,b[$. Alors :

$$f \text{ admet un extremum en } c \in]a,b[\quad \Rightarrow \quad f'(c) = 0.$$

Mise en garde.

La réciproque est fausse : par exemple, la fonction $f : x \in \mathbb{R} \mapsto x^3$ satisfait $f'(0) = 0$, mais f n'admet pas d'extremum global en 0.



Courbe représentative de la fonction $x \mapsto x^3$.

Théorème 11 (Condition nécessaire d'extremum global)

On suppose f de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$.
Si f admet un extremum global en c, **alors** $\nabla(f)(c) = (0,\ldots,0)$.

Preuve.

□

Définition.

On dit que c est un *point critique* de f lorsque $\nabla(f)(c) = (0, \dots, 0)$.

Mise en garde.

Si f admet un extremum en c, **alors** c est un point critique de f.

Attention, la réciproque est **fausse** : si c est un point critique de f, f n'admet pas nécessairement un extremum global en c.

Exemple. La fonction $f : (x, y) \in \mathbb{R}^2 \mapsto x^2 - y^2$ est polynomiale, donc de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$. Pour tout $(x, y) \in \mathbb{R}^2$:

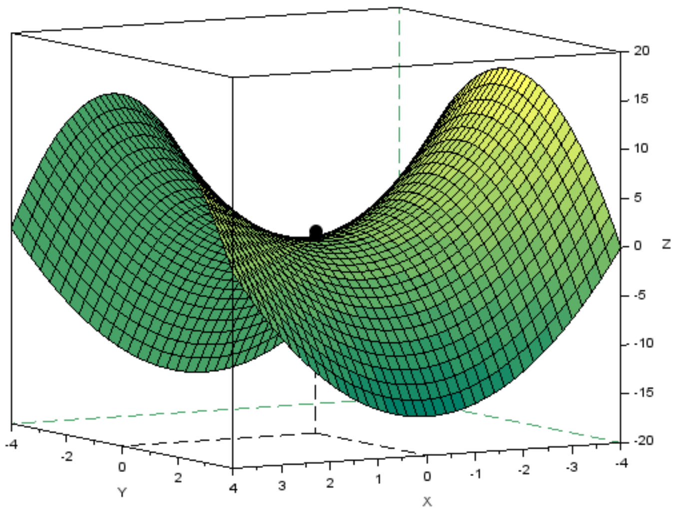
$$\nabla f(x, y) = (2x, -2y).$$

Ainsi $(x, y) \in \mathbb{R}^2$ est un point critique si, et seulement si :

$$\nabla f(x, y) = (0, 0) \quad \Leftrightarrow \quad (x, y) = (0, 0).$$

f a donc un unique point critique en $(0, 0)$, qui vaut $f(0, 0) = 0$. Mais 0 n'est ni un maximum, ni un minimum pour f puisque :

$$f(0, 1) = -1 < 0 < f(1, 0) = 1.$$



Graphe de $f : (x, y) \mapsto x^2 - y^2$.

Exemple. La fonction $f : (x, y) \in \mathbb{R}^2 \mapsto -x^2 - y^2$ est polynômiale donc $\mathscr{C}^1$. Son gradient est donné par $\nabla f(x, y) = (-2x, -2y)$ pour tout $(x, y) \in \mathbb{R}^2$. On cherche les points critiques de f :

$$\nabla f(x, y) = (0, 0) \quad \Leftrightarrow \quad (x, y) = (0, 0).$$

Donc f admet un unique point critique en $a = (0, 0)$. Et on a vu que f admet bien un maximum global en a.

Remarque. On avait calculé et représenté le plan tangent de f en $(0, 0)$: c'est le plan horizontal d'équation $z = 0$. Ca sera toujours le cas en un point critique. En effet, c est un point critique de f si, et seulement si, $\nabla f(c) = (0, \dots, 0)$, soit si, et seulement si, le plan tangent de f en c a pour équation :

$$x_{n+1} = f(c) + \langle \nabla f(c), x - c \rangle = f(c),$$

ce qui équivaut au fait que le plan tangent en c est horizontale.

Méthode. Comment déterminer les extremas d'une fonction ?

Pour déterminer les extremas d'une fonction f de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$:

Étape 1 : Classe $\mathscr{C}^1$ et calcul du gradient.
On justifie par théorèmes opératoires que f est de classe $\mathscr{C}^1$ et on calcule son gradient ∇f.

Étape 2 : Recherche des points critiques.
Afin de déterminer les points critiques de f, on résout l'équation :

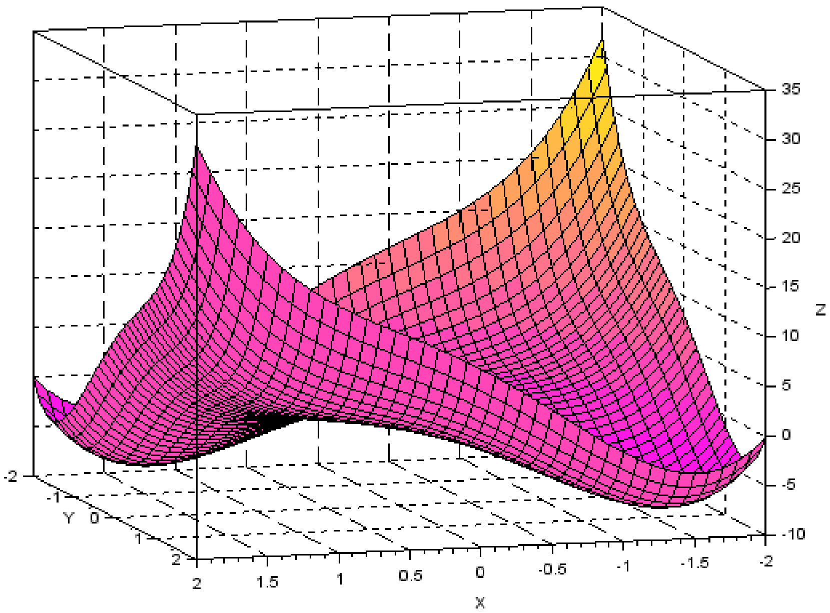
$$\nabla f(x) = (0, \ldots, 0).$$

Étape 3 : Nature de chacun des points critiques.
Parmi les points critiques se trouvent ceux en lesquels f atteint ses extrema globaux (s'il en existe). Il faut alors faire une étude particulière pour chaque point critique a de f.

- *Pour montrer que f n'admet pas d'extremum global en a, on peut selon les cas :*
  - *trouver des points $b \in \mathbb{R}^n$ tels que $f(b) > f(a)$ ou $f(b) < f(a)$, ce qui montrerait que f n'admet pas de maximum ou de minimum global en a.*
  - *étudier la limite de f dans une direction judicieusement choisie : si pour un certain $u \in \mathbb{R}^n$, la fonction $t \mapsto f(t \cdot u)$ tend vers $+\infty$ (resp. $-\infty$) quand $t \to +\infty$, alors f n'admet pas de maximum global (resp. minimum global). En pratique, on testera $u = (1, 0 \ldots, 0), \ldots, (0, \ldots, 0, 1), (1, \ldots, 1), \ldots$*
- *Pour montrer que f admet un extremum global en a, on déterminera le signe de $f(x) - f(a)$. Pour cela, on pourra essayer de faire apparaitre des carrés. Il pourra être utile d'effectuer le changement de variables $h = x - a$ afin de faciliter la mise en carrés.*

  Si le signe de $f(x) - f(a)$ reste constant pour tout $x \in \mathbb{R}^n$, alors f admet un extremum global en a.

Exercice. Déterminer les extremas éventuels de $f : (x, y) \in \mathbb{R}^2 \mapsto x^4 + y^4 - 2(x - y)^2$.



Graphe de f.

Remarque. Nous ne disposons pas d'outil plus élaboré pour déterminer si un point critique est ou non un extremum global. Nous verrons en TD d'autres méthodes pour obtenir la nature d'un point critique, mais aucune ne fonctionnera pour toutes les fonctions. Il faudra donc s'adapter selon les situations rencontrées, et prendre en compte les indications qui pourraient être fournies par l'énoncé.