Chapitre 37

Fonctions de deux variables

Le but de ce chapitre est essentiellement pratique : poser les bases du calcul différentiel sur $\mathbb{R}^2$.Vous referez tout dans un cadre plus général l'an prochain, et donc nous passerons rapidement sur certains passages théoriques, mais bien comprendre ces exemples vous aidera tout d'abord à avoir une intuition géométrique solide, et vous aidera également à acquérir des réflexes de calcul, à commencer par la règle de la chaîne (cf. paragraphe IV.1) qui doit, elle, être absolument maîtrisée en première année, ainsi que le calcul des dérivées partielles.

On munit $\mathbb{R}^2$ et $\mathbb{R}^3$ de leurs structures euclidiennes canoniques, on identifiera le plan à $\mathbb{R}^2$ et l'espace à $\mathbb{R}^3$, et quand on parlera de représentations graphiques, il sera sous-entendu qu'on se place dans les repères $(O, \vec{i}, \vec{j})$ pour $\mathbb{R}^2$, ou $(O, \vec{i}, \vec{j}, \vec{k})$ pour $\mathbb{R}^3$, avec O le point $(0,0)$ ou $(0,0,0)$, et $(\vec{i}, \vec{j})$ la base canonique de $\mathbb{R}^2$ et $(\vec{i}, \vec{j}, \vec{k})$ la base canonique de $\mathbb{R}^3$.

I Fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$

I.1 Rudiments de topologie sur $\mathbb{R}^2$

Définition. On note $\| \ \|$ la norme euclidienne canonique sur $\mathbb{R}^2$, c'est-à-dire :

$$\forall a = (x, y) \in \mathbb{R}^2, \|a\| = \sqrt{x^2 + y^2}$$

Remarque : Nous avons déjà rencontré cette norme au chapitre 34 : c'est la norme euclidienne associée au produit scalaire canonique. On a de plus vu qu'elle vérifie les propriétés suivantes :

- **(Positivité et séparation)** $\|a\| \geqslant 0$, avec égalité si et seulement si $a = 0$.
- **(Absolue homogénéité)** $\|\lambda a\| = |\lambda| \times \|a\|$. En particulier, si $a \neq 0$, alors $a/\|a\|$ est unitaire c'est-à-dire de norme 1.
- **(Inégalité triangulaire)** $\|a + b\| \leqslant \|a\| + \|b\|$. On a également les inégalités triangulaires « inversées », avec un $-$.

On l'appelle parfois $\| \ \|_2$. Cela aura plus de sens l'an prochain quand vous parlerez également de la norme $\| \ \|_1$ et de la norme $\| \ \|_\infty$.

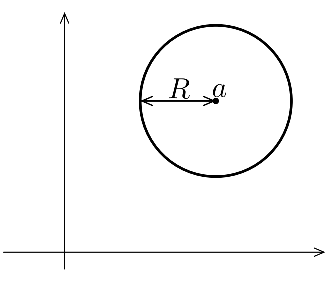
Remarques :

- Comme ci-dessus, nous noterons souvent les vecteurs a, b etc. pour noter les coordonnées x ou y, notations habituelles pour l'abscisse et l'ordonnée puisqu'on est sur $\mathbb{R}^2$.
- Comme dans le chapitre 34, la norme d'un vecteur représente sa distance à l'origine, et la norme d'une différence représente la différence entre les deux éléments.

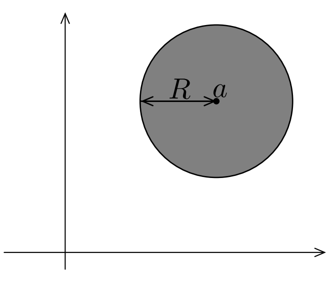
Définition. Soit $a \in \mathbb{C}$, soit $R \in \mathbb{R}_+^*$. On définit les ensembles suivants :

- $\{b \in \mathbb{R}^2 \mid \|a - b\| = R\}$ est la sphère de centre a de rayon R. On la note $S(a, R)$.
- $\{b \in \mathbb{R}^2 \mid \|a - b\| \leqslant R\}$ est la boule (fermée) de centre a de rayon R. On la note $B_F(a, R)$.
- $\{b \in \mathbb{R}^2 \mid \|a - b\| < R\}$ est la boule ouverte (c'est-à-dire privé de la sphère) de centre a de rayon R. On la note $B_O(a, R)$.

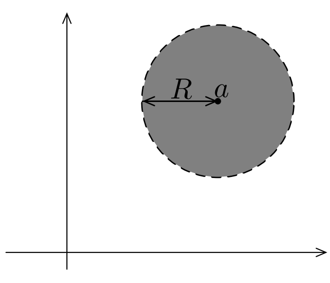
Interprétation géométrique :



Sphère



Boule fermée



Boule ouverte

Remarque : Pourquoi parle-t-on de sphère au lieu de cercle, de boule au lieu de disque (comme dans le chapitre 7) ? Parce que le but de ce chapitre est de se généraliser à $\mathbb{R}^n$ et même, pour certains résultats, à des espaces vectoriels quelconques, dans lesquels il est plus naturel de parler de boule (à commencer par $\mathbb{R}^3$). On a donc pris une même formulation pour tous les espaces, donc nous parlerons bien de boules et de sphère en dimension 2.

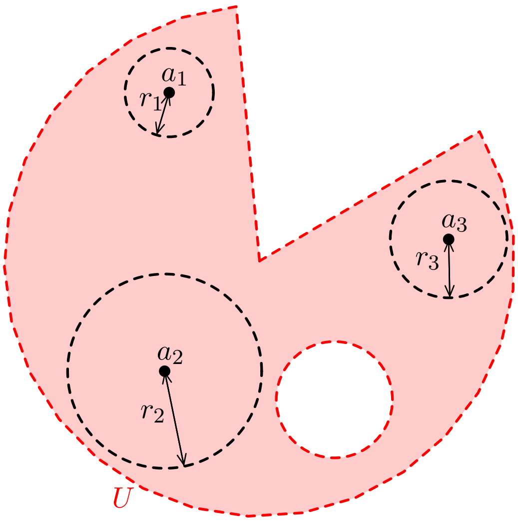
Définition. Soit U une partie de $\mathbb{R}^2$. On dit que U est un ouvert de $\mathbb{R}^2$ si :

$$\forall a \in U, \exists r > 0, B_O(a,r) \subset U$$

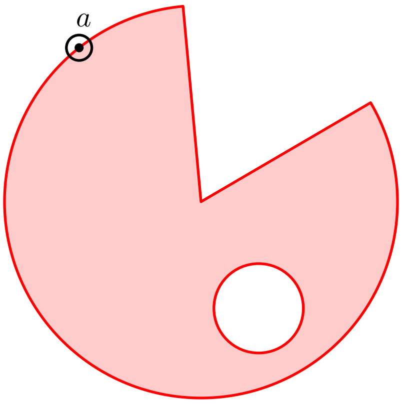
Remarques :

- En clair : un ensemble est un ouvert lorsque, pour tout point, il existe une petite boule pcentrée en ce point incluse dans l'ouvert, qui « ne déborde pas de l'ouvert », c'est-à-dire que pour tout point de l'ouvert, si on ne s'éloigne pas trop, on reste dans l'ouvert. Ci-dessous un ouvert (sans la frontière !) :

Un ouvert n'a aucune raison d'être convexe, ni même « connexe », c'est-à-dire en un seul morceau (ni même sans trou).



En effet, si on ajoute la frontière, ce n'est pas un ouvert car, si on prend un point a de la frontière, peu importe le rayon $r > 0$ qu'on prendra, la boule $B_O(a,r)$ ne sera pas incluse dans U :



- Intuitivement, pour tout point, il existe un voisinage de ce point inclus dans l'ouvert. Encore faut-il définir ce qu'est le voisinage d'un point... Ce n'est pas très difficile : une partie V est un voisinage de a s'il existe une boule centrée en a incluse dans V, c'est-à-dire s'il existe $r > 0$ tel que $B(a,r) \subset V$. On peut donc reformuler la définition d'un ouvert de la façon suivante : un ouvert est un ensemble qui est voisinage de tous ses points.

- Dans un ouvert, le rayon de la boule dépend du point considéré ! Certains points sont loin du bord et peuvent s'autoriser des boules de grand rayon, tandis que d'autres sont proches du bord et n'ont droit qu'à un petit rayon (mais toujours strictement positif).
- On pourrait prendre des boules fermées au lieu de boules ouvertes, la définition serait tout-à-fait équivalentes, c'est-à-dire que :

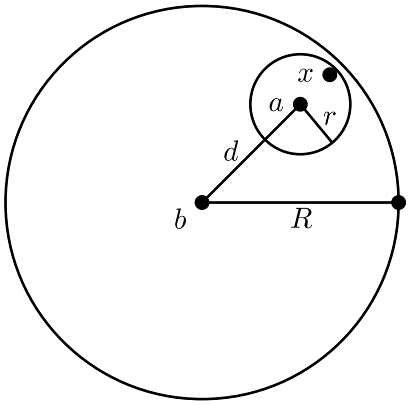
$$\forall a \in U, \exists r > 0, B_O(a,r) \subset U \iff \forall a \in U, \exists r > 0, B_F(a,r) \subset U$$

  En effet, s'il existe une boule fermée de rayon $r > 0$ incluse dans U alors la boule ouverte de rayon r est incluse dans U, et réciproquement, s'il existe une boule ouverte de rayon $r > 0$ dans U, alors la boule fermée de rayon $r/2 > 0$ est incluse dans U.

Attention, on ne dit pas qu'une boule ouverte est incluse si et seulement si la boule fermée de même rayon est aussi incluse dans l'ensemble ! On dit qu'il existe une boule ouverte dans l'ensemble si et seulement s'il existe une boule fermée (dont le rayon peut être différent) dans l'ensemble.

Activité : Montrons qu'une boule ouverte... est un ouvert.

Soit donc $B = B(b,R)$ une boule ouverte de centre $b \in \mathbb{R}^2$ et de rayon $R > 0$. Soit $a \in B(b,R)$. Montrons qu'il existe $r > 0$ tel que $B_O(a,r) \subset B$.



Notons $d = \|a-b\| < R$. On cherche donc r tel que $B_O(a,r) \subset B$. On pense à prendre $r = R - d$. Soit donc $r = R - d$ et soit $x \in B_O(a,r)$, et montrons que $x \in B$. Alors $\|x-a\| < r$. Par inégalité triangulaire,

$$\|x-b\| \leqslant \|x-a\| + \|a-b\| < r + d = R$$

si bien que $x \in B$ donc $B_O(a,r) \subset B$: B est un ouvert.

Même si les bords ne sont pas en pointillés, les boules sont ouvertes.

Vous ferez surtout ce genre de choses en deuxième année. Nous nous contenterons d'utiliser les ouverts pour y définir des fonctions de deux variables, et non pas pour les étudier eux précisément.

I.2 Fonctions définies sur un ouvert de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$

I.2.a Introduction

Dans ce chapitre, nous ne définirons que des fonctions définies sur une partie de $\mathbb{R}^2$ et à valeurs réelles. Vous manipulez des fonctions de ce type depuis des années :

- La fonction

$$a : \begin{cases} (\mathbb{R}_+^*)^2 \longrightarrow \mathbb{R} \\ (x,y) \longmapsto xy \end{cases}$$

  qui, à deux réels strictement positifs, renvoie leur produit : on peut voir $a(x,y)$ comme l'aire d'un rectangle dont la largeur et la longueur valent x et y.

- La fonction

$$p\colon \begin{cases} (\mathbb{R}_+^*)^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & 2x+2y \end{cases}$$

qui, à deux réels strictement positifs, renvoie la somme de leurs doubles : on peut voir $s(x,y)$ comme le périmètre du même rectangle.
- La fonction qui, à un point du plan (par exemple sur une carte IGN, cf. paragraphe IV.3) associe son altitude.

Dans notre chapitre (à part dans la dernière partie), nous ne nous intéresserons qu'à des fonctions dont le domaine de définition est un **ouvert** : en effet, nous allons parler de continuité et de dérivées (partielles) : il est donc indispensable de pouvoir définir des voisinages, et donc que le domaine de définition soit ouvert. C'est tout à fait analogue aux fonctions d'une seule variable : le domaine de définition ne pouvait pas contenir de points isolés.

On se donne dans la suite un ouvert U de $\mathbb{R}^2$.

I.2.b Rappel sur la représentation graphique d'une fonction d'une variable

Définition. Soit C un ensemble de points du plan. Soit (E) une équation à deux inconnues notées x et y (faisant éventuellement intervenir des paramètres notés $a, b\dots$). On dit que (E) est une équation de C si C est l'ensemble des points dont les coordonnées (x,y) sont solutions de (E). En d'autres termes :

$$M(x,y)\in C \iff x \text{ et } y \text{ sont solutions de } (E).$$

Encore en d'autres termes, un point appartient à un certain ensemble (une droite, un cercle etc.) si et seulement si ses coordonnées vérifient l'équation de cet ensemble.

Exemples :

- Une droite verticale a une équation du type $x=a$.
- Le cercle de centre $\Omega(x_0,y_0)$ et de rayon $R\geqslant 0$ a pour équation :

$$(x-x_0)^2+(y-y_0)^2=R^2$$

- Une droite non verticale a une équation du type $y=ax+b$.

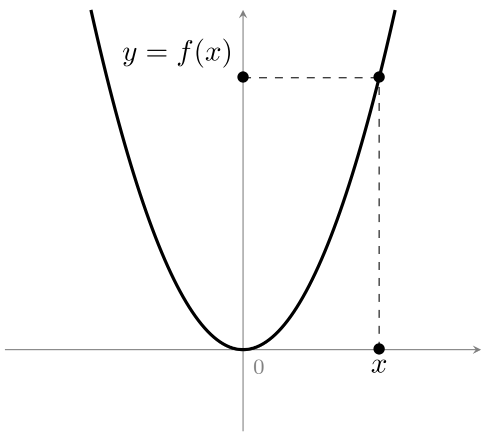
Remarque : ⚠ Attention a ne pas confondre l'ensemble C et l'équation (E). Par exemple, parler de la droite $y=x$ n'a aucun sens ! On parlera de la droite **d'équation** $y=x$.

Définition (courbe représentative). Soit $f : E \longrightarrow \mathbb{R}$. On appelle courbe représentative de f, et on note $\mathscr{C}_f$, l'ensemble $\mathscr{C}_f=\{(x,f(x))\,|\,x\in E\}$. En d'autres termes :

- $\mathscr{C}_f$ est l'ensemble des points du plan de coordonnées $(x,f(x))$, pour tout $x\in E$.
- $\mathscr{C}_f$ est l'ensemble des points du plan dont les coordonnées sont solution de l'équation $y=f(x)$.

En clair : $M(x,y)\in\mathscr{C}_f \iff y=f(x)$.

Remarque : En clair : $M(x,y)\in\mathscr{C}_f \iff y=f(x)$. $\mathscr{C}_f$ est aussi appelé le graphe de f.



I.2.c Représentation graphique d'une fonction de deux variables

On peut aisément généraliser ce qui précède à des fonctions de deux variables :

Définition. Soit S un ensemble de points de l'espace. Soit (E) une équation à trois inconnues notées (x, y, z) (faisant éventuellement intervenir des paramètres notés $a, b \dots$). On dit que (E) est une équation de S si S est l'ensemble des points dont les coordonnées (x, y, z) sont solutions de (E). En d'autres termes :

$$M(x, y, z) \in S \iff x, \ y \text{ et } z \text{ sont solutions de } (E).$$

Rappelons que x est l'abscisse, y l'ordonnée, et z la cote, c'est-à-dire « l'altitude ».

Exemples :

- Un plan parallèle au plan (Oxy) a pour équation $z = k$ (altitude constante : cf. paragraphe IV.3)
- La sphère de centre $\Omega(x_0, y_0, z_0)$ et de rayon $R \geqslant 0$ a pour équation :

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = R^2$$

Cet exemple ne nous sera pas utile en pratique car on ne s'intéressera qu'à des surfaces de la forme $z = f(x, y)$, voir ci-dessous.

Définition (surface représentative). Soit $f : U \longrightarrow \mathbb{R}$. On appelle surface représentative de f, et on note $\mathscr{S}_f$, l'ensemble $\mathscr{S}_f = \{(x, y, f(x, y)) \mid (x, y) \in U\}$. En d'autres termes :

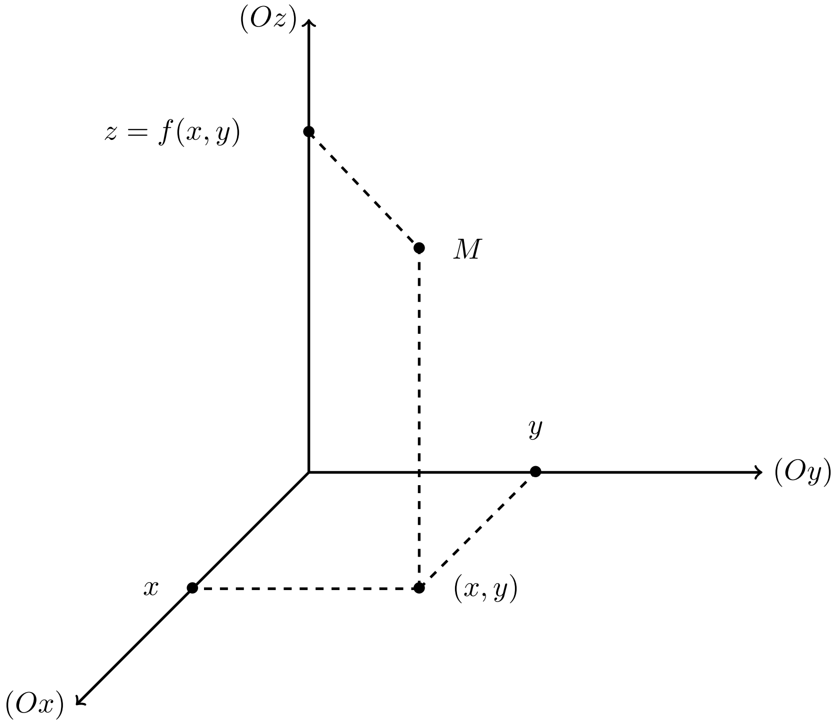
- $\mathscr{S}_f$ est l'ensemble des points du plan de coordonnées $(x, y, f(x, y))$, pour tout $(x, y) \in U$.
- $\mathscr{S}_f$ est l'ensemble des points du plan dont les coordonnées sont solution de l'équation $z = f(x, y)$.

Rappelons que U est un ouvert de $\mathbb{R}^2$.

Remarque : En d'autres termes, $\mathscr{S}_f$ est l'ensemble des points de l'espace dont les coordonnées sont solution de l'équation $z = f(x, y)$. En clair : $M(x, y, z) \in \mathscr{S}_f \iff z = f(x, y)$. $\mathscr{S}_f$ est aussi appelé le graphe de f.

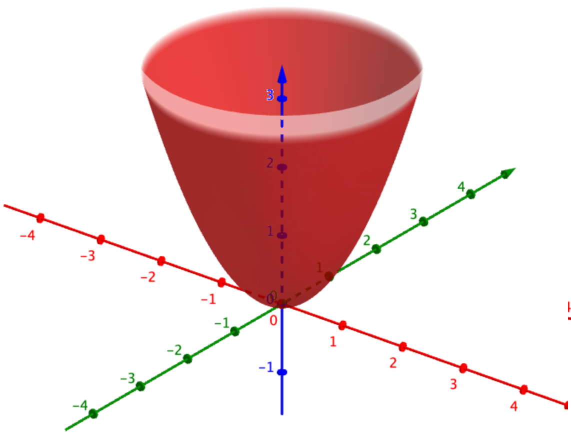
En clair : $M(x, y, z) \in \mathscr{S}_f \iff z = f(x, y, z)$.

De façon imagée : une surface, c'est une « nappe ».

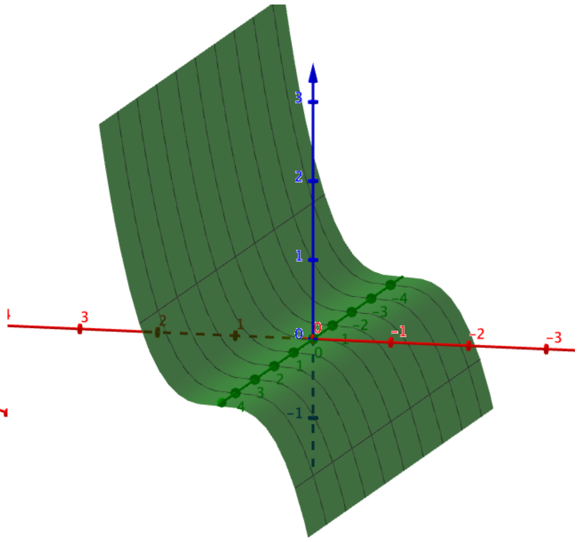


Exemples :

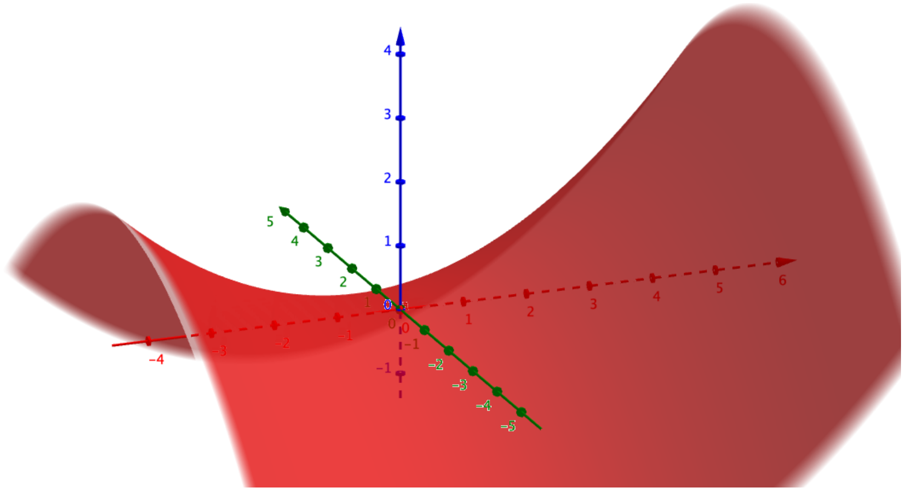
> On n'attend pas de vous que vous sachiez tracer une surface à main levée...



$z = x^2 + y^2$

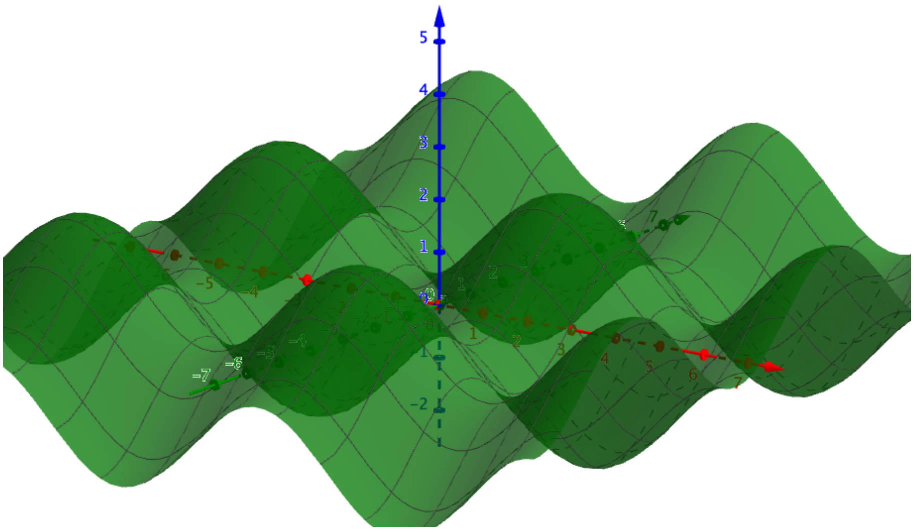


$z = x^3$

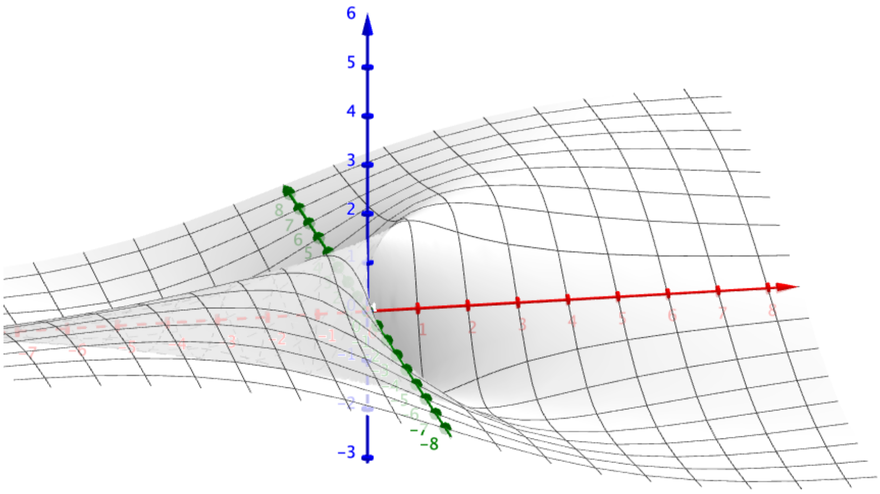


$z = x^2 - y^2$

> On a un effet « selle de cheval », cf. paragraphe V.



$z = \sin(x) + \sin(y)$



$$z = \frac{3xy}{x^2+y^2}$$

On remarque « qu'il y a un problème en $(0,0)$ » : cf. paragraphe II.2.

II Continuité d'une fonction de deux variables

On se donne dans la suite une fonction $f : U \to \mathbb{R}$ avec U un ouvert de $\mathbb{R}^2$.

II.1 Notion de limite

La définition est tout à fait analogue à celle pour des fonctions d'une variable : intuitivement, f va admettre une limite L en X_0 si $f(X)$ s'approche autant qu'on veut de L sans plus s'en éloigner donc si, peu importe la précision voulue, notée ε, l'écart entre $f(X)$ et L finit par être plus petit que ε dès que X est suffisamment proche de X_0. Cela justifie la définition suivante :

Nous ne parlerons que de limite réelles en un point de $\mathbb{R}^2$. On pourrait définie

Définition. Soit $L \in \mathbb{R}$ et soit $X_0 = (x_0, y_0) \in U$. On dit que f admet L pour limite en $X_0 = (x_0, y_0)$ ou tend vers L en X_0 si :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall X \in U, \|X - X_0\| \leqslant \eta \Rightarrow |f(X) - L| \leqslant \varepsilon$$

On note alors $f(X) \xrightarrow[X \to X_0]{} L$.

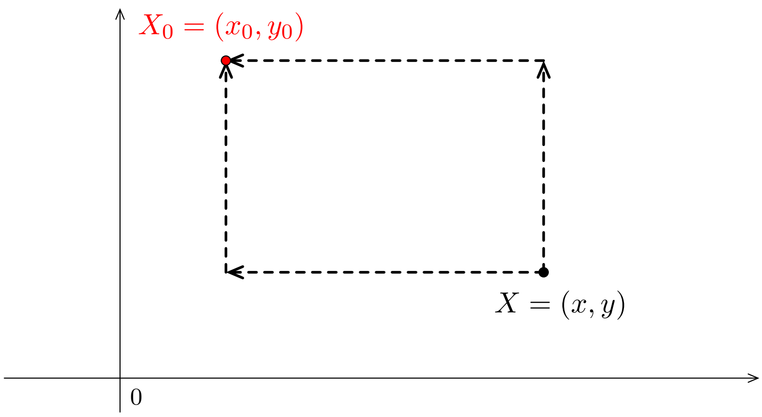
Remarques :

- La plupart des résultats vus au chapitre 13 (unicité de la limite, caractère localement borné, opérations sur les limites, théorème d'encadrement, l'inégalité large passe à la limite etc.) sont encore vrais pour des fonctions de deux variables : il suffit, dans toutes les preuves, de remplacer (pour les antécédents) la valeur absolue par la norme euclidienne de $\mathbb{R}^2$.
- En particulier, le théorème de composition de limites : si φ_1 et φ_2 vont de $\mathbb{R}$ dans U tendent vers (x_0, y_0) en t_0, et si $g : U \to \mathbb{R}$ tend vers L en (x_0, y_0), alors $t \mapsto g(\varphi_1(t), \varphi_2(t))$ tend vers L en (x_0, y_0).

Ce résultat est particulièrement utile pour prouver qu'une fonction n'admet PAS de limite, cf. paragraphe II.3.

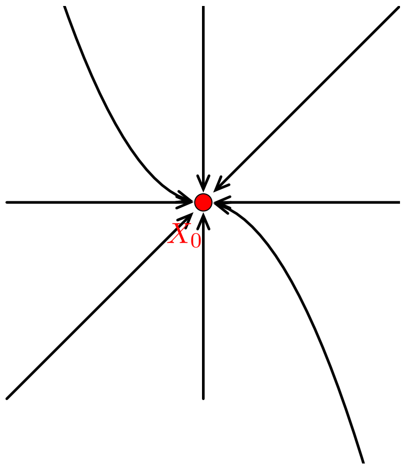
Exemple : En prenant $\eta = \varepsilon$, il est immédiat que $\|X - X_0\| \xrightarrow[X \to X_0]{} 0$. Ce résultat est souvent utile pour montrer la continuité de fonctions.

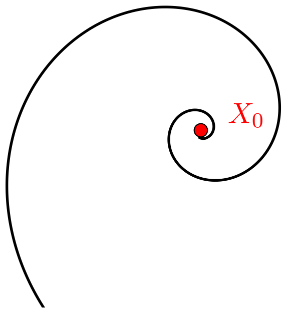
Remarque : Il n'y a pas de notion de limite à droite ou à gauche, car on peut arriver « de tous les côtés à la fois ». De plus, faire tendre x vers x_0 puis faire tendre y vers y_0 (ou le contraire) ne suffit pas, car alors cela revient à venir par la droite ou la gauche, ou par le haut ou le bas :



Or, dans le plan, la situation est beaucoup plus générale ! On n'est même pas obligé d'arriver de façon rectiligne, on peut arriver « de façon courbée » ou même en spirale !

Ce n'est pas parce que f admet la même limite en arrivant de toutes les directions que f tend effectivement vers cette limite : cf. exercice 6.





En fait, c'est le couple (x, y) qui doit tendre vers (x_0, y_0) « de façon globale », de toutes les façons possibles, mais pas forcément à la même vitesse... C'est un peu avec les mains, mais tout ça pour dire qu'il n'y a pas grand-chose d'autre que la définition.

II.2 Continuité

Définition. Soit $f : U \to \mathbb{R}$.

- Soit $X_0 = (x_0, y_0) \in U$. On dit que f est continue en $X_0 = (x_0, y_0)$ si $f(X) \xrightarrow[X \to X_0]{} f(X_0)$ c'est-à-dire :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall X \in U, \|X - X_0\| \leqslant \eta \Rightarrow |f(X) - f(X_0)| \leqslant \varepsilon$$

- f est continue sur U si f est continue en tout point de U.

Comme pour une fonction d'une seule variable, la continuité est une notion locale : une fonction est continue sur un ensemble lorsqu'elle est continue en chacun des points de l'ensemble.

Remarques :

- La norme ci-dessus est évidemment la norme euclidienne canonique i.e. $\|(x, y)\| = \sqrt{x^2 + y^2}$.
- À droite, c'est tout simplement la valeur absolue réelle puisque f est à valeurs dans $\mathbb{R}$.
- Cette définition est tout à fait analogue à celle sur $\mathbb{R}$: vous la généraliserez encore l'an prochain.
- Là aussi, on peut parler de prolongement par continuité : si f n'est pas définie en X_0 mais admet une limite finie L en X_0, alors on peut prolonger par continuité en X_0 en posant $f(X_0) = L$.

- Comme dans le paragraphe précédent, les règles usuelles (somme, produit, quotient lorsque le quotient n'est pas nul) sont encore valables, avec des preuves similaires (en remplaçant la valeur absolue pour les antécédents par la norme). De plus, on montre de même que si $g : U \to \mathbb{R}$ est continue en X_0 et si $f : \mathbb{R} \to \mathbb{R}$ est continue en $f(X_0)$, alors $g \circ f$ est continue en X_0.
- De même, si φ_1 et φ_2 vont de $\mathbb{R}$ dans U et sont continues en t_0, et si $g : U \to \mathbb{R}$ est continue en $(\varphi_1(t_0), \varphi_2(t_0))$, alors $t \mapsto g(\varphi_1(t), \varphi_2(t))$ est continue en t_0.

Ce résultat est particulièrement utile pour prouver qu'une fonction n'est PAS continue, cf. paragraphe II.3.

II.3 Exemples

Lemme. Les fonctions $p_1 : (x, y) \mapsto x$ et $p_2 : (x, y) \mapsto y$ sont continues sur $\mathbb{R}^2$.

Démonstration. Soit $X_0 = (x_0, y_0) \in \mathbb{R}^2$. Soit $\varepsilon > 0$. Soit $X = (x, y) \in \mathbb{R}^2$ (on fait comme dans le chapitre 13).

$$|p_1(x, y) - p_1(x_0, y_0)| \leqslant \varepsilon \quad \iff \quad |x - x_0| \leqslant \varepsilon$$

Or, par croissance de la racine carrée,

$$\|X - X_0\| = \sqrt{(x - x_0)^2 + (y - y_0)^2} \geqslant \sqrt{(x - x_0)^2} = |x - x_0| \qquad \square$$

Posons $\eta = \varepsilon$: si $\|X - X_0\| \leqslant \eta$ alors $p_1(X) - p_1(X_0)| \leqslant \varepsilon$: p_1 est continue en X_0 et X_0 est quelconque donc p_1 est continue sur $\mathbb{R}^2$. De même pour p_2.

Remarque : Par conséquent, par composition, les fonctions qui qui renvoient des fonctions continues en les coordonnées (et en particulier les fonctions polynomiales en les coordonnées) sont continues.

Exemple : La fonction $X \mapsto \|X\|$ est continue sur $\mathbb{R}^2$.

cf. paragraphe III.2 pour son graphe.

Notons f cette fonction. Alors $f : (x, y) \mapsto \sqrt{x^2 + y^2}$, c'est-à-dire que

$$f(x, y) = \sqrt{p_1(x, y)^2 + p_2(x, y)^2}$$

Or, p_1 est continue, la fonction carré est continue donc ${p_1}^2$ est continue, de même pour ${p_2}^2$. Par somme, ${p_1}^2 + {p_2}^2$ est continue et la racine carrée est continue donc f est continue.

Remarque : La continuité de f dépend fortement de la continuité des fonctions p_1 et p_2. Cependant, en pratique, il n'est pas nécessaire de justifier autant (ni même de citer ce lemme). Puisque l'étude de la continuité n'est pas un objectif du programme, on pourra se contenter de dire « par composition de fonctions continues ».

Exemples :

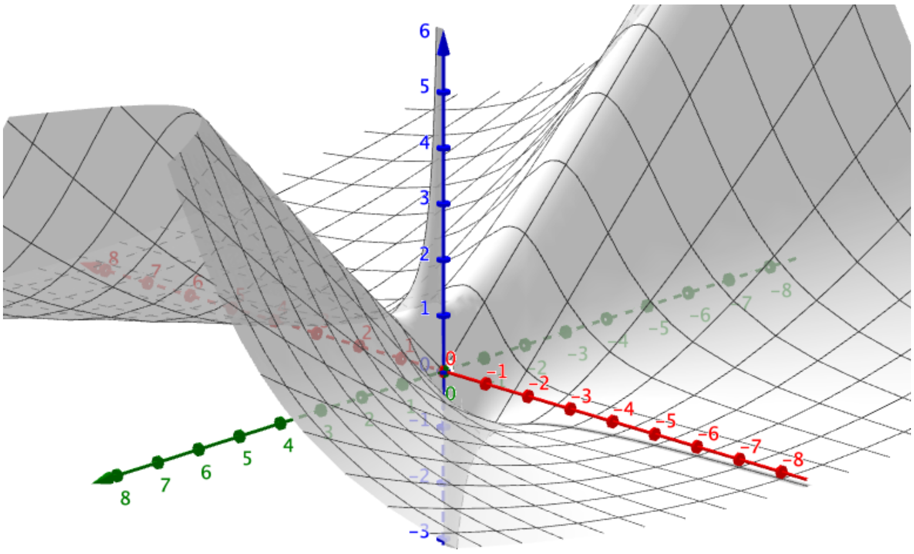
- La fonction $(x, y) \mapsto \sin(y + \cos(xy))$ est continue sur $\mathbb{R}^2$ car composée de fonctions continues.
- $f : (x, y) \mapsto \dfrac{x + y^2}{|y| + x^2}$ est continue sur $\mathbb{R}^2 \backslash \{(0, 0)\}$.

Remarque : On peut se demander si la question f ci-dessus est prolongeable par continuité en $(0, 0)$. On se demande donc si elle a une limite en $(0, 0)$.

- Pour montrer qu'une fonction n'a pas de limite en un point, on s'en tire en général avec le dernier résultat du paragraphe II.2 (qui n'est rien d'autre qu'une composition de limites). Par conséquent, « on pose $x = \varphi_1(t)$ et $y = \varphi_2(t)$ » (c'est-à-dire qu'on se restreint à une courbe du plan, on parle d'arc paramétré) avec φ_1 et φ_2 bien choisies, et on s'arrange pour ne pas avoir de limite en ce point, ou on le fait deux fois pour avoir deux limites différentes.
- Pour montrer qu'une fonction a une limite, pour la prolonger par continuité, on revient à la définition ou on peut faire un DL, passer en polaires etc.

Donnons quelques exemples mais rappelons que ce n'est pas un objectif du programme.
Exemples :

- Reprenons le dernier exemple. Pour tout $x \neq 0$, $f(x,0) = 1/x$ n'a pas de limite en 0 donc f n'a pas de limite en $(0,0)$ (en effet, d'après le résultat du paragraphe II.1), si f a une limite L en $(0,0)$ alors $f(x,0) \xrightarrow[x\to 0]{} f(0,0) = L$. Cela se voit très bien sur le graphe ci-dessous : quand on vient par les x négatifs (les rouges négatifs), « on plonge », et quand on vient par les x positifs (les rouges positifs), « on s'envole ».

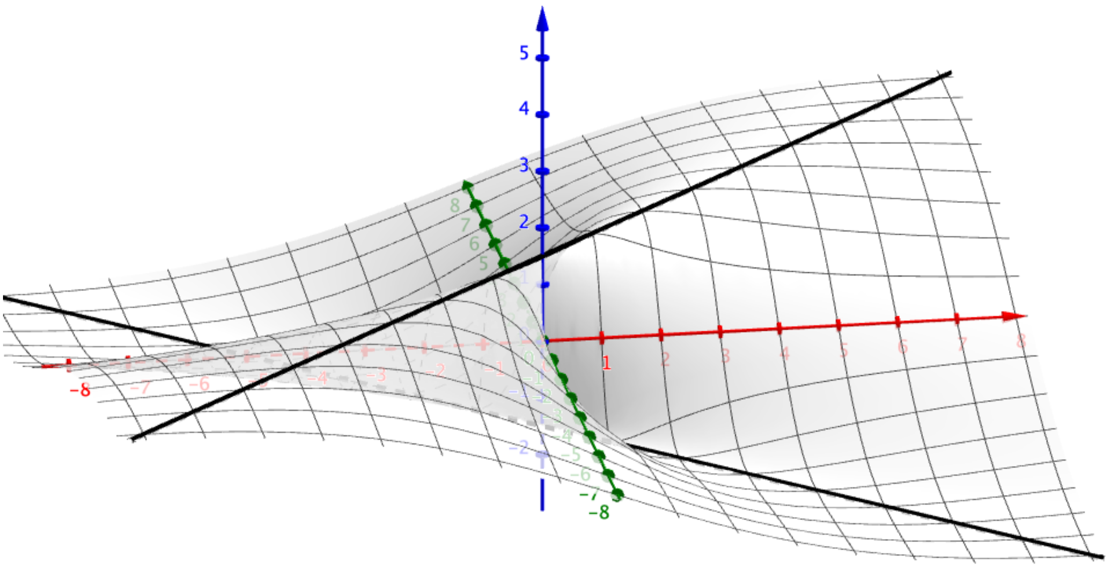


- Montrons que $f : (x,y) \mapsto \dfrac{x^2 - y^2}{x^2 + y^2}$ si $(x,y) \neq (0,0)$ n'a pas de limite en $(0,0)$.

  On remarque que, pour tout $x \neq 0$, $f(x,0) = 1 \xrightarrow[x\to 0]{} 1$ et $f(0,x) = -1 \xrightarrow[x\to 0]{} -1$ donc f n'a pas de limite en 0.
- Montrons que $f(x,y) \mapsto \dfrac{3xy}{x^2 + y^2}$ n'a pas de limite en $(0,0)$. D'une part, si $t \neq 0$,

  $$f(t,t) \quad = \quad 3/2 \xrightarrow[t\to 0]{} \frac{3}{2}$$

  mais, d'autre part, $f(t,-t) - \dfrac{3}{2} \xrightarrow[t\to 0]{} -\dfrac{3}{2}$. Les fonctions $t \mapsto f(t,t)$ et $t \mapsto f(t,-t)$ admettent des limites différentes donc f n'a pas de limite en $(0,0)$. Cela se voit bien sur le dessin suivant :



Quand on arrive en $(0,0)$ selon la droite d'équation $y=x$, alors les images tendent vers $3/2$, mais si on arrive selon la droite d'équation $y=-x$, les images tendent vers $-3/2$.

Remarque : Ici, le même exemple n'aurait pas marché : en effet, pour tout x non nuls, $f(x,0)=0 \xrightarrow[x\to 0]{} 0$ et $f(0,x)=0 \xrightarrow[x\to 0]{} 0$. On voit que faire tendre x vers 0 puis y vers 0 est insuffisant, cela revient (comme dit dans le paragraphe II.1) à arriver verticalement ou horizontalement, ce qui n'est pas suffisant pour affirmer que f est continue. Il faut parfois tenter des trucs. Il faut parfois avoir LA bonne idée : on tente $f(x,0), f(0,y), f(x,x), f(x,-x), f(x,2x), f(x,x^2)$ etc.

Si rien ne marche comme contre-exemple, alors on se dit que la fonction est continue : rien ne marche alors à part la définition : $f(x,y) \xrightarrow[(x,y)\to(x_0,y_0)]{} f(x_0,y_0)$, mais il est parfois plus facile de travailler en valeur absolue et avec des normes (car on travaille alors avec des réels et on peut parfois utiliser le théorème d'encadrement, et utiliser le fait que $\|(h,k)\| \xrightarrow[(h,k)\to(0,0)]{} 0$) et prouver que :

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

Les inégalités $|x| \leqslant \|(x,y)\|, |y| \leqslant \|(x,y)\|$ et $|xy| \leqslant \dfrac{x^2+y^2}{2} = \dfrac{\|(x,y)\|^2}{2}$ sont souvent utiles.

Exemple : Soit $f : (x,y) \mapsto \dfrac{xy^2}{x^2+y^2}$ si $(x,y)\neq(0,0)$, et $f(0,0)=0$. Alors f est continue sur $\mathbb{R}^2\backslash\{(0,0)\}$ car composée et quotients de fonctions continues Montrons que f est continue en 0. Soit $(x,y)\in\mathbb{R}^2$.

$$|f(x,y)| = \frac{|x|\times y^2}{\|(x,y)^2\|} \leqslant \frac{\|(x,y)\|^3}{\|(x,y)\|^2} = \|(x,y)\|$$

Or, $\|(x,y)\| \xrightarrow[(x,y)\to(0,0)]{} 0$ donc, d'après le théorème d'encadrement, on en déduit que $f(x,y) \xrightarrow[(x,y)\to(0,0)]{} 0 = f(0,0)$ donc f est continue sur $\mathbb{R}^2$.

Remarque : Si utiliser des inégalités ne permet pas de conclure, on peut parfois revenir à la définition (des ε) ou utiliser des outils d'analyse genre TAF : il y a beaucoup de méthodes différentes, mais nous n'allons voir qu'un exemple, l'étude de la continuité n'étant pas un objectif du programme.

Vous en referez en deuxième année.

Exemple : Soit g de classe $\mathscr{C}^1$ sur $\mathbb{R}$. Soit Δ la droite d'équation $y=x$. On définit f sur $\mathbb{R}^2$ par :

$$f : (x,y) \mapsto \begin{cases} \dfrac{g(x)-g(y)}{x-y} & \text{si } x\neq y \\ \\ g'(x) & \text{si } x=y \end{cases}$$

Montrons que g est continue sur $\mathbb{R}^2$. g est continue sur $\mathbb{R}^2\backslash\Delta$. Soit $(x_0,x_0)\in\Delta$ et prouvons que f est continue en (x_0,x_0). Soit $(x,y)\in\mathbb{R}^2$. Si $x=y$ alors $g(x,y)=f'(x)$, et sinon, d'après le TAF, il existe $c_{x,y}\in\,]\,x\,;y\,[$ (on explicite la dépendance en (x,y)) tel que

$$g(x,y) = \frac{f(x)-f(y)}{x-y} = f'(c_{x,y})$$

Dans tous les cas, il existe $c_{x,y}\in[\,x\,;y\,]$ tel que $g(x,y)=f'(c_{x,y})$. Or, $c_{x,y} \xrightarrow[(x,y)\to(x_0,y_0)]{} x_0$ et f' est continue car f est $\mathscr{C}^1$ si bien que $g(x,y) \xrightarrow[(x,y)\to(x_0,y_0)]{} f'(x_0) = g(x_0,x_0)$: g est bien continue sur $\mathbb{R}^2$.

III Dérivées partielles

III.1 Dérivées partielles

Définition. Soit $X_0 = (x_0, y_0) \in U$.

- On dit que f admet une dérivée partielle par rapport à x en X_0 si la fonction $f_1 : x \mapsto f(x, y_0)$ est dérivable en x_0, et on note alors :

$$\frac{\partial f}{\partial x}(x_0, y_0) = {f_1}'(x_0)$$

- On dit que f admet une dérivée partielle par rapport à y en X_0 si la fonction $f_2 : y \mapsto f(x_0, y)$ est dérivable en y_0, et on note alors :

$$\frac{\partial f}{\partial y}(x_0, y_0) = {f_2}'(x_0)$$

Interprétation géométrique : la dérivée par rapport à x est la dérivée de la fonction « qui va d'Est en Ouest » (i.e. parallèlement à l'axe des abscisses) et la dérivée par rapport à y est la dérivée de la fonction « qui va du Nord au Sud » (i.e. parallèlement à l'axe des ordonnées) : voir le dessin du point Pringle au paragraphe V.

Remarques :

- Sous réserve d'existence :

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

  et idem pour la dérivée partielle par rapport à y :

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

- On a vu la notation $\mathrm{d}/\mathrm{d}x$ pour la dérivée d'une fonction d'une variable. Le fait qu'on utilise ici un d rond signifie qu'il y a plusieurs variables.
- Parfois, surtout quand on manipulera des composées pour appliquer la règle de la chaîne (cf. paragraphe IV.1), il y aura beaucoup de variables, et ce sera assez difficile de s'y retrouver (par exemple, on aura une fonction qui s'appellera x). Il faut impérativement comprendre que, quand on écrit

$$\frac{\partial f}{\partial x} \quad \text{et} \quad \frac{\partial f}{\partial y}$$

  alors il est en fait écrit :

$$\frac{\partial f}{\partial \text{ première variable}} \quad \text{et} \quad \frac{\partial f}{\partial \text{ deuxième variable}}$$

  En général, évidemment, on note x la première variable et y la deuxième, mais parfois il y aura confusion donc il faut garder en tête ce que cela signifie, et ne pas hésiter à les écrire

$$\frac{\partial f}{\partial u} \quad \text{et} \quad \frac{\partial f}{\partial v}$$

  pour que tout cela devienne plus clair. On trouve également (dans certains livres, ce n'est pas une notation au programme) ∂_1 pour $\partial/\partial x$ et ∂_2 pour $\partial/\partial y$.
- L'existence de dérivées partielles n'est pas l'analogue de la dérivabilité pour les fonctions de deux variables ! En particulier, une fonction peut admettre des dérivées partielles en un point... sans être continue en ce point ! L'existence des dérivées partielles n'entraîne pas la continuité ! Si on note par exemple

L'analogue de la dérivée pour les fonctions de deux variables est la différentielle, au programme de deuxième année.

$$f : \begin{cases} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \begin{cases} 1 & \text{si } x = 0 \text{ ou } y = 0 \\ 0 & \text{si } x \neq 0 \text{ et } y \neq 0 \end{cases} \end{cases}$$

c'est-à-dire que f est la fonction qui vaut 1 sur les deux axes et 0 ailleurs, alors f n'est pas continue en $(0,0)$ (car, par exemple, , pour tout $x > 0$, $f(x,x) = 0 \xrightarrow[x \to 0^+]{} 0 \neq f(0,0) = 1$). Pourtant, pour tout x, $x \mapsto f(x,0) = 0$ est dérivable de dérivée nulle en 0, donc f admet une dérivée partielle par rapport à x en $(0,0)$ et $\dfrac{\partial f}{\partial x}(0,0) = 0$, et de même (en étudiant $y \mapsto f(0,y)$), f admet une dérivée partielle nulle par rapport à y en $(0,0)$.

- Les règles de calcul (dérivée partielle d'un produit, etc.) sont les mêmes que pour une dérivée classique... puisqu'une dérivée partielle est la dérivée d'une fonction d'une variable ! Plus précisément, le fait de se restreindre à une droite sur laquelle y est fixé permet de calculer la dérivée partielle par rapport à x comme la dérivée d'une fonction classique, en considérant que y est constante. De même, la dérivée partielle par rapport à y se calcule en dérivant par rapport à la variable y comme une fonction usuelle, en considérant que x est une constante.

Exemples :

- Soit $f : (x,y) \mapsto e^{x\cos(xy)}$. Fixons $y \in \mathbb{R}$. Alors $x \mapsto e^{x\cos(xy)}$ est dérivable, et si on fixe x, alors $y \mapsto f(x,y)$ est aussi dérivable. f admet donc des dérivées partielles en tout point, et pour tout $(x,y) \in \mathbb{R}^2$,

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

III.2 Fonctions de classe $\mathscr{C}^1$

Définition. On dit que f est de classe $\mathscr{C}^1$ sur U si f admet des dérivées partielles en tout point de U et si celles-ci sont des fonctions continues sur U. L'ensemble des fonctions $\mathscr{C}^1$ sur U à valeurs dans $\mathbb{R}$ est noté $\mathscr{C}^1(U,\mathbb{R})$.

Exemple : On a vu que, pour tout $(x,y) \in \mathbb{R}^2$, $f : (x,y) \mapsto e^{x\cos(xy)}$ admet des dérivées partielles en tout point, et les fonctions

$$(x,y) \mapsto \frac{\partial f}{\partial x}(x,y) \qquad \text{et} \qquad (x,y) \mapsto \frac{\partial f}{\partial y}(x,y)$$

sont continues donc f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$.

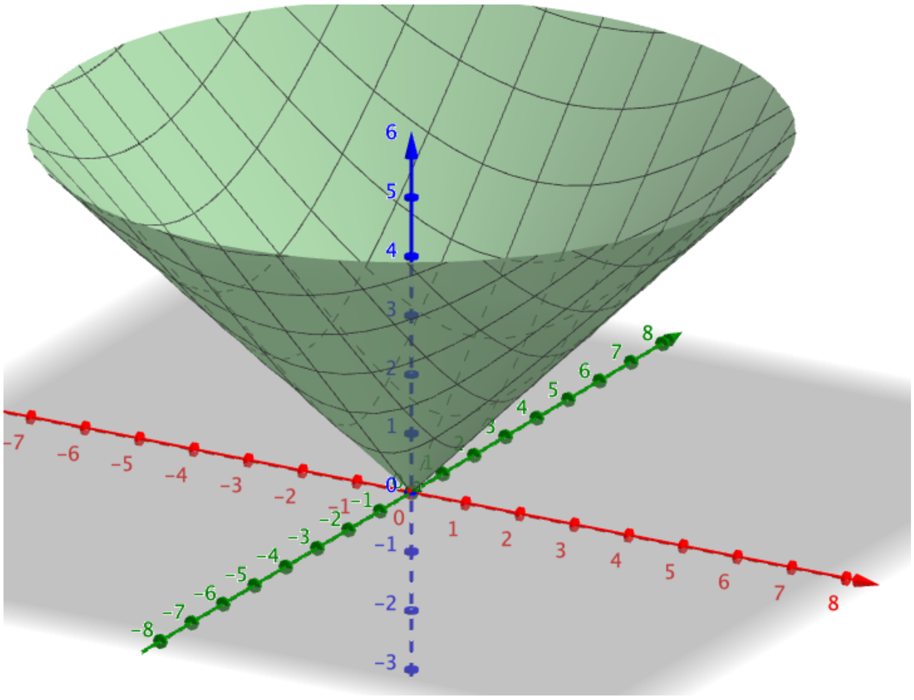
Remarque : Les propriétés de dérivabilité (somme de fonctions dérivables etc.), qui sont encore vraies pour les fonctions admettant des dérivées partielles, et les propriétés pour les fonctions continues vues au paragraphe II.2, permettent d'affirmer qu'une somme, un produit etc. de fonctions $\mathscr{C}^1$ est encore une fonction $\mathscr{C}^1$.

Lemme. Les fonctions $p_1 : (x,y) \mapsto x$ et $p_2 : (x,y) \mapsto y$ sont $\mathscr{C}^1$ sur $\mathbb{R}^2$.

Démonstration. Pour tout $y \in \mathbb{R}$, la fonction $x \mapsto p_1(x,y)$ n'est rien d'autre que la fonction $x \mapsto x$ qui est évidemment dérivable, donc p_1 admet une dérivée partielle égale à 1 en tout point $(x,y) \in \mathbb{R}^2$ c'est-à-dire que $\partial p_1/\partial x$ est la fonction constante égale à 1 sur $\mathbb{R}^2$ qui est continue. De même, p_2 admet une dérivée partielle en tout point par rapport à y, et celle-ci est la fonction constante égale à 0, également continue, donc p_1 est $\mathscr{C}^1$. De même pour p_2.

Remarque : Par conséquent, par composition, les fonctions qui renvoient des fonctions $\mathscr{C}^1$ en les coordonnées sont de classe $\mathscr{C}^1$. Comme ci-dessus pour la continuité, il ne sera pas nécessaire d'exhiber les fonctions p_1 et p_2 (même si, encore une fois, tout est fondé sur le caractère $\mathscr{C}^1$ de p_1 et p_2) et on pourra aller plus vite.

Exemple : La fonction $f : X \mapsto \|X\|$ est $\mathscr{C}^1$ sur $\mathbb{R}^2\backslash\{(0,0)\}$. En effet, f est la fonction $f : (x,y) \mapsto \sqrt{x^2+y^2}$. Or, $(x,y) \mapsto x^2+y^2$ est $\mathscr{C}^1$ sur $\mathbb{R}^2\backslash\{(0,0)\}$ et la racine carrée est $\mathscr{C}^1$ sur $\mathbb{R}_+^*$ donc f est $\mathscr{C}^1$ sur $\mathbb{R}^2\backslash\{(0,0)\}$. Et en $(0,0)$? Fixons $y=0$. Alors la fonction $x \mapsto \sqrt{x^2+0^2} = |x|$ n'est pas dérivable donc f n'a pas de dérivée partielle selon x en 0 : la norme n'est pas $\mathscr{C}^1$ sur $\mathbb{R}^2$: il y a une « pointe » en 0 !



III.3 Développement limité d'une fonction de deux variables

Seul le DL à l'ordre 1 est au programme :

Proposition/Définition (admise). Soit $f : U \to \mathbb{R}$ de classe $\mathscr{C}^1$. Alors, pour tout $(x_0,y_0) \in U$:

$$f(x_0+h, y_0+k) \underset{(h,k)\to(0,0)}{=} f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0) \times h + \frac{\partial f}{\partial y}(x_0,y_0) \times k + \mathrm{o}\left(\|(h,k)\|\right)$$

On dit que f admet un DL à l'ordre 1 au voisinage de (x_0,y_0).

Méthode : En pratique, le DL de f à l'ordre 1 s'obtient en se ramenant à des DL de fonctions d'une seule variable, x, y ou une autre variable en x et y (par exemple $x+y$: en effet, si $(x,y) \to (0,0)$, alors $x+y \to 0$). Il faut ensuite bien gérer le o, en remarquant qu'avec les notations précédentes, $\|(h,k)\| = \sqrt{h^2+k^2}$ et donc, par croissance de la racine carrée, $\|(h,k)\| \geqslant \max(|h|,|k|)$. En particulier : **un $\mathbf{o}(h)$ ou un $\mathbf{o}(k)$ est un $\mathbf{o}((h,k))$**, mais aussi, par exemple, un $\mathrm{o}(h+k)$ (car $|h+k| \leqslant 2\max(|h|,|k|) \leqslant 2\|(h,k)\|$), un terme en hk, car $|hk| \leqslant \|(h,k)\|^2 = \mathrm{o}(\|(h,k)\|)$, ou plus généralement un produit d'au moins deux termes en h et k (h^2, hk, h^2k) etc.

Exemples :

- Donnons le DL à l'ordre 1 en $(0,0)$ de $f : (x,y) \mapsto e^{\cos(x+2y)}(1+x+xy)$.

  f est $\mathscr{C}^1$ car produit et composée de fonctions $\mathscr{C}^1$. Soit $(x,y) \in \mathbb{R}^2$.

$$\begin{array}{rcl} f(x,y) & = & e^{1+\mathrm{o}(x+2y)} \times (1+x+xy) \\ & = & e \times e^{\mathrm{o}(\|(x,y)\|)} \times (1+x+\mathrm{o}(\|(x,y)\|)) \\ & = & e \times (1+\mathrm{o}(\|(x,y)\|)) \times (1+x+\mathrm{o}(\|(x,y)\|)) \\ & = & e + ex + \mathrm{o}(\|(x,y)\|) \end{array}$$

DL du cos : les termes suivants sont d'ordre 2.

car $xy = \mathrm{o}((x,y))$.

- Donnons le DL à l'ordre 1 en $(0,0)$ de $f : (x,y) \mapsto e^{x+\ln(2+y)}$.

  f est $\mathscr{C}^1$ par théorèmes généraux. Soit $(x,y) \in \mathbb{R}^2$.

$$\begin{array}{rcl} f(x,y) & = & e^{x+\ln(2)+\ln(1+y/2)} \\ & = & 2 \times e^{x+\ln(1+y/2)} \\ & = & 2 \times e^{x+y/2+\mathrm{o}(y)} \\ & = & 2 \times e^{x+y/2+\mathrm{o}(\|(x,y)\|)} \\ & = & 2 \times (1+x+y/2+\mathrm{o}(\|(x,y)\|)\quad) \\ & = & 2+2x+y+\mathrm{o}(\|(x,y)\|) \end{array}$$

(Les termes suivants sont en x^2, xy etc. donc $\mathrm{o}(\|(x,y)\|)$.

- Le DL de $f : (x,y) \mapsto x^3$ à l'ordre 1 en $(0,0)$ est $f(x,y) = \mathrm{o}(\|(x,y)\|)$.

Remarques :

- La définition est analogue à celle d'un DL d'une fonction d'une variable : on fait une approximation polynomiale de f au voisinage de (x_0, y_0), et puisqu'on fait un DL à l'ordre 1, on ne garde que les termes d'ordre 1, les autres termes étant négligeables. On peut encore sommer des DL, prouver qu'il y a unicité etc. mais il suffit de savoir effectuer un DL.
- Rappelons que le DL d'une fonction d'une seule variable en 0 est $f(x) = f(0) + xf'(0) + \mathrm{o}(x)$ et que son DL en x_0 est $f(x_0+h) = f(x_0) + hf'(x_0) + \mathrm{o}(h)$. La définition est donc tout à fait analogue !
- Remarquons que

$$(h,k) \mapsto \frac{\partial f}{\partial x}(x_0,y_0) \times h + \frac{\partial f}{\partial y}(x_0,y_0) \times k$$

  est une application linéaire, et donc

$$(h,k) \mapsto f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0) \times h + \frac{\partial f}{\partial y}(x_0,y_0) \times k$$

  est ce qu'on appelle une application affine. Là aussi, la situation est tout à fait analogue à ce qui se passe pour une fonction d'une variable : la partie principale du DL à l'ordre 1 est la meilleure approximation par une fonction affine, et c'est exactement ce qui se passe avec deux variables.
- Géométriquement, cela signifie que, au voisinage de (x_0, y_0), le graphe de f « ressemble » au plan (affine) d'équation

$$z = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0) \times (x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0) \times (y-y_0)$$

  Par analogie avec les fonctions d'une variable :

Définition. Soit $(x_0, y_0) \in U$. Si f est $\mathscr{C}^1$, on appelle plan tangent à la surface d'équation $z = f(x,y)$ en (x_0, y_0) le plan d'équation

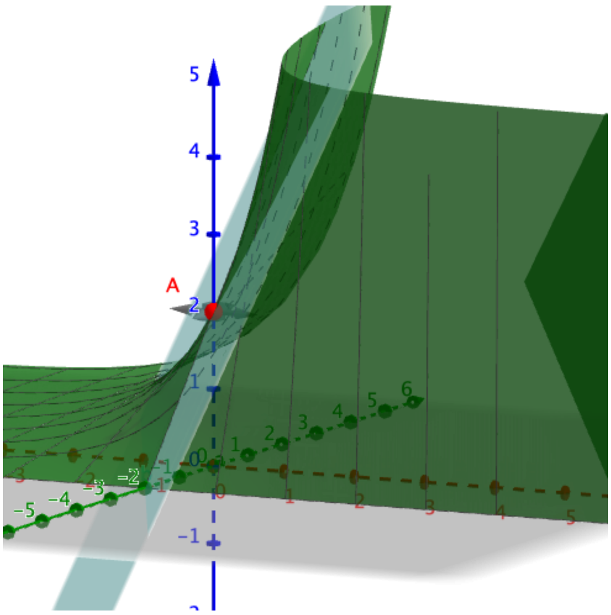
$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \times (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \times (y - y_0)$$

Remarque : Le fait (évoqué plus haut, et admis mais la démonstration n'est pas très difficile) qu'il y ait unicité du DL lorsqu'il y a existence permet de donner l'équation du plan tangent lorsqu'on a le DL à l'ordre 1 : si f est $\mathscr{C}^1$ et si f admet un DL à l'ordre 1 donné par $f(x_0 + h, y_0 + k) = a + bh + ck + \text{o}\left(\|(h,k)\|\right)$, alors $f(x_0, y_0) = a$, $\dfrac{\partial f}{\partial x}(x_0, y_0) = b$ et $\dfrac{\partial f}{\partial y}(x_0, y_0) = c$ si bien que le plan tangent admet comme équation $z = a + b(x - x_0) + c(y - y_0)$.

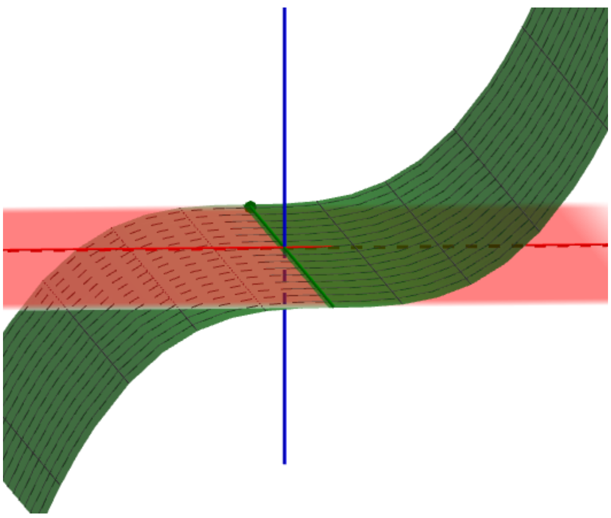
Le fait que, localement, le graphe de f ressemble au plan tangent se voit bien sur les exemples ci-dessous :

Attention, de même qu'une tangente n'est pas « la droite qui ne coupe le graphe qu'en un point », de même que la tangente peut traverser le graphe, le plan tangent peut recouper ailleurs la surface, et peut très bien la traverser, comme sur les exemples ci-contre :

- $f : (x,y) \mapsto e^{x + \ln(2+y)}$, dont le plan tangent en $(0,0)$ est le plan d'équation $z = 2 + 2x + y$:



- $f : (x,y) \mapsto x^3$, dont le plan tangent en $(0,0)$ est le plan d'équation $z = 0$:



Remarque : On a vu qu'une fonction admettant des dérivées partielles n'est pas forcément continue. Cependant, une fonction $\mathscr{C}^1$ est continue. En effet, à l'aide de l'expression du DL de f en $X_0 = (x_0, y_0)$, on en déduit que $f(M) \xrightarrow[M \to X_0]{} f(X_0)$.

Le vrai analogue du « dérivable ⇒ continue » pour les fonctions de deux variables est le « différentiable ⇒ continue » que vous verrez l'année prochaine.

III.4 Gradient d'une fonction $\mathscr{C}^1$

Définition. On suppose que f est de classe $\mathscr{C}^1$. Le gradient de f est la fonction notée ∇f définie sur U par :

$$\forall (x_0, y_0) \in U, \nabla f(x_0, y_0) = \begin{pmatrix} \dfrac{\partial f}{\partial x}(x_0, y_0) \\ \\ \dfrac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix}$$

Remarques :

- Attention, le gradient de f en un point est un vecteur ! ∇f n'est pas une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ mais une fonction de $\mathbb{R}^2$ dans $\mathbb{R}^2$.
- L'opérateur ∇ se lit : « nabla ».
- Si on note $\overrightarrow{i}$ et $\overrightarrow{j}$ la base canonique de $\mathbb{R}^2$ alors, pour tout $(x_0, y_0) \in \mathbb{R}^2$:

$$\nabla f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)\overrightarrow{i} + \frac{\partial f}{\partial y}(x_0, y_0)\overrightarrow{j}$$

  De même, si on note cette base (e_1, e_2) :

$$\nabla f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)e_1 + \frac{\partial f}{\partial y}(x_0, y_0)e_2$$

- En travaillant coordonnée par coordonnée et en utilisant les règles de dérivation pour les fonctions d'une variable, on obtient immédiatement les règles de calcul suivantes (avec des notations et des hypothèses évidentes, que nous ne précisons pas pour ne pas alourdir la chose) :

Proposition.

- **(Linéarité du gradient)** $\nabla(\lambda f + \mu g) = \lambda \nabla(f) + \mu \nabla(g)$.
- $\nabla(f \times g) = f \times \nabla g + g \times \nabla f$.
- $\nabla \left(\dfrac{1}{f}\right) = \dfrac{-1}{f^2}\nabla f$.
- $\nabla(\varphi \circ f) = (\nabla f) \times \varphi' \circ f$.

Moyen mnémotechnique : se dire dans sa tête que le gradient est l'analogue de la dérivée pour les fonctions d'une variable.

Proposition. On suppose que f est de classe $\mathscr{C}^1$. Alors, pour tout $X_0 = (x_0, y_0) \in U$, le DL de f en (x_0, y_0) (à l'ordre 1) s'écrit :

$$f(X) \underset{X \to X_0}{=} f(X_0) + \langle X - X_0, \nabla f(X_0)\rangle + \mathrm{o}\left(\|X - X_0\|\right)$$

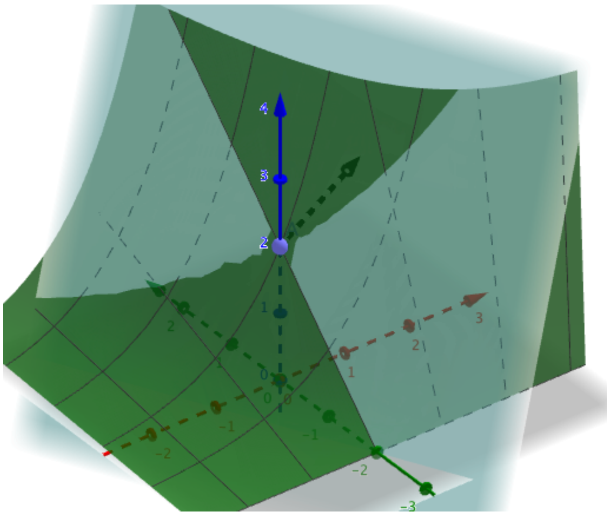
ou encore :

$$f(X_0 + H) \underset{H \to 0}{=} f(X_0) + \langle H, \nabla f(X_0)\rangle + \mathrm{o}\left(\|H\|\right)$$

où $\langle , \rangle$ est le produit scalaire canonique de $\mathbb{R}^2$.

Démonstration. Immédiat.

Remarque : Cette écriture permet de donner une interprétation géométrique du gradient. D'après l'inégalité de Cauchy-Schwarz, $\langle H, \nabla f(X_0)\rangle$ est maximal lorsque H et $\nabla f(X_0)$ sont colinéaires de même sens, et alors $f(X_0+H)$ est aussi maximal, et alors l'écart entre $f(X_0+H)$ et $f(X_0)$ est maximal, c'est donc dans la direction commune de H et de $\nabla f(X_0)$ que la croissance est la plus forte. En conclusion : le gradient indique la direction de la plus forte croissance de f au voisinage de X_0. Cela se comprend particulièrement bien lorsqu'on fait partir le vecteur $\nabla f(X_0)$ de X_0 : il est orienté dans la direction du plan tangent où la croissance est la plus forte (mais parallèlement au « plancher » i.e. au plan d'équation $z=0$, car c'est un vecteur du plan i.e. de $\mathbb{R}^2$).
Exemple : Avec $f:(x,y)\mapsto e^{x+\ln(2+y)}$, $\nabla(f)(0,0)=(2,1)$:



tandis qu'avec $f:(x,y)\mapsto x^3$, on a $\nabla f(0,0)=(0,0)$, ce qui est cohérent : toutes les directions sont identiques puisque le plan est horizontal. C'est le cas en tout point critique (cf. paragraphe V).

IV Dérivées partielles et composées

IV.1 Règle de la chaîne :

La maîtrise pratique du résultat qui suit est l'objectif principal de ce chapitre. Une fois n'est pas coutume, pour ne pas alourdir l'énoncé, donnons les hypothèses à part : on se donne I un intervalle de $\mathbb{R}$, non vide, non réduit à un point, x et y deux fonctions de I dans $\mathbb{R}$ telles que (x,y) soit à valeurs dans U.

Théorème (Régle de la chaîne). Si les fonctions considérées sont de classe $\mathscr{C}^1$, la fonction $t\mapsto f(x(t),y(t))$ est de classe $\mathscr{C}^1$ et :

$$\forall t\in I,\ \frac{\mathrm{d}}{\mathrm{d}t}\left(f(x(t),y(t))\right) = x'(t)\times\frac{\partial f}{\partial x}(x(t),y(t)) + y'(t)\times\frac{\partial f}{\partial y}(x(t),y(t))$$

Nous en donnerons une interprétation géométrique (difficile) dans le paragraphe IV.3.

Démonstration. Notons $g:t\mapsto f(x(t),y(t))$ et $\gamma:t\mapsto(x(t),y(t))$ (si bien que $g=f\circ\gamma$). Soient t et h deux réels. Tout d'abord :

$$\|\gamma(t+h)-\gamma(t)\| = \sqrt{(x(t+h)-x(t))^2+(y(t+h)-y(t))^2}$$

Or, x et y sont $\mathscr{C}^1$ donc on peut effectuer un DL au voisinage de t (quand $h\to 0$, donc) :

$$\begin{aligned}\|\gamma(t+h)-\gamma(t)\| &= \sqrt{(hx'(t)+\mathrm{o}(h))^2+(hx'(t)+\mathrm{o}(h))^2}\\ &= \sqrt{h^2x'(t)^2+h^2y'(t)^2+\mathrm{o}(h^2)}\\ &= |h|\sqrt{x'(t)^2+y'(t)^2+\mathrm{o}(1)}\\ &= \mathrm{O}(h)\end{aligned}$$

En particulier, on en déduit que $\gamma(t+h) \xrightarrow[h\to 0]{} \gamma(t)$, et qu'un $\mathrm{o}\left(\|\gamma(t+h)-\gamma(t)\|\right)$ est un $\mathrm{o}(h)$ (une quantité négligeable devant $\|\gamma(t+h)-\gamma(t)\|$ sera négligeable devant h). f étant elle aussi de classe $\mathscr{C}^1$, elle admet un DL à l'ordre 1 et donc

$$\begin{aligned}
g(t+h) &= f(\gamma(t+h)) \\
&= f(\gamma(t)) + \langle \gamma(t+h)-\gamma(t), \nabla f(\gamma(t))\rangle + \mathrm{o}\left(\|\gamma(t+h)-\gamma(t)\|\right) \\
&= g(t) + (x(t+h)-x(t)) \times \frac{\partial f}{\partial x}(x(t),y(t)) + (y(t+h)-y(t)) \times \frac{\partial f}{\partial y}(x(t),y(t)) + \mathrm{o}(h) \\
&= g(t) + (hx'(t)+\mathrm{o}(h)) \times \frac{\partial f}{\partial x}(x(t),y(t)) + (hy'(t)+\mathrm{o}(h)) \times \frac{\partial f}{\partial y}(x(t),y(t)) + \mathrm{o}(h) \\
&= g(t) + h\left(x'(t) \times \frac{\partial f}{\partial x}(x(t),y(t)) + y'(t) \times \frac{\partial f}{\partial y}(x(t),y(t))\right) + \mathrm{o}(h)
\end{aligned}$$

En d'autres termes, on a un DL de g à l'ordre 1. Or, rappelons que g va de I dans $\mathbb{R}$: l'existence du DL à l'ordre 1 est équivalente à la dérivabilité en t et le coefficient devant h est égal à $g'(t)$, donc g est dérivable, et

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

□

Enfin, par composition, g est continue donc g est $\mathscr{C}^1$: la règle de la chaîne est démontrée.

Corollaire. Avec les mêmes hypothèses, en posant encore une fois $\gamma : t \mapsto (x(t),y(t))$, il vient :

$$(f\circ\gamma)'(t) = \langle \nabla f(\gamma(t)), \gamma'(t)\rangle$$

C'est une simple réécriture du résultat, il suffit de l'écrire.

Remarques :

- Comme dit plus haut, il faut IMPÉRATIVEMENT maîtriser cette règle (c'est écrit explicitement dans le programme) et savoir l'appliquer sans se poser de question et sans faire d'erreur.
- On pourrait se mélanger les pinceaux entre le x de $\dfrac{\partial f}{\partial x}$ et le x de la fonction $t\mapsto x(t)$. Il est impératif de bien comprendre que $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont en fait, respectivement :

$$\frac{\partial f}{\partial \text{ première variable}} \qquad \text{et} \qquad \frac{\partial f}{\partial \text{ deuxième variable}}$$

  Si on a peur de se tromper, on peut nommer u et v les variables de f et les dérivées partielles $\dfrac{\partial f}{\partial u}$ et $\dfrac{\partial f}{\partial v}$.
- Comme pour la dérivée d'une composée de fonctions d'une variable, il suffit de retenir le principe suivant : on dérive ce qu'il y a à l'intérieur (la fonction dans la première variable et la fonction dans la deuxième variable), puis on dérive f par rapport à la première coordonnée et par rapport à la deuxième coordonnée, et on multiplie chaque dérivée par la dérivée partielle correspondante (et on somme). C'est la même chose que pour une variable : la dérivée de f est remplacée par son gradient (i.e. le vecteur des dérivées partielles) et le produit des deux dérivées est remplacé par le produit scalaire.

Exemples :

- Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathscr{C}^1$. Donner la dérivée de $F : t \mapsto f(t^2, \sin(t))$.

  D'après la règle de la chaîne, F est $\mathscr{C}^1$ et, pour tout $t \in \mathbb{R}$:

$$F'(t) = 2t\frac{\partial f}{\partial x}(t^2, \sin(t)) + \cos(t)\frac{\partial f}{\partial y}(t^2, \sin(t))$$

- Soit $f : (x, y) \mapsto xe^{xy^2}$ et soit $F : t \mapsto f(\ln(t), t^2)$ définie sur $\mathbb{R}_+^*$. Donner la dérivée de F.

  f est de classe $\mathscr{C}^1$ et la fonction ln et la fonction carré sont $\mathscr{C}^1$ donc, d'après la règle de la chaîne, F est $\mathscr{C}^1$. Soit $t > 0$.

$$F'(t) = \frac{1}{t} \times \frac{\partial f}{\partial x}(\ln(t), t^2) + 2t \times \frac{\partial f}{\partial y}(\ln(t), t^2)$$

  Or, pour tout $(x, y) \in \mathbb{R}^2$:

$$\frac{\partial f}{\partial x}(x, y) = (1 + xy^2)e^{xy^2} \qquad \text{et} \qquad \frac{\partial f}{\partial y}(x, y) = 2x^2ye^{xy^2}$$

  Finalement :

$$F'(t) = \frac{1}{t} \times \left(1 + t^4\ln(t)\right)e^{t^4\ln(t)} + 2t^3(\ln(t))^2e^{t^4\ln(t)}$$

La règle de la chaîne n'est pas indispensable ici : on pourrait écrire F explicitement et la dérivée « normalement ».

On peut utiliser cette formule pour donner la dérivée lorsqu'on compose des fonctions de deux variables. Encore une fois, nous donnons les hypothèses à part :

- f est de classe $\mathscr{C}^1$ sur un ouvert U de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$.
- φ et ψ sont deux fonctions de classe $\mathscr{C}^1$ sur un ouvert V de $\mathbb{R}^2$ à valeurs dans U (on a donc des fonctions de deux variables qui renvoient des éléments de $\mathbb{R}^2$).

Théorème. La fonction $(u, v) \mapsto f(\varphi(u, v), \psi(u, v))$ est $\mathscr{C}^1$ et, pour tout $(u, v) \in V$:

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

et

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

Démonstration. On fixe une variable et on applique la règle de la chaîne pour l'autre variable, ce qui donne la dérivée partielle correspondante.

Remarques :

- Ne vous embêtez pas trop avec les variables u, v, x, y etc. (surtout que les noms des variables risquent de changer dans les exercices, sans oublier le fait que les variables sont muettes) et remplacez-les dans votre tête par « première variable » et « deuxième variable », ou plutôt ∂_1 et ∂_2 pour gagner de la place :

$$\frac{\partial f(\varphi(u,v),\psi(u,v))}{\partial u} = \frac{\partial \varphi}{\partial u}(u,v) \times \underbrace{\partial_1 f(\varphi(u,v),\psi(u,v))}_{\text{on a dérivé ce qui est en emplacement 1}} + \frac{\partial \psi}{\partial u}(u,v) \times \underbrace{\partial_2 f(\varphi(u,v),\psi(u,v))}_{\text{on a dérivé ce qui est en emplacement 2}}$$

  et

$$\frac{\partial f(\varphi(u,v),\psi(u,v))}{\partial v} = \frac{\partial \varphi}{\partial v}(u,v) \times \underbrace{\partial_1 f(\varphi(u,v),\psi(u,v))}_{\text{on a dérivé ce qui est en emplacement 1}} + \frac{\partial \psi}{\partial v}(u,v) \times \underbrace{\partial_2 f(\varphi(u,v),\psi(u,v))}_{\text{on a dérivé ce qui est en emplacement 2}}$$

- Cette expression peut faire un peu peur, mais il faut bien comprendre que le principe est toujours le même : disons qu'on veuille dériver par rapport à u (la première variable), alors on dérive tout par rapport à u. On dérive d'abord ce qui se trouve dans la première variable de f, dans son premier emplacement, le $\partial\varphi/\partial u$, qu'on multiplie par rapport à la dérivée partielle de f correspondante, c'est-à-dire par rapport à sa première variable, $\partial f/\partial x$, et on recommence avec la deuxième variable, $\partial\psi/\partial u \times \partial f/\partial y$. En clair : on dérive tout par rapport à u, on commence par ce qu'il y a dans la première variable de f, en appliquant le théorème de dérivation d'une composée, et on recommence avec ce qu'il y a dans la deuxième variable, et idem si on veut dériver par rapport à v.

Exemple : Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathscr{C}^1$ et soit $F : (x,y) \mapsto f(x^2+y, xy-y)$. Alors F est $\mathscr{C}^1$ car f est $\mathscr{C}^1$ et $\varphi : (x,y) \mapsto x^2+y$ et $\psi : (x,y) \mapsto xy-y$ sont $\mathscr{C}^1$. Soit $(x,y) \in \mathbb{R}^2$.

$$\begin{aligned}\frac{\partial F}{\partial x}(x,y) &= \frac{\partial \varphi}{\partial x}(x,y) \times \frac{\partial f}{\partial x}(x^2+y, xy-y) + \frac{\partial \psi}{\partial x}(x,y) \times \frac{\partial f}{\partial y}(x^2+y, xy-y) \\ &= 2x \times \frac{\partial f}{\partial x}(x^2+y, xy-y) + y \times \frac{\partial f}{\partial y}(x^2+y, xy-y)\end{aligned}$$

et

$$\begin{aligned}\frac{\partial F}{\partial y}(x,y) &= \frac{\partial \varphi}{\partial y}(x,y) \times \frac{\partial f}{\partial x}(x^2+y, xy-y) + \frac{\partial \psi}{\partial y}(x,y) \times \frac{\partial f}{\partial y}(x^2+y, xy-y) \\ &= 1 \times \frac{\partial f}{\partial x}(x^2+y, xy-y) + (x-1) \times \frac{\partial f}{\partial y}(x^2+y, xy-y)\end{aligned}$$

IV.2 Dérivée directionnelle selon un vecteur

Définition. Soit $v \in \mathbb{R}^2$ et soit $X_0 \in U$. On dit que f admet une dérivée directionnelle selon le vecteur v si la fonction $F : t \mapsto f(X_0 + tv)$ est dérivable en 0 : $F'(0)$ est alors appelé dérivée directionnelle de f en X_0 selon le vecteur v et est noté $D_v f(X_0)$.

Il est bien évident que F n'a rien à voir avec une quelconque primitive de f.

Remarques :

- On dit aussi que f admet une dérivée selon le vecteur v ou une dérivée dans la direction de v.
- Si v est le vecteur nul, alors F est constante donc dérivable et sa dérivée est nulle : toute fonction admet une dérivée directionnelle (nulle) selon le vecteur nul.
- Là aussi, une dérivée directionnelle étant une dérivée d'une fonction d'une seule variable, les règles habituelles s'appliquent (en particulier la linéarité).
- En d'autres termes (dérivée d'une fonction d'une variable), f admet une dérivée directionnelle en X_0 selon le vecteur v si la fonction

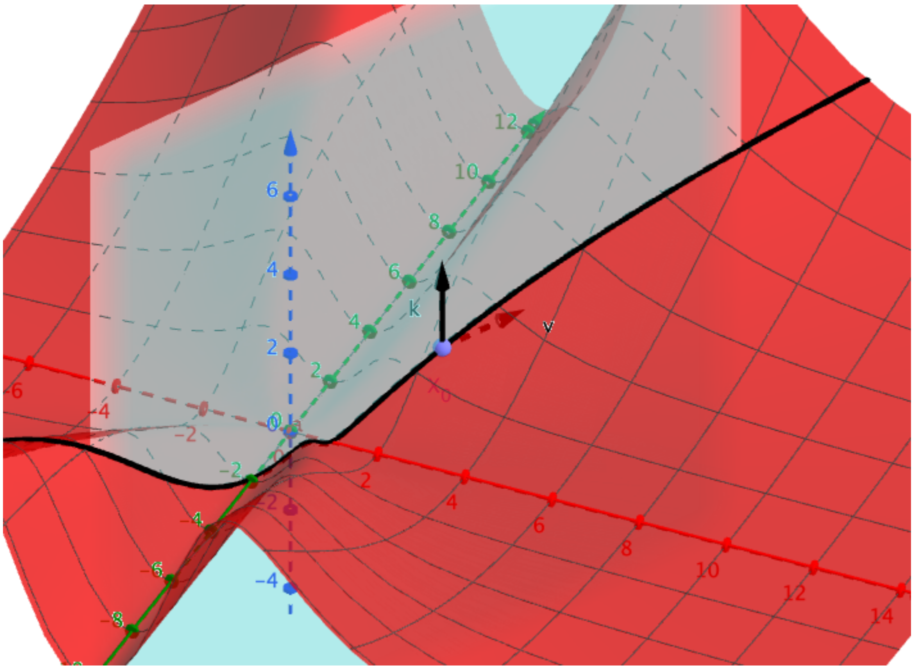
$$\tau : t \mapsto \frac{f(X_0+tv) - f(X_0)}{t}$$

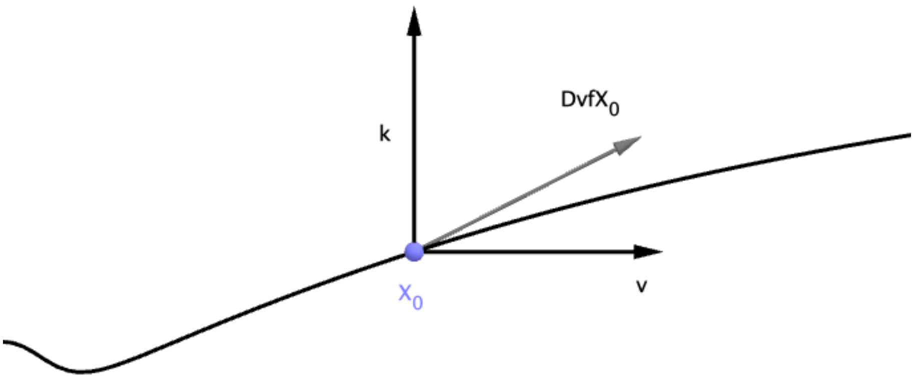
admet une limite finie en 0, et alors

$$D_v f(X_0) = \lim_{t \to 0} \frac{f(X_0+tv) - f(X_0)}{v}$$

- Sous réserve d'existence, les dérivées partielles $\dfrac{\partial f}{\partial x}(x_0, y_0)$ et $\dfrac{\partial f}{\partial y}(x_0, y_0)$ sont les dérivées directionnelles selon les vecteurs $\overrightarrow{i}$ et $\overrightarrow{j}$, c'est-à-dire les dérivées dans les directions respectivement de l'axe des abscisses et de l'axe des ordonnées.

- Interprétation géométrique (lorsque v n'est pas le vecteur nul) : on trace la surface représentative de f, on l'intersecte avec le plan vertical passant par A dirigé par v : l'idée est que v va représenter le vecteur $\overrightarrow{i}$ dans ce nouveau plan (i.e. le vecteur unitaire dirigeant l'axe des abscisses). On prend donc un vecteur $\overrightarrow{j}$ orthogonal à $\overrightarrow{i} = v$ de même norme (pour avoir un repère orthonormal) et comme origine le point $(X_0, f(X_0))$. L'intersection de la surface avec le plan donne une courbe, et $D_v f(X_0)$ est la dérivée en X_0 i.e. le coefficient directeur de la tangente en X_0 de cette courbe (en prenant v comme échelle de longueur) :





En particulier, les dérivées partielles sont les dérivées des courbes obtenues lorsqu'on intersecte avec un plan vertical parallèle à l'axe des abscisses (pour la dérivée partielle selon x) et l'axe des ordonnées (dérivée partielle selon y). Voir un dessin dans le paragraphe V.

On peut se demander quand f admet des dérivées directionnelles :

Proposition. On suppose que f est de classe $\mathscr{C}^1$. Alors f admet des dérivées directionnelles en tout point et dans toutes les directions. De plus, pour tous $(x_0, y_0) \in U$ et

$v = (v_1, v_2) \in \mathbb{R}^2$:

$$\begin{aligned} D_v f(x_0, y_0) &= \langle \nabla f(x_0, y_0), v \rangle \\ &= \frac{\partial f}{\partial x}(x_0, y_0) \times v_1 + \frac{\partial f}{\partial y}(x_0, y_0) \times v_2 \end{aligned}$$

DÉMONSTRATION. Soit $X_0 = (x_0, y_0) \in U$ et soit $v = (v_1, v_2) \in \mathbb{R}^2$. Pour tout t en lequel cela a un sens, $F(t) = f(x_0 + tv_1, y_0 + tv_2)$. D'après la règle de la chaîne, les fonctions f, $x : t \mapsto x_0 + tv_1$ et $y : t \mapsto y_0 + tv_1$ étant $\mathscr{C}^1$, F est $\mathscr{C}^1$ et en particulier dérivable, d'où l'existence de la dérivée directionnelle, et :

$$\begin{aligned} F'(0) &= x'(0)\frac{\partial f}{\partial x}(x(0), y(0)) + y'(0)\frac{\partial f}{\partial y}(x(0), y(0)) \\ &= v_1 \frac{\partial f}{\partial x}(x_0, y_0) + v_2 \frac{\partial f}{\partial y}(x_0, y_0) \end{aligned}$$

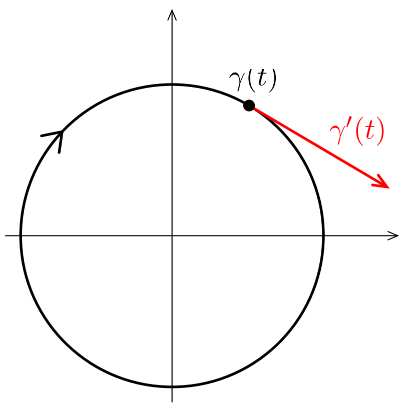
□

IV.3 Notion de courbe paramétrée

Si x et y sont deux fonctions définies sur une partie D de $\mathbb{R}$ à valeurs dans $\mathbb{R}$, alors $\gamma = (x, y)$ est une fonction de D dans $\mathbb{R}^2$. En d'autres termes, c'est une fonction qui à tout élément de D associe un point du plan, et donc, quand t varie, $\gamma(t) = (x(t), y(t))$ décrit un ensemble qu'on appelle une courbe paramétrée. Par exemple, $\gamma : t \mapsto (\cos(t), \sin(t))$ décrit, lorsque t décrit $\mathbb{R}$, le cercle trigonométrique. Si f est une fonction de D dans $\mathbb{R}$, alors la courbe décrite par $t \mapsto (t, f(t))$ est le graphe de f : une courbe paramétrée permet donc de donner des courbes plus générales que des graphes de fonctions (plusieurs points avec les mêmes abscisses par exemple).

cf. chapitre 5, on a parlé de paramétrisation du cercle unité.

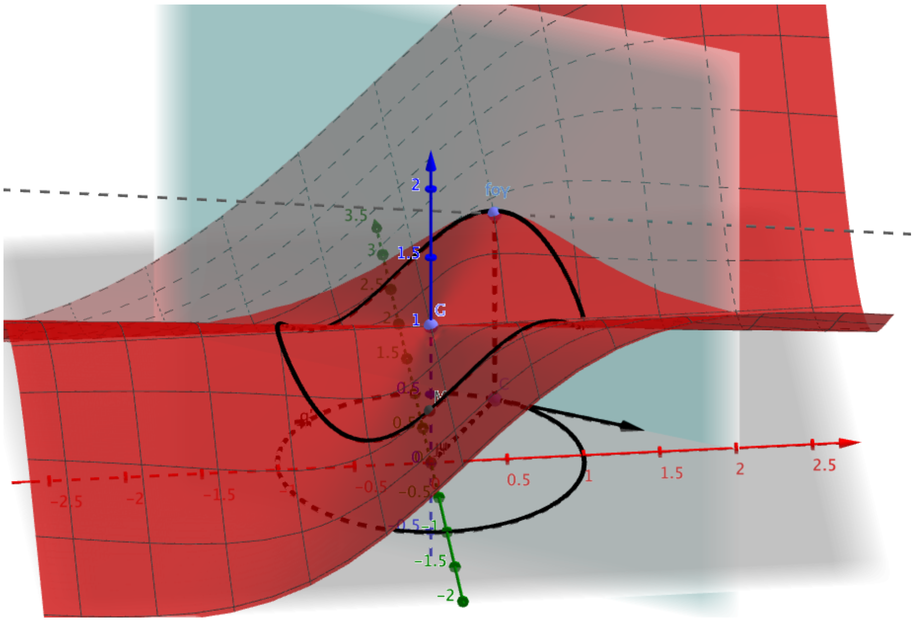
Nous n'allons pas rentrer dans les détails, ce n'est pas au programme, juste en parler un peu avec les mains pour donner des interprétations géométriques de certains résultats (interprétations qui sont, elles, au programme... sans commentaire). Disons juste que si x et y sont $\mathscr{C}^1$, alors on dit que γ est $\mathscr{C}^1$ avec $\gamma' = (x', y')$, et alors γ' dirige la tangente à γ (quand ce n'est pas le vecteur nul) :



On peut alors interpréter $(f \circ \gamma)'(t)$ comme la dérivée de f le long de la courbe paramétrée γ, et alors la règle de la chaîne

$$(f \circ \gamma)'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle$$

ne dit rien d'autre que la dérivée de $f \circ \gamma$ est la dérivée directionnelle de f en γ selon γ', c'est-à-dire que si on prend les images de f sur γ, alors les pentes sont égales à la fois à $(f \circ g)'$ et à $D_{\gamma'(t)} f(\gamma(t))$:

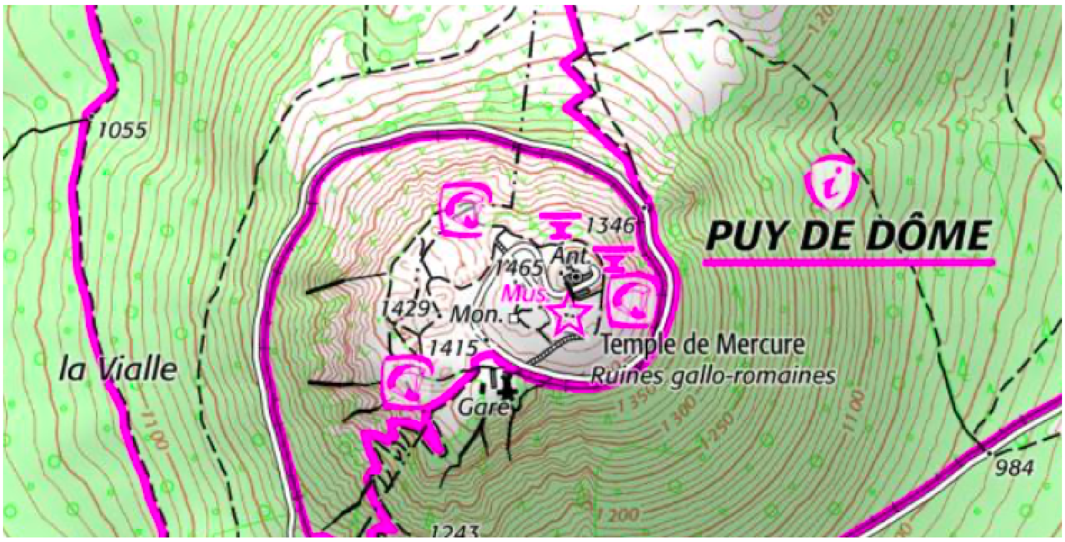


Expliquons un peu la figure ci-dessus : le cercle sur le plancher est la courbe γ, et on a tracé au-dessus son image par f (le petit serpentin). En le point $f \circ \gamma(t)$, le coefficient directeur de la tangente à la courbe (la droite en pointillée) est égal à $(f \circ \gamma)'(t)$ mais aussi à la dérivée directionnelle de $f \circ \gamma$ selon la direction γ'...

Donnons une dernière utilisation des arcs paramétrés nous permettant de mieux comprendre géométriquement les gradients. Tout d'abord :

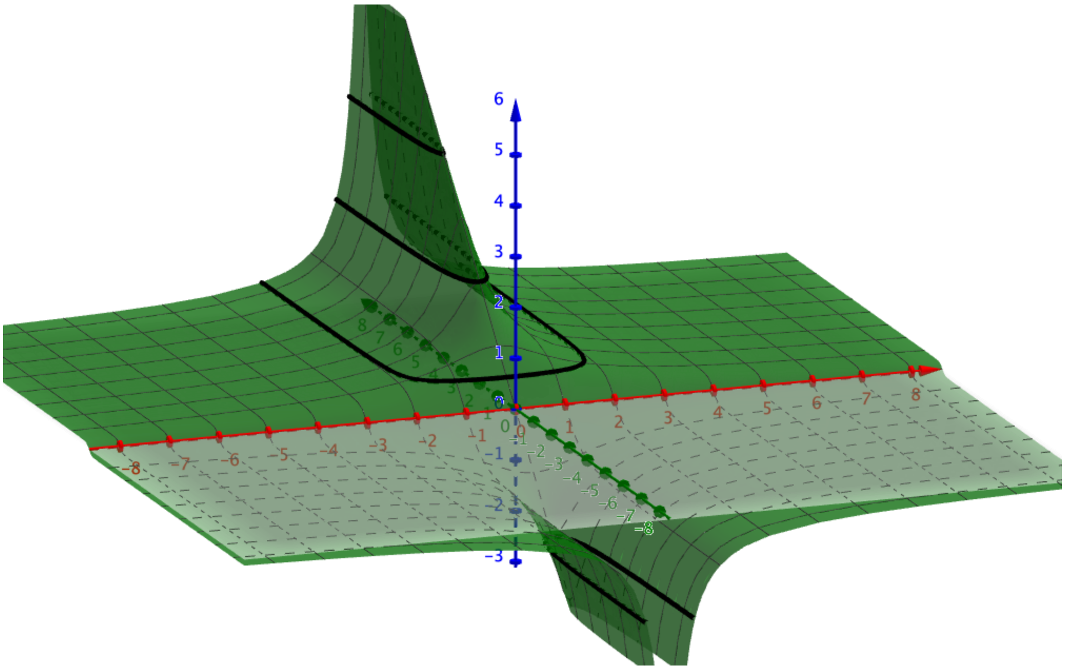
Définition. Soit $a \in \mathbb{R}$. La ligne de niveau de hauteur a est l'ensemble des points $(x, y) \in \mathbb{R}^2$ tels que $h(x, y) = a$.

Remarque : En d'autres termes, il s'agit de $f^{-1}(a)$. Encore en d'autres termes, ce sont les points du plan dont l'altitude de l'image par f vaut a. Les lignes de niveau sont donc les courbes dont l'altitude de l'image est constante. Elles sont représentées sur les cartes IGN, par exemple sur celle du Puy de Dôme (l'Auvergne est une très belle région) :



Les lignes sont les points de même altitude, sur la carte ci-dessus, elles vont de 10 mètres en 10 mètres : plus elles sont rapprochées, plus l'altitude augmente vite, et donc plus la porte est forte (et plus les cuisses et les genoux travaillent). Un exemple dans un cadre plus mathématique (les lignes de niveau étant les points ayant la même altitude, on les obtient en prenant l'intersection du graphe de f avec un plan horizontal) : on a tracé ci-dessous trois lignes de niveau (pour $a = -2, 1/2, 2$ et 5) :

Bon, en fait, les lignes de niveau sont plutôt ces courbes mais dans le plan (c'est comme si on les regardait par dessus) : on les trace ci-dessous.

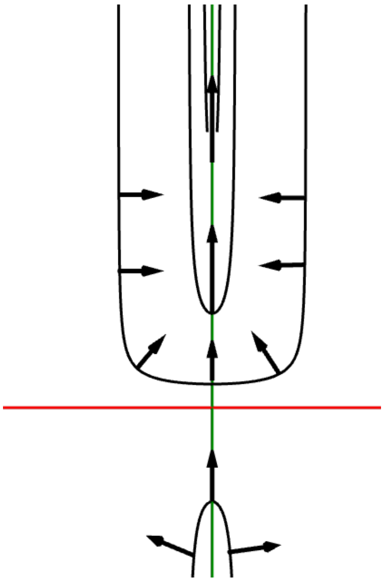


À l'aide d'un paramétrage $\mathscr{C}^1$ (dont l'existence est très difficile à prouver), on obtient le résultat suivant :

Théorème (admis). Les gradients d'une fonction $\mathscr{C}^1$ sont orthogonaux aux lignes de niveau.

Remarque : Ainsi, les gradients sont à la fois orthogonaux aux lignes de niveau, et dirigés vers la direction de plus forte pente (là où ça monte le plus), ce qui les rend particulièrement faciles à tracer. Si on reprend l'exemple ci-dessus, cela donne (précisons que les gradients n'ont aucune raison d'être de même norme, ce n'est que par commodité graphique qu'on a pris des vecteurs de même norme) :

Par conséquent, si vous faites une randonnée, vous repérez les endroits où la pente est plus forte : ce sont les directions orthogonales aux lignes de niveau, orientées vers la direction où l'altitude augmente.



V Recherche d'extrema

Définition. Soit $f : D \to \mathbb{R}$ où D est une partie de $\mathbb{R}^2$. Soit $X_0 = (x_0, y_0) \in \mathbb{R}^2$. On dit que f admet en X_0 :

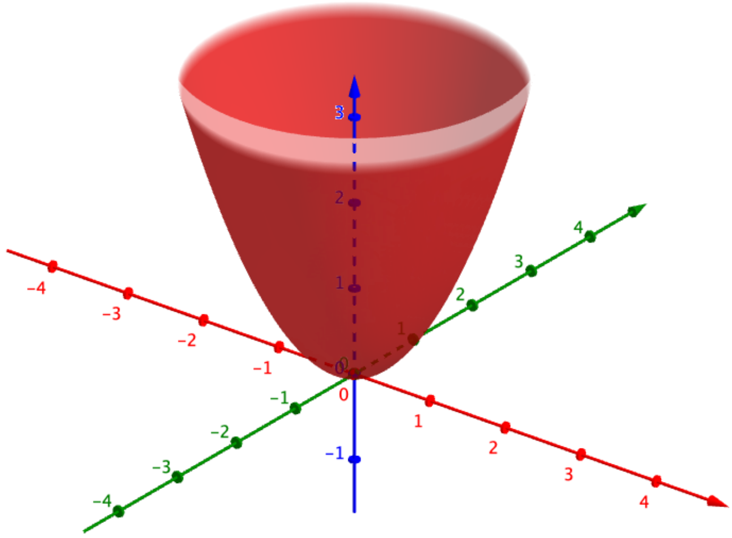
- un maximum global si, pour tout $X \in D, f(X) \leqslant f(X_0)$.
- un minimum global si, pour tout $X \in D, f(X) \geqslant f(X_0)$.
- un maximum local s'il existe $\varepsilon > 0$ tel que pour tout $X \in D \cap B(X_0, \varepsilon), f(X) \leqslant$

$f(X_0)$.

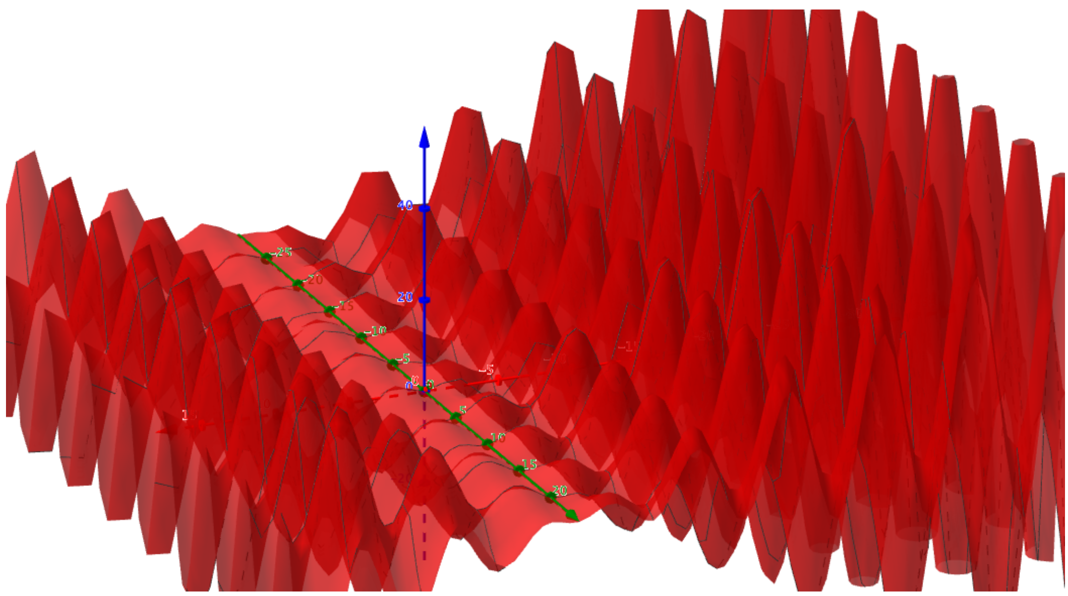
- un minimum local s'il existe $\varepsilon > 0$ tel que pour tout $X \in D \cap B(X_0, \varepsilon), f(X) \geqslant f(X_0)$.
- un extremum global (respectivement local) si f admet en X_0 un minimum ou un maximum global (respectivement local).

Définition analogue à celle pour les fonctions d'une variable : une fonction admet un extremum local en X_0 si f admet en X_0 un extremum sur un voisinage de X_0. On ne précise pas si la boule est ouverte ou fermée car cela n'a aucune importance, même si en général on prend une boule ouverte.

Remarque : Géométriquement, il est facile de repérer les extrema globaux : ce sont les endroits où l'attitude est maximale (pour les maxima) ou minimale (pour les minima), bien sûr quand de tels points existent. Par exemple, $f : (x, y) \mapsto x^2 + y^2$ admet en $(0, 0)$ un minimum global puisque, pour tout $(x, y), f(x, y) \geqslant 0 = f(0, 0)$.



Géométriquement, les extrema locaux sont les « tas de sables » (pour les maxima) ou les « cuvettes » pour les minima : ci-dessous le graphe d'une fonction admettant une infinité de maxima et de minima locaux mais aucun extremum global.



Pour se simplifier la vie, là aussi, nous nous restreindrons au cas où la fonction f est définie sur un ouvert U, ce qui nous permettra d'utiliser les outils des paragraphes précédents (typiquement les dérivées partielles). On se redonne donc dans la suite une fonction $f : U \to \mathbb{R}$ avec U un ouvert de $\mathbb{R}^2$.

Définition. On suppose que f est de classe $\mathscr{C}^1$. Soit $X_0 = (x_0, y_0) \in U$. On dit que X_0 est un point critique de f si $\nabla f(X_0) = 0$ c'est-à-dire si :

$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0$$

Remarque : En d'autres termes, un point critique est un point en lequel les deux dérivées partielles de f s'annulent, ce qui est tout à fait analogue à la définition d'un point critique pour les fonctions d'une variable. La condition d'extremum local EN UN POINT INTÉRIEUR est elle aussi encore valable :

On dit qu'un point est intérieur lorsqu'il existe un voisinage de ce point inclus dans l'ensemble : puisqu'on travaille sur un ouvert, tout point est intérieur, ce qui simplifie grandement les choses, sinon il faut aussi regarder sur la frontière, ce qui est plus compliqué. C'est pour cela qu'on travaille sur un ouvert en première année. L'année prochaine, grâce à la compacité, vous pourrez travailler sur des domaines non ouverts.

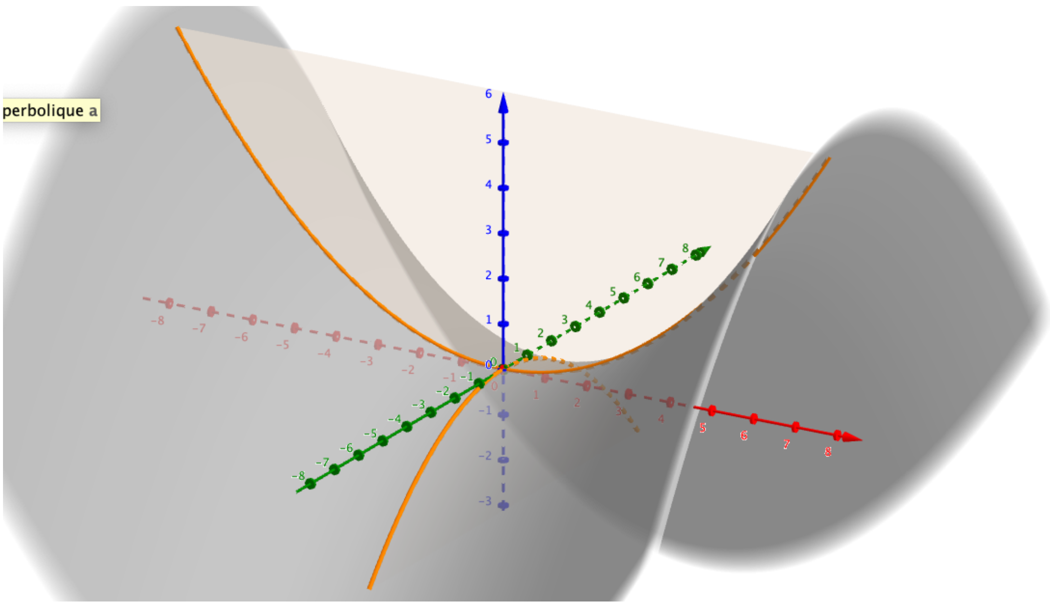
Proposition (Condition nécessaire d'extremum local en un point intérieur). Soit $X_0 = (x_0, y_0) \in U$ avec U ouvert. Si f est $\mathscr{C}^1$ et si f admet un extremum local en X_0, alors X_0 est un point critique, c'est-à-dire que

$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0$$

Démonstration. Supposons que f admette en X_0 un minimum local (raisonnement analogue dans l'autre cas). La fonction $f_1 : x \mapsto f(x, y_0)$ admet donc un minimul local en x_0 donc sa dérivée est nulle (condition nécessaire d'extremum local en un point intérieur), c'est-à-dire que $\dfrac{\partial f}{\partial x}(x_0, y_0) = 0$. De même pour l'autre.

Remarque : Comme pour les fonctions d'une seule variable, la réciproque est fausse ! On peut avoir ce qu'on appelle « un point col » ou « un point selle » (penser à un col de montagne ou une selle de cheval) : les deux dérivées partielles peuvent être nulles mais la fonction peut monter d'un côté et descendre d'un autre. L'interprétation d'un point critique est la suivante : quand on arrive parallèlement à l'axe des abscisses ou parallèlement à l'axe des ordonnées, alors la dérivée de la courbe est nulle, mais cela ne suffit pas pour avoir un maximum ou un minimum, même local :

Appellation personnelle : un point Pringle. À l'aide de l'interprétation géométrique des dérivées partielles vue au paragraphe III.1 : la dérivée selon x est la dérivée de la fonction obtenue « en allant d'Est en Ouest », tandis que celle selon y est la dérivée de la fonction obtenue en allant « du Nord au Sud ». S'il y a un extremum local, alors c'est aussi le cas de ces deux fonctions, donc les dérivées partielles s'annulent, mais on voit ci-contre que ce n'est pas équivalent.



Méthode : On commence par vérifier qu'on est sur un ouvert (ce sera systématiquement le cas cette année mais pas l'année prochaine) et ensuite on cherche les points critiques c'est-à-dire les points où les deux dérivées partielles s'annulent. Ensuite, on fait au cas par cas :

- on essaye de voir si ce ne serait pas trivialement un extremum global (par exemple si $f(X_0) = 0$ et si f est positive, alors f admet en X_0 un minimum global).
- si on essaye de prouver que ce n'est pas un extremum local, on essaye de faire varier une seule coordonnée, et on regarde si c'est plus grand (alors ce n'est pas un minimum) ou plus petit (et alors ce n'est pas un maximum), ce qui permet de montrer que ce n'est pas un extremum local (parfois, il n'est pas facile de savoir dans quelle direction chercher, mais en général l'énoncé donne une indication).
- si on a un extremum local et si on veut prouver que ce n'est pas un extremum global, on essaye de faire tendre f vers $\pm\infty$ dans une certaine direction.
- parfois, rien ne marche : on cherche donc le signe de $f(x_0 + h, y_0 + k) - f(x_0, y_0)$ (méthode pour comparer deux quantités : on fait la différence et on donne le signe).

Faire un changement de variable, comme pour les DL en un point différent de 0, permet de se ramener au voisinage de 0, ce qui simplifie les choses.

Exemples :

- Donner les extrema locaux éventuels de $f : (x, y) \mapsto x^2 + y^2$.

  f est de classe $\mathscr{C}^1$. Soit $(x, y) \in \mathbb{R}^2$. Alors

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

  Alors

  $$\left\{\begin{array}{rcl} \dfrac{\partial f}{\partial x}(x, y) & = & 0 \\ \dfrac{\partial f}{\partial y}(x, y) & = & 0 \end{array}\right. \iff x = y = 0$$

  On en déduit que $(0, 0)$ est l'unique point critique de f. Or, $f(0, 0) = 0$ et f est positive donc f admet en $(0, 0)$ un minimum global et n'a pas d'autre extremum (local ou global).
- Donner les extrema locaux éventuels de $f : (x, y) \mapsto x^2 - y^2$.

  De même, $(0, 0)$ est l'unique point critique de f. Or, $f(0, 0) = 0$. Si $x \in \mathbb{R}^*$, $f(x, 0) = x^2 > 0$ donc f n'admet pas en $(0, 0)$ un maximum local, et $f(0, x) = -x^2 < 0$ donc f n'admet pas en $(0, 0)$ un minimum local. On en déduit donc que f n'admet aucun extremum local.
- Donner la nature des points critiques de f définie sur $\mathbb{R}^2$ par $f(x, y) = xe^{-x-y^4}$.

  f est $\mathscr{C}^1$. Soit $(x, y) \in \mathbb{R}^2$.

  $$\frac{\partial f}{\partial x}(x, y) = e^{-x-y^4}(1 - x) \qquad \text{et} \qquad \frac{\partial f}{\partial y}(x, y) = -4y^3 xe^{-x-y^4}$$

  Alors :

$$\begin{cases} \dfrac{\partial f}{\partial x}(x,y) & = & 0 \\ \\ \dfrac{\partial f}{\partial y}(x,y) & = & 0 \end{cases} \iff \begin{cases} e^{-x-y^4}(1-x) & = & 0 \\ \\ -4y^3xe^{-x-y^4} & = & 0 \end{cases}$$

$$\iff \begin{cases} (1-x) & = & 0 \\ \\ -4y^3x & = & 0 \end{cases}$$

$$\iff \begin{cases} x & = & 1 \\ \\ -4y^3 & = & 0 \end{cases}$$

$$\iff x=1 \qquad \text{et} \qquad y=0$$

Le point $(1,0)$ est par conséquent l'unique point critique de f, et $f(1,0)=e^{-1}$. Or, pour tout $(x,y)\in\mathbb{R}^2, f(x,y)\leqslant xe^{-x}$ et une rapide étude de fonction montre que la fonction $x\mapsto xe^{-x}$ admet un maximum en $x=1$. Ainsi, $f(x,y)\leqslant xe^{-x}\leqslant 1\times e^{-1}=f(1,0)$: f admet un maximum global en $(1,0)$.

- Soit $f:(x,y)\mapsto x\times\left((\ln(x))^2+y^2\right)$, définie sur $\mathbb{R}_+^*\times\mathbb{R}$. f admet-elle un minimum global ? un maximum global ? Donner les points critiques de f et les étudier.

  f est définie sur $\mathbb{R}^{+*}\times\mathbb{R}$: f est par conséquent à valeurs positives, et $f(1,0)=0$: f admet donc en $(1,0)$ un minimum global. f n'admet cependant pas de maximum global car $f(1,y)=y^2\xrightarrow[y\to+\infty]{}+\infty$. Cherchons à présent les extrema locaux. f est $\mathscr{C}^1$ et les dérivées partielles de f sont données par

  $$\frac{\partial f}{\partial x}(x,y)=\ln(x)^2+y^2+x\times\frac{2\ln(x)}{x}=\ln(x)^2+y^2+2\ln(x) \qquad \text{et} \qquad \frac{\partial f}{\partial y}(x,y)=2xy$$

  Alors (rappelons que $x>0$) :

  $$\begin{cases} \dfrac{\partial f}{\partial x}(x,y) & = & 0 \\ \\ \dfrac{\partial f}{\partial y}(x,y) & = & 0 \end{cases} \iff \begin{cases} \ln(x)^2+y^2+2\ln(x) & = & 0 \\ \\ 2xy & = & 0 \end{cases}$$

  $$\iff \begin{cases} \ln(x)\times(\ln(x)+2) & = & 0 \\ \\ y & = & 0 \end{cases}$$

  $$\iff \begin{cases} x=1 \text{ ou } x=e^{-2} \\ \\ y=0 \end{cases}$$

  Ainsi, f admet deux points critiques $(1,0)$ et $(e^{-2},0)$. On a déjà vu que le premier est un minimum global. Étudions l'autre point critique. Tout d'abord, $f(e^{-2},0)=4e^{-2}$. Or, pour tout $y\in\mathbb{R}$, $f(e^{-2},y)=4e^{-2}+e^{-2}y^2\geqslant f(e^{-2},0)$ et donc ce n'est pas un maximum local. De plus, une rapide étude de fonction montre que la fonction $x\mapsto x(\ln(x))^2$ est croissante sur $\left]0\,;e^{-2}\right]$ puis décroissante sur $\left[e^{-2}\,;1\right]$. Par conséquent, pour tout $x\in\left]0\,;1\right], f(x,0)\leqslant f(e^{-2},0)$ et donc $(e^{-2},0)$ n'est pas non plus un minimum : c'est un point col.

- Donner la nature des points critiques de $f:(x,y)\mapsto x^2-3x+xy+y^2$.

f est de classe $\mathscr{C}^1$. Soit $(x,y)\in\mathbb{R}^2$.

$$\frac{\partial f}{\partial x}(x,y)=2x-3+y \qquad \text{et} \qquad \frac{\partial f}{\partial y}(x,y)=x+2y$$

Alors (rappelons que $x>0$) :

$$\left\{\begin{array}{lcl}\dfrac{\partial f}{\partial x}(x,y) & = & 0\\[2ex] \dfrac{\partial f}{\partial y}(x,y) & = & 0\end{array}\right. \iff \left\{\begin{array}{lcl}2x-3+y & = & 0\\[2ex] x+2y & = & 0\end{array}\right.$$

$$\iff \left\{\begin{array}{lcl}x & = & 2\\[2ex] y & = & -1\end{array}\right.$$

Le seul point critique de f est en $(2,-1)$. Aucune des méthodes précédentes ne semble fonctionner. Soit $(h,k)\in\mathbb{R}^2$ et étudions le signe de $f(2+h,-1+k)-f(2,-1)$:

$$\begin{array}{rcl}f(2+h,-1+k)-f(2,-1) & = & (2+h)^2-3(2+h)+(2+h)(-1+k)+(-1+k)^2+3\\[2ex] & = & h^2+hk+k^2\end{array}$$

Or, on sait que $-2hk\leqslant h^2+k^2$ si bien que

$$f(2+h,-1+k)-f(2,-1)\geqslant\frac{h^2+k^2}{2}\geqslant 0$$

On en déduit que f admet en $(2,-1)$ un minimum global.