Chapitre 40

Fonctions de deux variables

Plan du chapitre

Hypothèse

Dans tout ce chapitre, $\mathbb{R}^2$ est muni de leur sturcture d'espace euclidien (avec le produit scalaire canonique).
On notera notamment d la distance euclidienne et $||\cdot||$ la norme euclidienne.
De plus U est un ouvert non vide de $\mathbb{R}^2$ (cf plus loin pour une définition).

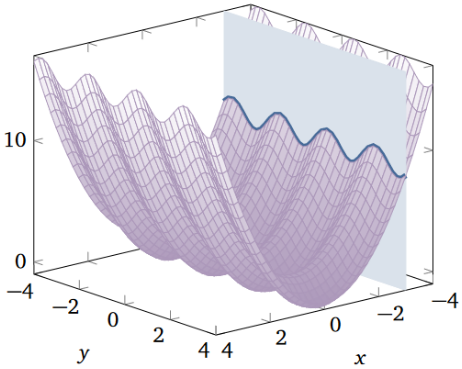
1 Graphe d'une fonction de deux variables

Il est usuel de représenter une fonction $f : \mathbb{R} \to \mathbb{R}$ (ou plutôt son graphe) par une courbe d'équation $y = f(x)$ dans le plan $\mathbb{R}^2$. Pour une fonction $f : \mathbb{R}^2 \to \mathbb{R}$, on la représentera dans $\mathbb{R}^3$ par une surface d'équation $z = f(x, y)$.

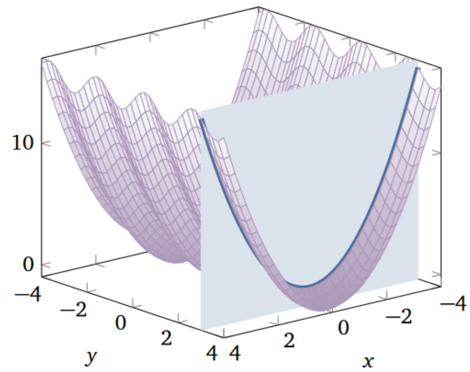
Exemple 1. Soit $f : (x, y) \mapsto x^2 + \sin(3y)$. Pour se représenter mentalement (ou construire) son graphe $\mathscr{S}$ en 3D, il est fréquent de réaliser des "coupes" de cette surface par des plans qui balaient l'espace $\mathbb{R}^3$ tout entier. Ainsi, pour tout $\lambda \in \mathbb{R}$:

- L'intersection de $\mathscr{S}$ avec le plan d'équation $x = \lambda$ est la courbe d'équation $z = \lambda^2 + \sin(3y)$ dans ce plan.
- L'intersection de $\mathscr{S}$ avec le plan d'équation $y = \lambda$ est la courbe d'équation $z = x^2 + \sin(3\lambda)$ dans ce plan.
- L'intersection de $\mathscr{S}$ avec le plan d'équation $z = \lambda$ est la courbe d'équation $\lambda = x^2 + \sin(3y)$ dans ce plan. On appelle cette courbe la ligne de niveau λ de f.

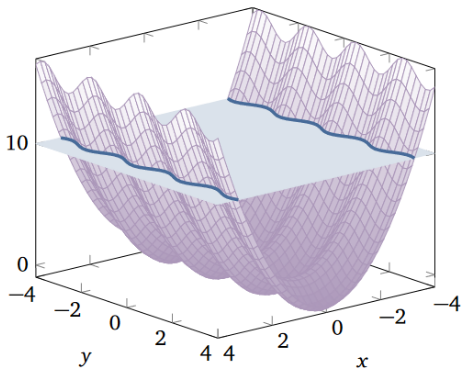
Intersection de $\mathscr{S}$ avec $x = -3$



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

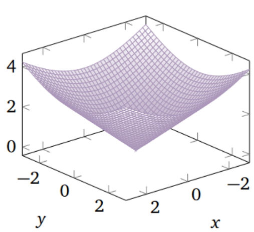


Intersection de $\mathscr{S}$ avec $z = 10$

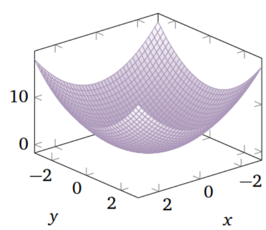


Voici d'autres exemples de surfaces.

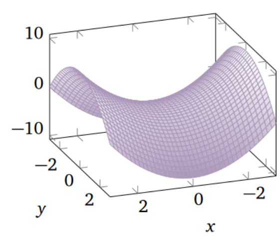
$z = \sqrt{x^2 + y^2}$



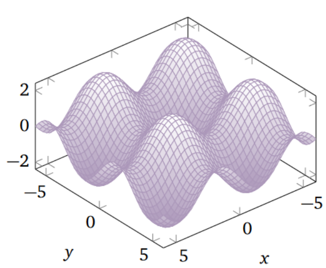
$z = x^2 + y^2$



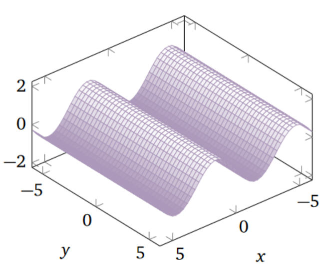
$z = x^2 - y^2$



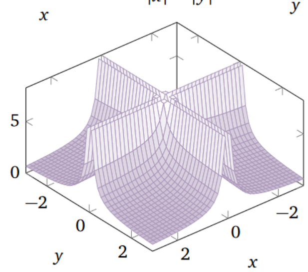
$z = \sin x + \sin y$



$z = \sin x$



$z = \frac{1}{|x|} + \frac{1}{|y|}$



2 Introduction à la topologie de $\mathbb{R}^2$

2.1 Boules ouvertes

Étant donné $M = (x,y) \in \mathbb{R}^2$ et $A = (x_0, y_0) \in \mathbb{R}^2$, on note

$$\begin{aligned} AM = d(A,M) &= ||M - A|| \\ &= ||\,(x - x_0,\, y - y_0)\,|| \\ &= \sqrt{(x-x_0)^2 + (y-y_0)^2} \end{aligned}$$

Il s'agit de la distance "usuelle" entre deux points.

Définition 40.1

Soit $A = (x_0, y_0) \in \mathbb{R}^2$ et $r > 0$. On appelle boule ouverte de centre A et de rayon r l'ensemble noté

$$B(A,r) := \left\{M \in \mathbb{R}^2 \mid AM < r\right\}$$

et on appelle boule fermée de centre A et de rayon r l'ensemble noté

$$\overline{B}(A,r) := \left\{M \in \mathbb{R}^2 \mid AM \leq r\right\}$$

On écrira parfois $B_{\mathbb{R}^2}(A,r)$ pour rappeler qu'il s'agit d'une boule dans $\mathbb{R}^2$.
Dans le plan $\mathbb{R}^2$, ces boules représentent en fait des disques (comme ce qu'on a vu dans $\mathbb{C}$ avec les disques ouverts ou fermés). Toutefois, ce formalisme se généralise sans difficulté à $\mathbb{R}^n$ avec $n \geq 3$, donc on introduit dès maintenant ce vocabulaire.

De même, pour tout réel x, on a

$$B_{\mathbb{R}}(x,r) = \ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\overline{B}_{\mathbb{R}}(x,r) = \ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Définition 40.2

Soit $A \in \mathbb{R}^2$. On appelle voisinage de A toute partie V de $\mathbb{R}^2$ qui contient une boule ouverte de centre A.

Notation. On notera dans ce cours $\mathcal{V}_{\mathbb{R}^2}(A)$ l'ensemble des voisinages de A dans $\mathbb{R}^2$. On a donc

$$V \in \mathcal{V}_{\mathbb{R}^2}(A) \iff \exists r > 0 \qquad B_{\mathbb{R}^2}(A,r) \subset V$$

La définition d'un voisinage de $\mathbb{R}$ est similaire : pour tout réel x,

$$V \in \mathcal{V}_{\mathbb{R}}(x) \iff \exists \varepsilon > 0 \qquad]x - \varepsilon, x + \varepsilon[\subset V$$

2.2 Ouvert de $\mathbb{R}^2$

Définition 40.3 (Ouvert)

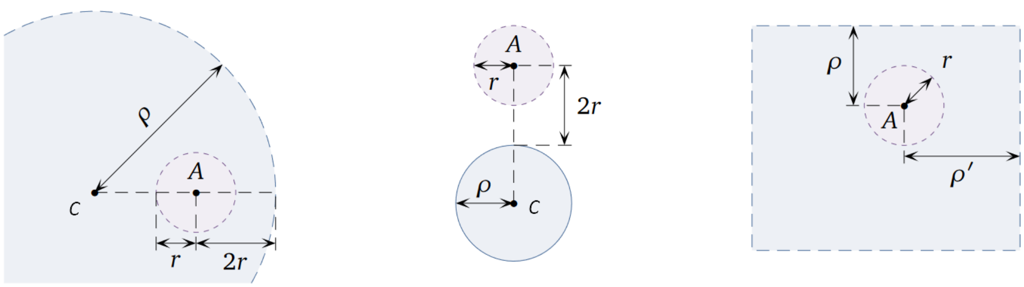
Soit Ω une partie de $\mathbb{R}^2$. On dit que Ω est (un) ouvert si

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

Cela revient à dire que Ω est un voisinage de chacun de ces points. Bien entendu, le rayon r de la boule dépend du point A considéré.

Exemple 2.
- Toute boule ouverte est un ouvert.
- Le complémentaire d'une boule fermée est un ouvert.
- Si I et J sont des intervalles ouverts, alors $I \times J$ est un ouvert de $\mathbb{R}^2$

La preuve consiste à chaque fois à prendre un point quelconque A de ces ensembles et de déterminer $r > 0$ tel que $B(A,r)$ est contenu dans cet ensemble. On se contentera d'un dessin pour se donner l'intuition géométrique de la preuve.



À gauche : $\Omega = B(C,\rho)$. Au milieu : $\Omega = \mathbb{R}^2 \setminus \overline{B}(C,\rho)$. À droite : $\Omega = I \times J$

2.3 Limite finie en un point

On rappelle que dans ce chapitre, on suppose que Ω est un ouvert non vide de $\mathbb{R}^2$.

Définition 40.4 (Limie finie en un point)

Soit $\Omega \subset \mathbb{R}^2$ et $f : \Omega \to \mathbb{R}$ une fonction. Soit $A \in \Omega$ et $\ell \in \mathbb{R}$. On dit que f tend vers ℓ en A si :

$$\forall V_\ell \in \mathcal{V}_{\mathbb{R}}(\ell) \qquad \exists V_A \in \mathcal{V}_{\mathbb{R}^2}(A) \qquad \forall M \in V_A \cap \Omega \qquad f(M) \in V_\ell$$

ou encore :

$$\forall \varepsilon > 0 \qquad \exists r > 0 \qquad \forall M \in \Omega \qquad AM < r \implies |f(M) - \ell| < \varepsilon$$

Cette dernière assertion est beaucoup plus proche de la notion de limite qu'on connait. En posant $M = (x,y)$ et $A = (x_0,y_0)$, on peut aller encore plus loin dans le parallèle :

$$\forall \varepsilon > 0 \qquad \exists r > 0 \qquad \forall (x,y) \in \Omega \qquad ||(x,y) - (x_0,y_0)|| < r \implies |f(x,y) - \ell| < \varepsilon$$

Notation. Lorsque f tend vers ℓ en A, on notera

$$\lim_A f = \lim_{M \to A} f(M) = \ell$$

ou encore, en posant $M = (x, y)$ et $A = (x_0, y_0)$:

$$\lim_{(x_0,y_0)} f = \lim_{(x,y)\to(x_0,y_0)} f(x, y) = \ell$$

Propriété 40.5

La majorité des résultats vus pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ restent valides : unicité de la limite, opérations sur les limites (combinaison linéaire, produit, quotient, composition à gauche avec une fonction de $\mathbb{R}$ dans $\mathbb{R}$), passage dans une inégalité large / stricte, théorème d'encadrement.

Démonstration. Dans la majorité des cas, il suffit de remplacer les valeurs absolues de l'ensemble de départ par la norme de $\mathbb{R}^2$. $\square$

On pourrait également définir le cas d'une limite $+\infty$ ou $-\infty$ en un point, mais cela n'est pas l'objet du chapitre

Méthode

Pour montrer qu'une fonction f tend vers ℓ en (x_0, y_0), on peut repartir de la définition. On peut parfois également majorer $|f(x, y) - \ell|$ par une expression qui dépend de $||(x - x_0, y - y_0)||$ et qui tend vers 0 quand cette norme tend vers 0.
Pour montrer que f n'admet pas de limite (finie ou non) en (x_0, y_0), il suffit d'exhiber deux couples de suites (x_n, y_n) et (x'_n, y'_n) qui tendent tous les deux vers (x_0, y_0) mais pour lesquels

$$\lim_{n\to+\infty} f(x_n, y_n) \neq \lim_{n\to+\infty} f(x'_n, y'_n)$$

Cette dernière approche est similaire à la caractérisation séquentielle de la limite pour une fonction f de $\mathbb{R}$ dans $\mathbb{R}$. Très souvent, on a $(x_0, y_0) = (0, 0)$ et on obtient une contradiction en calculant :

$$\lim_{x\to 0} f(x, 0) \qquad \lim_{y\to 0} f(0, y) \qquad \lim_{x\to 0} f(x, x)$$

ce qui correspond par exemple aux couples de suites $\left(\frac{1}{n}, 0\right)$, $\left(0, \frac{1}{n}\right)$, et $\left(\frac{1}{n}, \frac{1}{n}\right)$ qui tendent tous vers $(0, 0)$.

Remarque. Attention ! Faire tendre $M = (x, y)$ vers $A = (x_0, y_0)$ ne revient pas à faire tendre x vers x_0, puis y vers y_0 (ou l'inverse). Le point M peut se rapprocher de A de bien des manières, il n'est pas obligé de le faire par les directions horizontale et verticale.

Exemple 3. On définit la fonction

$$f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$$
$$(x, y) \mapsto \frac{xy}{x^2 + y^2}$$

Montrer que f n'a pas de limite en 0.

2.4 Continuité d'une fonction $f : \mathbb{R}^2 \to \mathbb{R}$

Définition 40.6 (Continuité en un point)

Soit $\Omega \subset \mathbb{R}^2$ et $f : \Omega \to \mathbb{R}$ une fonction. Soit $A \in \Omega$. On dit que f est continue en A si $\lim\limits_{M \to A} f(M) = f(A)$, c'est-à-dire :

$$\forall \varepsilon > 0 \qquad \exists r > 0 \qquad \forall M \in \Omega \qquad AM < r \implies |f(M) - f(A)| < \varepsilon$$

En posant $A = (x_0, y_0)$, on peut réécrire $\lim\limits_{(x,y) \mapsto (x_0,y_0)} f(x,y) = f(x_0,y_0)$ d'une autre manière :

$$\forall \varepsilon > 0 \qquad \exists r > 0 \qquad \forall (x,y) \in \Omega \qquad ||(x,y) - (x_0,y_0)|| < r \implies |f(x,y) - f(x_0,y_0)| < \varepsilon$$

Définition 40.7 (Continuité sur une partie de $\mathbb{R}^2$)

Soit $\Omega \subset \mathbb{R}^2$ et $f : \Omega \to \mathbb{R}$. On dit que f est continue sur Ω si elle est continue en tout point de Ω.
On notera $\mathcal{C}(\Omega, \mathbb{R})$ l'ensemble des fonctions continues sur Ω.

Propriété 40.8

La majorité des résultats vus pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ restent valides : la combinaison linéaire de fonctions continues est une fonction continue. De même pour le produit, le quotient, la composition à gauche avec une fonction continue de $\mathbb{R}$ dans $\mathbb{R}$.

Exemple 4. Les fonctions $(x,y) \mapsto x$ et $(x,y) \mapsto y$ sont continues sur $\mathbb{R}^2$.

Exemple 5. A fortiori, par produit et combinaison linéaire, toute fonction polynômiale en les deux variables x et y est continue sur $\mathbb{R}^2$. Par exemple
$$(x,y) \mapsto x^2 - 4x^3y + 2y^2 - x + 1$$
est continue sur $\mathbb{R}^2$.

Exemple 6. La norme $u \mapsto ||u||$ est continue sur $\mathbb{R}^2$.

Exemple 7. On définit $f : \mathbb{R}^2 \to \mathbb{R}$ par

$$f(x,y) = \begin{cases} 0 & \text{si } (x,y) = (0,0) \\ \dfrac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \end{cases}$$

Par ce qu'on a vu à l'exemple 3, f n'est pas continue en 0. En effet, $f(t,t) = \dfrac{1}{2} \xrightarrow[t\to 0]{} \dfrac{1}{2} \neq f(0,0)$.

3 Dérivée partielle

3.1 Dérivée directionnelle (selon un vecteur)

On aimerait définir la dérivée d'une fonction $f : \Omega \to \mathbb{R}$ de manière similaire au cas d'une fonction à une variable. On pourrait être tenté de définir
$$\text{„} \quad \lim_{M\to A} \frac{f(M)-f(A)}{M-A} \quad \text{„}$$
Mais il y a plusieurs problèmes. D'une part, on divise par $M - A \in \mathbb{R}^2$ ce qui n'a pas de sens. D'autre part, même pour des fonctions f très "gentilles", cette limite n'est pas forcément unique, car cela dépend de la manière dont M tend vers A. La définition qui suit permet toutefois de se ramener au cas réel en imposant à M de tendre vers A en restant dans la même "direction".

Dans la définition ci-dessous, on emploie volontairement une notation vectorielle $\overrightarrow{v}$ mais cela est purement arbitraire.

Définition 40.9 (Dérivée directionnelle selon un vecteur)

Soit $f : \Omega \to \mathbb{R}$, $A = (x_0, y_0) \in \Omega$ et $\overrightarrow{v} \in \mathbb{R}^2$. On dit que <u>$f$ est dérivable en A dans la direction $\overrightarrow{v}$</u> si la limite
$$\lim_{t\to 0} \frac{f(A+t\overrightarrow{v}) - f(A)}{t}$$
existe et est finie. Lorsque c'est le cas, on appelle <u>dérivée de f en A dans la direction $\overrightarrow{v}$</u> le réel
$$D_{\overrightarrow{v}} f(A) = \lim_{t\to 0} \frac{f(A+t\overrightarrow{v}) - f(A)}{t} \quad \in \mathbb{R}$$

Dire que f est dérivable en A dans la direction $\overrightarrow{v}$ revient à dire que la fonction $t \mapsto f(A + t\overrightarrow{v})$ est dérivable en 0.

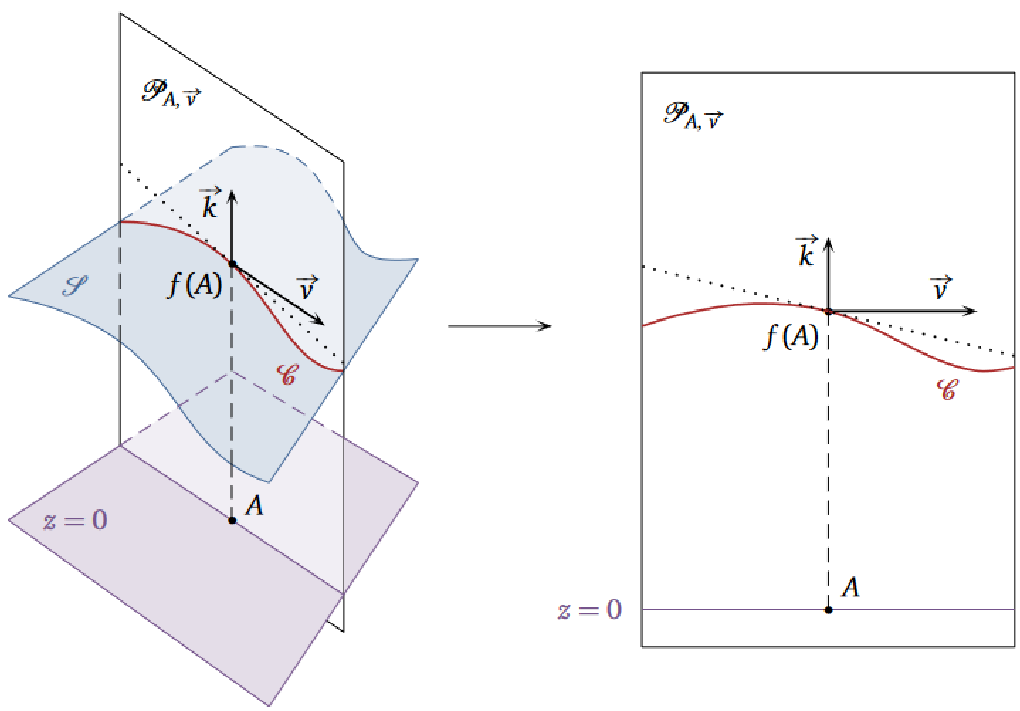
- Si $\overrightarrow{v}$ est le vecteur nul, alors $t \mapsto f(A+t\overrightarrow{v}) = f(A)$ est constante sur $\mathbb{R}$ tout entier, donc dérivable en 0. On a donc $D_{\overrightarrow{0}} f(A) = 0$.

- Si $\overrightarrow{v}$ est non nul, alors la fonction $t \mapsto f(A+t\overrightarrow{v})$ est définie *au voisinage de* 0. En effet, comme Ω est ouvert, il existe $r > 0$ tel que $B(A,r) \subset \Omega$. Alors, en posant $\alpha = \dfrac{r}{||\overrightarrow{v}||}$, on montre que

$$\forall t \in]-\alpha,\alpha[\qquad A+t\overrightarrow{v} \in B(A,r)$$

car $d(A,A+t\overrightarrow{v}) = ||t\overrightarrow{v}|| = |t| \times ||\overrightarrow{v}|| < \alpha||\overrightarrow{v}|| = r$. Ainsi, la fonction $t \mapsto f(A+t\overrightarrow{v})$ est définie sur $]-\alpha,\alpha[$ et considérer sa dérivée en 0 a bien un sens.

Que représente concrètement $D_{\overrightarrow{v}}f(A)$? Le principe est de considérer le graphe (i.e. la surface) qui représente f autour de A, de faire une coupe de cette surface par le plan vertical $\mathscr{P}_{A,\overrightarrow{v}}$ qui passe par A et qui contient le vecteur $\vec{v}$ (en toute rigueur, $\overrightarrow{v} \in \mathbb{R}^2$ mais on le voit ici comme un vecteur de $\mathbb{R}^3$ dont la composante en z est nulle). La coupe permet d'obtenir une courbe, celle de $t \mapsto f(A+t\overrightarrow{v})$. La pente de la tangente en 0 donnera le nombre $D_{\overrightarrow{v}}f(A)$.



Attention toutefois, on pourrait croire que $D_{\overrightarrow{v}}f(A)$ ne dépend que de la direction de $\overrightarrow{v}$ et non de la norme de $\overrightarrow{v}$, mais ce n'est pas le cas. Au contraire, $D_{\overrightarrow{v}}f(A)$ est proportionnel à $||\overrightarrow{v}||$ comme le montre l'exemple ci-dessous.

Exemple 8. Calculer la dérivée directionnelle de $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = e^x + y$ au point $A = (0,0)$ selon la direction $v = (1,1)$, puis selon le vecteur $w = (2,2)$.

3.2 Dérivée partielle

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

Définition 40.10

Soit $f : \Omega \to \mathbb{R}$ et $A = (x_0, y_0) \in \Omega$.

- La dérivée directionnelle de f en A selon $\overrightarrow{i} = (1,0)$, lorsqu'elle existe, est appelée dérivée partielle en A selon (la direction) x. On la note

$$\frac{\partial f}{\partial x}(x_0, y_0) := \lim_{t \to 0} \frac{f(x_0 + t, y_0) - f(x_0, y_0)}{t}$$

- La dérivée directionnelle de f en A selon $\overrightarrow{j} = (0,1)$, lorsqu'elle existe, est appelée dérivée partielle en A selon (la direction) y. On la note

$$\frac{\partial f}{\partial y}(x_0, y_0) := \lim_{t \to 0} \frac{f(x_0, y_0 + t) - f(x_0, y_0)}{t}$$

Attention à bien utiliser des "d ronds" donc $\dfrac{\partial f}{\partial x}$, et non des "d droits" tels que $\dfrac{\mathrm{d}f}{\mathrm{d}x}$.

Méthode

Dans la pratique, calculer $\dfrac{\partial f}{\partial x}(x_0, y_0)$ revient à calculer la dérivée de $x \mapsto f(x, y_0)$ en x_0 : la valeur de y est donc fixée à y_0 et n'influence pas la dérivation selon x.
Les règles de calcul de dérivation sont exactement les mêmes (combinaison linéaire, produit, inverse, composition, etc.)

Exemple 9. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = e^{xy^2}$. On peut vérifier que f admet des dérivées partielles en tout point $(x,y) \in \mathbb{R}^2$, et

$$\frac{\partial f}{\partial x}(x,y) = \ldots\ldots\ldots\ldots\ldots\ldots \qquad \frac{\partial f}{\partial y}(x,y) = \ldots\ldots\ldots\ldots\ldots\ldots$$

Attention, dans l'écriture $\dfrac{\partial f}{\partial x}(x,y)$, la lettre x joue deux rôles très différents : dans ∂x, elle précise qu'on dérive dans la direction horizontale. Dans (x,y), elle donne l'abscisse du point en lequel on évalue $\dfrac{\partial f}{\partial x}$. On est donc en droit d'écrire, avec (x_1,y_1) et (p,q) dans $\mathbb{R}^2$:

$$\frac{\partial f}{\partial x}(1,3) \qquad \frac{\partial f}{\partial x}(x_1,y_1) \qquad \frac{\partial f}{\partial x}(p,q) \qquad \cdots$$

Enfin, les notations $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ signifient qu'on calcule la dérivée de f par rapport aux variables x et y : ce sont les notations du programme. Elles sont adaptées aux fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$, mais d'autres notations plus générales existent :

$$\frac{\partial f}{\partial x_1} \quad \text{et} \quad \frac{\partial f}{\partial x_2} \qquad \left(\text{voire } \frac{\partial f}{\partial x_3}, \cdots, \frac{\partial f}{\partial x_m} \text{ si } f : \mathbb{R}^m \to \ldots\right)$$

$$\partial_1 f \quad \text{et} \quad \partial_2 f \qquad (\text{voire } \partial_3 f, \cdots, \partial_m f \text{ si } f : \mathbb{R}^m \to \ldots)$$

Exemple 10. On définit $f : \mathbb{R}^2 \to \mathbb{R}$ par

$$f(x,y) = \begin{cases} 0 & \text{si } (x,y) = (0,0) \\ \dfrac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \end{cases}$$

On a vu que f n'est pas continue en 0. Pourtant, f admet des dérivées partielles en tout point de $\mathbb{R}^2$.

4 Fonction de classe $\mathcal{C}^1$

4.1 Définition

Si une fonction $f : \Omega \to \mathbb{R}$ admet des dérivées partielles en tout point de Ω, alors on peut définir les fonctions :

$$\frac{\partial f}{\partial x} : \begin{array}{l} \Omega \to \mathbb{R} \\ (x,y) \mapsto \dfrac{\partial f}{\partial x}(x,y) \end{array} \qquad \frac{\partial f}{\partial y} : \begin{array}{l} \Omega \to \mathbb{R} \\ (x,y) \mapsto \dfrac{\partial f}{\partial y}(x,y) \end{array}$$

Définition 40.11

La fonction $f : \Omega \to \mathbb{R}$ est dite <u>de classe $\mathcal{C}^1$</u> si elle admet des dérivées partielles en tout point de Ω et si les fonctions $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont continues sur Ω.
On note $\mathcal{C}^1(\Omega, \mathbb{R})$ ou $\mathcal{C}^1(\Omega)$ l'ensemble des fonctions de classe $\mathcal{C}^1$ sur Ω.

Exemple 11. Toute fonction polynômiale en les variables x et y est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$.
Toute fonction rationnelle en les variables x et y est de classe $\mathcal{C}^1$ sur son ensemble de définition (qui est un ouvert).

Exemple 12. On définit $f : \mathbb{R}^2 \to \mathbb{R}$ par

$$f(x,y) = \begin{cases} 0 & \text{si } (x,y) = (0,0) \\ \dfrac{2xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \end{cases}$$

On a vu que f n'est pas continue en 0. f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2 \setminus \{(0,0)\}$ en tant que fonction rationnelle en les variables x et y.

Toutefois, on va montrer que f n'est pas $\mathcal{C}^1$ en $(0,0)$. On a vu que

$$\forall (x,y) \in \mathbb{R}^2 \qquad \frac{\partial f}{\partial x}(x,y) = \begin{cases} 0 & \text{si } (x,y) = (0,0) \\ \dfrac{y(y^2-x^2)}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0) \end{cases}$$

On voit ainsi que $\dfrac{\partial f}{\partial x}$ n'est pas continue en $(0,0)$ car :

Propriété 40.12

Soit f, g deux fonctions de $\mathcal{C}^1(\Omega, \mathbb{R})$ et $\lambda \in \mathbb{R}$. Alors $f+g$, λf, fg sont encore de classe $\mathcal{C}^1$.

La composition avec des fonctions de classe $\mathcal{C}^1$ conserve le caractère $\mathcal{C}^1$ également, mais il y a tant de manières de composer et de formules de dérivations associées qu'on les traite dans une section à part.

4.2 Intermède hors programme : règle de la chaine

La section suivante, i.e. la section 4.3 est dédiée aux règles de dérivation par composition lorsqu'on manipule des fonctions de deux variables. Ces règles portent le nom de “règle de la chaine”. Trois versions sont au programme, mais ces règles découlent toutes de la même règle. Dans cette section 4.2, on donne la règle générale qui permet de déduire les trois règles particulières.

Définition 40.13 (Applications composantes)

Soit $f : \mathbb{R}^m \to \mathbb{R}^n$. Pour tout $j \in [\![1,n]\!]$, on définit l'application $f_j : \mathbb{R}^m \to \mathbb{R}$ qui à tout $x \in \mathbb{R}^m$ associe la j-ième composante de $f(x)$ selon la base canonique de $\mathbb{R}^n$.
Dit autrement, on a

$$\forall x \in E \qquad f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix}$$

Les applications $f_1, \cdots, f_n$ sont appelées les applications composantes de f.

On écrit plus simplement $f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = (f_1, \cdots, f_n)$.

Propriété 40.14

Soit $f : \mathbb{R}^m \to \mathbb{R}^n$. Alors f est de classe $\mathcal{C}^1$ si et seulement si ses applications composantes $f_1, \cdots, f_n$ le sont et

$$\forall A \in \mathbb{R}^m \quad \forall i \in [\![1,m]\!] \qquad \frac{\partial f}{\partial x_i}(A) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_i}(A) \\ \vdots \\ \dfrac{\partial f_n}{\partial x_i}(A) \end{pmatrix}$$

On a donc pour tout $i \in [\![1,m]\!]$ $\quad \dfrac{\partial f}{\partial x_i} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_i} \\ \vdots \\ \dfrac{\partial f_n}{\partial x_i} \end{pmatrix} = \left(\dfrac{\partial f_1}{\partial x_i}, \cdots, \dfrac{\partial f_n}{\partial x_i}\right)$.

Exemple 13. On définit $f : \mathbb{R}^2 \to \mathbb{R}^3$ par $f(x,y) = (2x^2 + xy, 3x - y^2, y)$. Les applications composantes de f sont définies par :

$$f_1(x,y) = 2x^2 + xy \qquad f_2(x,y) = 3x - y^2 \qquad f_3(x,y) = y$$

Ces applications sont de classe $\mathcal{C}^1$, donc f aussi et

$$\frac{\partial f}{\partial x_1}(x,y) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(x,y) \\ \dfrac{\partial f_2}{\partial x_1}(x,y) \\ \dfrac{\partial f_3}{\partial x_1}(x,y) \end{pmatrix} = \begin{pmatrix} 4x+y \\ 3 \\ 0 \end{pmatrix} \qquad \frac{\partial f}{\partial x_2}(x,y) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_2}(x,y) \\ \dfrac{\partial f_2}{\partial x_2}(x,y) \\ \dfrac{\partial f_3}{\partial x_2}(x,y) \end{pmatrix} = \begin{pmatrix} x \\ -2y \\ 1 \end{pmatrix}$$

Règle de la chaine générale On considère donc $m, n, p \in \mathbb{N}^*$, $f : \mathbb{R}^m \to \mathbb{R}^n$, et $g : \mathbb{R}^n \to \mathbb{R}^p$. On pose enfin $h = g \circ f$. On a donc

$$\underbrace{\mathbb{R}^m \xrightarrow{f} \mathbb{R}^n \xrightarrow{g} \mathbb{R}^p}_{h}$$

Ensemble	$\mathbb{R}^m$	$\mathbb{R}^n$	$\mathbb{R}^p$
Notation des dérivées partielles	$\dfrac{\partial}{\partial x_1}, \cdots, \dfrac{\partial}{\partial x_m}$	$\dfrac{\partial}{\partial y_1}, \cdots, \dfrac{\partial}{\partial y_n}$	
Indice associé	$1 \leq i \leq m$	$1 \leq j \leq n$	$1 \leq k \leq p$
Point considéré	A	$f(A)$	$g(f(A))$

Propriété 40.15

Si $f : \mathbb{R}^m \to \mathbb{R}^n$ est de classe $\mathcal{C}^1$ et si $g : \mathbb{R}^n \to \mathbb{R}^p$ est de classe $\mathcal{C}^1$, alors $h = g \circ f$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^m$, et pour tout $A \in \mathbb{R}^m$,

$$\forall i \in [\![1, m]\!] \quad \forall k \in [\![1, p]\!] \qquad \frac{\partial h_k}{\partial x_i}(A) = \sum_{j=1}^{n} \frac{\partial g_k}{\partial y_j}(f(A)) \times \frac{\partial f_j}{\partial x_i}(A)$$

Exemple 14. Soit $g : \mathbb{R}^2 \to \mathbb{R}^2$ une fonction de classe $\mathcal{C}^1$. On pose

$$h(x, y) = g(x^2 - y^2, 2xy)$$

Déterminer $\dfrac{\partial h_1}{\partial x_2}(x, y)$ en fonction des dérivées partielles de g.

Remarque. Si $m = 1$, alors l'ensemble de départ des fonctions f et h est $\mathbb{R}$. Cela entraine que $\dfrac{\partial f}{\partial x_1} = f'$ et $\dfrac{\partial h}{\partial x_1} = h'$: ces dérivées partielles correspondent aux dérivées classiques. De même, si $n = 1$, on a $\dfrac{\partial g}{\partial y_1} = g'$. En particulier, si $m = n = p = 1$, alors $f = f_1$, $g = g_1$ et $h = h_1$, et on a

$$h'(A) = g'(f(A)) \times f'(A)$$

On retrouve la formule usuelle pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

4.3 Dérivation de fonctions composées

On donne à présent les versions de la règle de la chaine qui sont au programme de MPSI.

Le premier cas concerne la situation où on compose f à gauche par une fonction $\theta : \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^1$, donc à $\mathbb{R}^2 \xrightarrow{f} \mathbb{R} \xrightarrow{\theta} \mathbb{R}$. Comme f et θ sont à valeurs dans $\mathbb{R}$, elles n'ont qu'une application composante : $f = f_1$ et $\theta = \theta_1$. Enfin, on a $\dfrac{\partial \theta}{\partial y_1} = \theta'$

Propriété 40.16 (Post-composition $\mathbb{R}^2 \xrightarrow{f} \mathbb{R} \xrightarrow{\theta} \mathbb{R}$)

Soit $f \in \mathcal{C}^1(\Omega, \mathbb{R})$, soit I un intervalle contenant $f(\Omega)$ et $\theta \in \mathcal{C}^1(I, \mathbb{R})$. Alors $\theta \circ f \in \mathcal{C}^1(\Omega, \mathbb{R})$ et pour tout $A \in \Omega$

$$\frac{\partial(\theta \circ f)}{\partial x_1}(A) = \ldots\ldots\ldots\ldots\ldots\ldots \qquad \frac{\partial(\theta \circ f)}{\partial x_2}(A) = \ldots\ldots\ldots\ldots\ldots\ldots$$

En particulier, avec $\theta : x \mapsto \dfrac{1}{x}$, on voit que si f ne s'annule par sur Ω, alors $\dfrac{1}{f}$ est de classe $\mathcal{C}^1$ sur Ω.

Le second cas concerne la situation où on compose g à droite par une fonction $\gamma : \mathbb{R} \to \mathbb{R}^2$ de classe $\mathcal{C}^1$, donc à $\mathbb{R} \xrightarrow{\gamma} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}$. On a donc $\gamma(t) = (\gamma_1(t), \gamma_2(t))$: cela représente un arc paramétré, $\gamma_1(t)$ étant l'abscisse et $\gamma_2(t)$ étant l'ordonnée (on trouve aussi la notation $\gamma(t) = (x(t), y(t))$ mais elle peut générer des confusions). Comme g est à valeurs dans $\mathbb{R}$, elle n'a qu'une application composante : $g = g_1$. De plus, comme γ part de $\mathbb{R}$, on a $\dfrac{\partial \gamma}{\partial x_1} = \gamma'$.

Propriété 40.17 (Pré-composition $\mathbb{R} \xrightarrow{\gamma} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}$)

Soit I un intervalle ouvert (non vide non singleton). Soit $g \in \mathcal{C}^1(\Omega, \mathbb{R})$ et $\gamma = (\gamma_1, \gamma_2) \in \mathcal{C}^1(I, \mathbb{R}^2)$. On suppose que $\gamma(I) \subset \Omega$. Alors $g \circ \gamma \in \mathcal{C}^1(I, \mathbb{R})$ et pour tout $a \in I$

$$(g \circ \gamma)'(a) = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Le dernier cas concerne la situation où on compose g à droite par une fonction $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ de classe $\mathcal{C}^1$, donc à $\mathbb{R}^2 \xrightarrow{\varphi} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}$. On a donc $\varphi(t) = (\varphi_1(t), \varphi_2(t))$ avec $\varphi_1 : \mathbb{R}^2 \to \mathbb{R}$ et $\varphi_2 : \mathbb{R}^2 \to \mathbb{R}$. Comme g est à valeurs dans $\mathbb{R}$, elle n'a qu'une application composante : $g = g_1$.

Propriété 40.18 (Pré-composition $\mathbb{R}^2 \xrightarrow{\varphi} \mathbb{R}^2 \xrightarrow{g} \mathbb{R}$)

Soit Ω, Ω' deux ouverts non vides de $\mathbb{R}^2$. Soit $\varphi \in \mathcal{C}^1(\Omega, \mathbb{R})$ et $g \in \mathcal{C}^1(\Omega', \mathbb{R})$. On suppose que $\varphi(\Omega) \subset \Omega'$ Alors $g \circ \varphi \in \mathcal{C}^1(\Omega, \mathbb{R})$ et pour tout $A \in \Omega$

$$\frac{\partial(g \circ \Phi)}{\partial x_1}(A) = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\frac{\partial(g \circ \Phi)}{\partial x_2}(A) = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

4.4 Développement limité à l'ordre 1

Propriété 40.19

Soit $f : \Omega \to \mathbb{R}$ et $A = (x_0, y_0) \in \Omega$. Si f est de classe $\mathcal{C}^1$, alors la fonction f admet un développement limité en A d'ordre 1 : il existe $\varepsilon : \Omega \to \mathbb{R}$ tel que pour tout $(h,k) \in \mathbb{R}^2$ vérifiant $(x_0+h, y_0+k) \in \Omega$, on a :

$$f(x_0+h, y_0+k) = f(x_0,y_0) + h\frac{\partial f}{\partial x}(x_0,y_0) + k\frac{\partial f}{\partial y}(x_0,y_0) + ||(h,k)||\varepsilon(h,k)$$

avec $\lim\limits_{(h,k)\to(0,0)} \varepsilon(h,k) = 0$.

Démonstration. Preuve admise et non exigible. □

Le terme $||(h,k)||\varepsilon(h,k)$ se réécrit parfois $\underset{(h,k)\to(0,0)}{o}(||(h,k)||)$. Enfin, on peut poser $x = x_0 + h$ et $y = y_0 + k$ pour réécrire ce DL en

$$\forall (x,y) \in \Omega \qquad f(x,y) = f(x_0,y_0) + (x-x_0)\frac{\partial f}{\partial x}(x_0,y_0) + (y-y_0)\frac{\partial f}{\partial y}(x_0,y_0) + \underset{(x,y)\to(x_0,y_0)}{o}(||(x,y)-(x_0,y_0)||)$$

Corollaire 40.20

Si $f \in \mathcal{C}^1(\Omega, \mathbb{R})$, alors f est continue sur Ω.

Démonstration. Montrons que pour tout $(x_0,y_0) \in \Omega$, f est continue en (x_0,y_0). Soit $(x,y) \in \Omega$. Lorsque $(x,y) \to (x_0,y_0)$ alors en particulier $x \to x_0$ et $y \to y_0$, si bien que

$$\lim_{(x,y)\to(x_0,y_0)} \left[(x-x_0)\frac{\partial f}{\partial x}(x_0,y_0) + (y-y_0)\frac{\partial f}{\partial y}(x_0,y_0) + \underset{(x,y)\to(x_0,y_0)}{o}(||(x,y)-(x_0,y_0)||)\right] = 0$$

d'où, par le DL ci-dessus, on a

$$\lim_{(x,y)\to(x_0,y_0)} [f(x,y) - f(x_0,y_0)] = 0$$

Donc f est continue en (x_0,y_0). D'où f est continue sur Ω par arbitraire sur (x_0,y_0). □

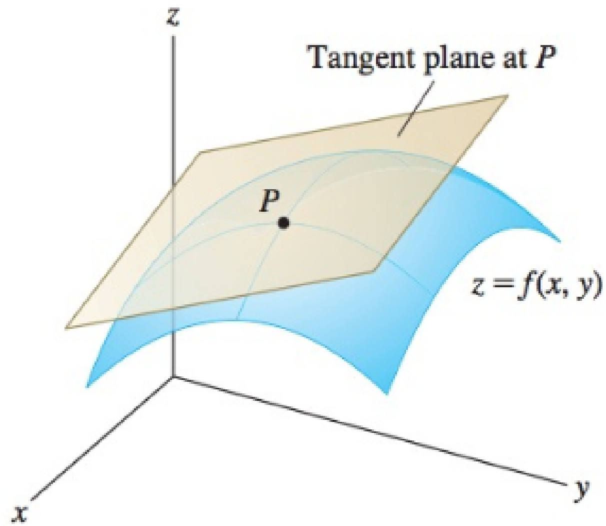
Remarque. On pose ici $A = (a,b) \in \Omega$ pour simplifier quelques notations. Pour $f \in \mathcal{C}^1(\Omega,\mathbb{R})$, un DL d'ordre 1 en A se réécrit

$$f(x,y) = f(a,b) + (x-a)\frac{\partial f}{\partial x}(a,b) + (y-b)\frac{\partial f}{\partial y}(a,b) + o([\ldots])$$

Ceci montre que, au premier ordre, pour des points (x,y) voisins de (a,b), la courbe de f est très proche du plan d'équation

$$z = f(a,b) + (x-a)\frac{\partial f}{\partial x}(a,b) + (y-b)\frac{\partial f}{\partial y}(a,b)$$

Ce plan est appelé le plan tangent (du graphe) de f en (a,b).



Exemple 15. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = \sin(x+2y)$. Déterminer le DL à l'ordre 1 en $(0,0)$ de f puis son plan tangent en $(0,0)$.

4.5 Gradient

Définition 40.21

Soit $f \in \mathcal{C}^1(\Omega, \mathbb{R})$. On définit le gradient de f comme étant l'application :

$$\begin{aligned} \nabla f : \quad & \Omega \to \mathbb{R}^2 \\ & (a,b) \mapsto \left(\frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b) \right) \end{aligned}$$

Ainsi, en tout point $(a,b) \in \Omega$, on dispose d'un vecteur $\nabla f(a,b)$ de $\mathbb{R}^2$. On dit que ∇f est un champ de vecteurs. Dans la suite, on note $\langle \cdot \mid \cdot \rangle$ le produit scalaire de $\mathbb{R}^2$.

Propriété 40.22

Soit $f \in \mathcal{C}^1(\Omega, \mathbb{R})$. Alors en tout point $A = (a,b) \in \Omega$ et pour tout $\overrightarrow{v} = (v_1, v_2) \in \mathbb{R}^2$, f admet une dérivée directionnelle en A selon $\overrightarrow{v}$, et on a :

$$D_{\overrightarrow{v}} f(a,b) = \langle \nabla f(a,b) \mid \overrightarrow{v} \rangle = v_1 \frac{\partial f}{\partial x}(a,b) + v_2 \frac{\partial f}{\partial y}(a,b)$$

Démonstration. Comme f est $\mathcal{C}^1$, f admet un développement limité en (a,b). On a donc (lorsque $(a+h, b+k) \in \Omega$) :

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

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

$\square$

Remarque (Interprétation du gradient)**.** Partons du DL de f en (a,b) :

$$f(x,y) = f(a,b) + (x-a)\frac{\partial f}{\partial x}(a,b) + (y-b)\frac{\partial f}{\partial y}(a,b) + o\left(||(x,y)-(a,b)||\right)$$

Sur le même principe que la preuve ci-dessus, en posant $A = (a,b)$ et $B = (x,y)$, le DL de f en A peut se réécrire

$$f(B) = f(A) + \langle \nabla f(A) \mid \overrightarrow{AB} \rangle + o\left(||\overrightarrow{AB}||\right)$$

Question. Supposons que B soit à distance $r \ll 1$ de A, donc que $||\overrightarrow{AB}|| = r \ll 1$. On a donc :

$$f(B) \approx f(A) + \langle \nabla f(A) \mid \overrightarrow{AB} \rangle$$

Quel est le point B qui donne la valeur $f(B)$ soit la plus grande possible ? la plus petite possible ? la plus proche de $f(A)$?

Ainsi, $\nabla f(A)$ **représente la direction où f croît le plus vite, localement autour de** A. En suivant la direction du gradient, on peut ainsi atteindre un maximum local de f (si f est majorée). Réciproquement, en suivant la direction inverse du gradient, on peut ainsi atteindre un minimum local de f (si f est minorée).

5 Extrema locaux

Définition 40.23

Soit $X \subset \mathbb{R}^2$ une partie non vide, et $f : X \to \mathbb{R}$. Soit $A = (a,b) \in X$

- On dit que f admet un <u>maximum</u> en A si : $\forall (x,y) \in X \qquad f(x,y) \leq f(a,b)$.
- On dit que f admet un <u>maximum local</u> en A si : $\exists r > 0 \quad \forall (x,y) \in X \cap B(A,r) \qquad f(x,y) \leq f(a,b)$
- On définit de même la notion de <u>minimum (local)</u> en A.
- On dit que f admet un <u>extremum (local)</u> en A si f admet un maximum (local) en A ou un minimum (local) en A.

Dit autrement, f admet un maximum en A si elle est majorée par $f(A)$, et admet un maximum local en A si elle est majorée par $f(A)$ au voisinage de A.

Définition 40.24

Soit $f \in \mathcal{C}^1(\Omega, \mathbb{R})$ et $A = (a,b) \in \Omega$. On dit que A est un <u>point critique</u> de f si $\nabla f(A) = 0$, càd si

$$\frac{\partial f}{\partial x}(a,b) = \frac{\partial f}{\partial y}(a,b) = 0$$

Propriété 40.25

Soit $f \in \mathcal{C}^1(\Omega, \mathbb{R})$ et $A \in \Omega$. Si f admet un extremum local en A, alors A est un point critique de f.

Méthode

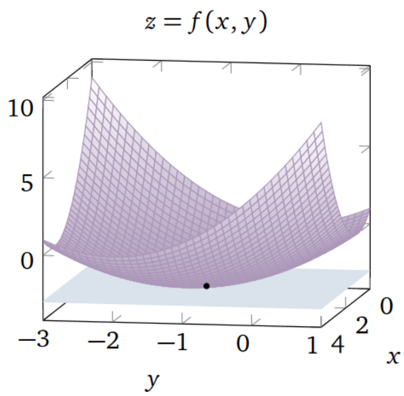
Pour trouver un extremum (local ou global) de $f \in \mathcal{C}^1(\Omega, \mathbb{R})$:

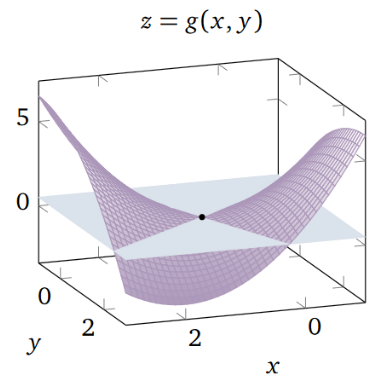
- On cherche les points critiques, qui sont des candidats pour être des extrema
- Pour chaque point critique (a,b), on étudie le signe de $f(x,y) - f(a,b)$, soit localement, soit globalement.

Exemple 16. Déterminer les extrema locaux de $f : (x,y) \mapsto x^2 - 3x + xy + y^2$. Pour chacun d'eux, dire s'il s'agit d'un minimum ou d'un maximum, local ou global.

La réciproque de la Propriété 40.25 est fausse ! Un point critique n'est pas toujours un extremum local.

Exemple 17. On pose $g : (x,y) \mapsto x^2 - x - xy - \dfrac{y^2}{2} + 2y$. Montrer que f admet un point critique, mais aucun extremum local sur $\mathbb{R}^2$.


$z = f(x, y)$
10
5
0
−3
−2
−1
0
1
4
2
0
y
x


$z = g(x, y)$
5
0
0
2
y
2
0
x