Chap. III

FONCTIONS DE PLUSIEURS VARIABLES.

" L'idée que les mathématiques devraient se faire "sans calculs", si elle vraie dans certains domaines, est en fait souvent mise en défaut. Ce qu'il faut supprimer, ce sont les calculs inutiles – mais pas les autres ! "

R. GODEMENT

Chap. II

FONCTIONS DE PLUSIEURS VARIABLES

" Il est vrai que les mathématiques deviennent un jeu "automatique" [illegible] certaines démarches, et ne font [illegible] tout au départ. Ce qu'il faut comprendre, ce sont les calculs menés – mais pas les calculs. "

R. [illegible]

1. CALCUL DIFFÉRENTIEL ET DÉVELOPPEMENT DU 1er ORDRE.

1.1 Préliminaires.

- En classe de Première, Terminale, et en L1 a été vue la dérivation de fonctions numériques de la variable réelle :

$$f : \underset{\nwarrow \text{ intervalle de } \mathbb{R}}{I \subset \mathbb{R}} \longrightarrow \mathbb{R} \qquad (1)$$
$$t \longmapsto f(t).$$

On sait ce que signifie " f est dérivable en t_0", ce qu'est la "dérivée de f en t_0" (notée $f'(t_0)$, $\frac{df}{dt}(t_0)$, $\left.\frac{df}{dt}\right|_{t_0}$), et on connaît le calcul différentiel qui va avec.

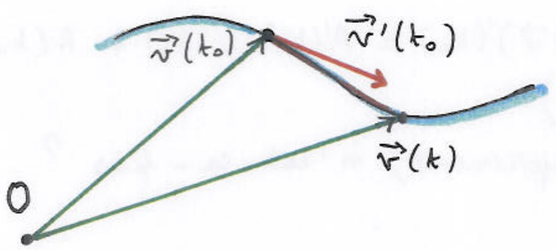
- En Terminales (un peu) et, surtout, en L1 a été vue la dérivation de fonctions vectorielles de la variable réelle. Rappelons brièvement de quoi il s'agit. Soit

$$\vec{v} : I \subset \mathbb{R} \longrightarrow \mathbb{R}^3 \text{ (ou } \mathbb{R}^2\text{)}$$
$$t \longmapsto \vec{v}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \leftarrow \text{composantes dans une base fixe de } \mathbb{R}^3 \qquad (2)$$

On met la flèche au-dessus de v (même si ce n'est pas nécessaire) pour ne pas oublier qu'il s'agit d'un vecteur (qui dépend de t).

On rappelle que $\vec{v}$ est dite dérivable en t_0 lorsque le quotient différentiel $\frac{\vec{v}(t)-\vec{v}(t_0)}{t-t_0}$ a une limite quand $t \to t_0$ (*); cette limite (qui est un vecteur) est notée $\vec{v}\,'(t_0)$, ou $\frac{d\vec{v}}{dt}(t_0)$, $\left.\frac{d\vec{v}}{dt}\right|_{t_0}$.

Penser à $\vec{v}\,'(t_0)$ comme à la vitesse (instantanée) en $t = t_0$ d'un mobile de coordonnées $(x(t), y(t), z(t))$ se déplaçant dans l'espace en fonction du temps t.



(*) Il faut bien intégrer ce que signifie "un vecteur $\vec{v}$ tend vers un vecteur $\vec{v}_0$" : cela signifie " $\|\vec{v}-\vec{v}_0\|$ (longueur de $\vec{v}-\vec{v}_0$) tend vers 0". d'une manière équivalente, chaque composante de $\vec{v}$ (un scalaire donc) tend vers la composante correspondante de $\vec{v}_0$.

Avec la représentation (2) de $\vec{v}(t)$, on a :

$(\vec{v}$ dérivable en $t_0) \Longleftrightarrow \left(\begin{array}{c}\text{chacune des (fonctions) composantes } x, y, z \text{ de} \\ \vec{v} \text{ est dérivable en } t_0.\end{array}\right)$;

de plus :

$$\vec{v}\,'(t_0) = \begin{pmatrix} x'(t_0) \\ y'(t_0) \\ z'(t_0) \end{pmatrix}. \qquad (3)$$

• <u>Exemples de règles de calcul différentiel</u>.

* Si $\vec{u}$ et $\vec{v}$ sont dérivables en t_0, il en est de même de la fonction <u>numérique</u> $t \longmapsto s(t) := \langle \vec{u}(t), \vec{v}(t) \rangle$ et

$$s'(t_0) = \langle \vec{u}\,'(t_0), \vec{v}(t_0) \rangle + \langle \vec{u}(t_0), \vec{v}\,'(t_0) \rangle ;$$

il en est de même de la fonction <u>vectorielle</u> $t \longmapsto w(t) := \vec{u}(t) \wedge \vec{v}(t)$ et

$$\vec{w}\,'(t_0) = \vec{u}(t_0) \wedge \vec{v}\,'(t_0) + \vec{u}\,'(t_0) \wedge \vec{v}(t_0).$$

* Si $\vec{u}, \vec{v}, \vec{w}$ sont dérivables en t_0, il en est de même du produit mixte $t \longmapsto m(t) := [\vec{u}(t), \vec{v}(t), \vec{w}(t)]$ et

$$m'(t_0) = [\vec{u}\,'(t_0), \vec{v}(t_0), \vec{w}(t_0)] + [\vec{u}(t_0), \vec{v}\,'(t_0), \vec{w}(t_0)] + [\vec{u}(t_0), \vec{v}(t_0), \vec{w}\,'(t_0)]$$

(i.e., on fait "défiler le prime" sur chaque facteur).

<u>Exercice 1</u> (adaptation du calcul différentiel précédent).

Soit $A : I \longrightarrow \mathcal{M}_3(\mathbb{R})$, $t \longmapsto A(t) = [a_{ij}(t)]$ dérivable en t_0 (ce qui revient à dire que tous les a_{ij} sont dérivables en t_0) ;

soit $\vec{v} : I \longrightarrow \mathbb{R}^3$ dérivable en t_0.

Alors, la fonction vectorielle

$$A\vec{v} : I \longrightarrow \mathbb{R}^3$$
$$t \longmapsto (A\vec{v})(t) := A(t)\,\vec{v}(t)$$

est dérivable en t_0 et $(A\vec{v})'(t_0) = A'(t_0)\,\vec{v}(t_0) + A(t_0)\,\vec{v}\,'(t_0)$ (où $A'(t_0) = [a'_{ij}(t_0)] \in \mathcal{M}_3(\mathbb{R})$).

Le résultat n'est pas surprenant, n'est-ce-pas ?

1.2 <u>Fonctions numériques de plusieurs variables</u>.

Là on change de registre... les choses sont un peu plus délicates.

Soit $f : U \subset \mathbb{R}^3$ (ou $\mathbb{R}^2$) $\longrightarrow \mathbb{R}$ une fonction <u>numérique</u> <u>des</u> variables x, y, z (ou x, y). C'est une situation très fréquente en

Sciences de l'ingénieur : les variables peuvent être des coordonnées, une température, un coefficient de pression, etc. Il faut donner un sens à "f est dérivable en $(x_0, y_0, z_0) \in \mathcal{O}$".

- Dérivées partielles

C'est simple : on gèle toutes les variables sauf une (mettons la 1ère), et on regarde si la fonction résultante est dérivable (comme fonction de la variable réelle) :

$$\varphi_1 : x \longmapsto f(x, \underbrace{y_0, z_0})$$ on agit comme s'il s'agissait de constantes.

Si la fonction (dite partielle) φ_1 est dérivable en x_0, on dit que f admet une dérivée partielle par rapport à la 1ère variable (ou par rapport à x) au point (x_0, y_0, z_0).

$$\frac{f(x_0+h, y_0, z_0) - f(x_0, y_0, z_0)}{h} \underset{h \to 0}{\longrightarrow} \varphi_1'(x_0).$$

On fait cela, si c'est possible, par rapport à chaque variable (x, y et z), et en tout point (x, y, z) de $\mathcal{O}$.

Notations diverses utilisées pour les dérivées partielles:

$\dfrac{\partial f}{\partial x}(x_0, y_0, z_0)$: dérivée partielle de f par rapport à la 1ère variable (appelée x) prise au point (x_0, y_0, z_0);

$\left.\dfrac{\partial f}{\partial x}\right|_{(x_0, y_0, z_0)}$;

$\dfrac{\partial f}{\partial x}$ seul, sans référence au point (x_0, y_0, z_0) où la (fonction) dérivée partielle est évaluée ; c'est un abus d'écriture (mais très fréquent) car $\dfrac{\partial f}{\partial x}$ est, elle aussi, une fonction des variables x, y, z.

$\partial_1 f(x_0, y_0, z_0)$; $\partial_x f(x_0, y_0, z_0)$, $f'_x(x_0, y_0, z_0)$, $D_1 f(x_0, y_0, z_0)$.

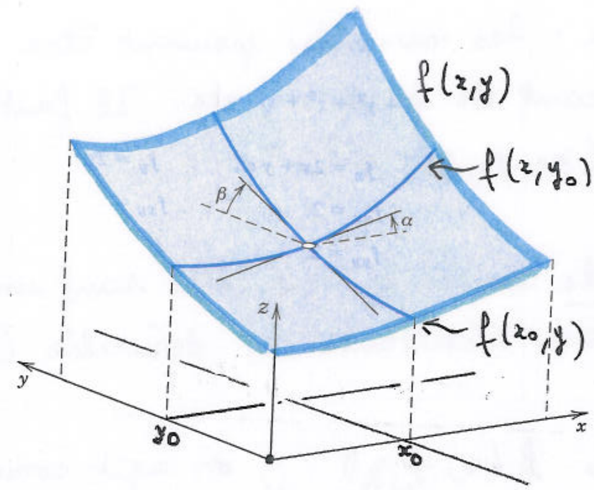
Exemple 1. $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$

$$(x, y) \longmapsto f(x, y) = x^2 y + x \sin y.$$

Alors

$$\frac{\partial f}{\partial x}(x_0, y_0) = 2 x_0 y_0 + \sin y_0,$$

$$\frac{\partial f}{\partial y}(x_0, y_0) = x_0^2 + x_0 \cos y_0.$$



<u>Interprétation</u> <u>géométrique</u> : $z = f(x,y)$ est représenté graphiquement par une surface de l'espace, en <u>gelant</u> la variable x_0, on fait une <u>coupe</u> le long du plan d'équation $x = x_0$, et on s'intéresse du coup à la <u>courbe</u> représentant la fonction partielle $y \longmapsto f(x_0, y)$.

On peut continuer à "dériver partiellement" (si la dérivation est licite) ; on obtient ainsi les dérivées partielles secondes (ou dérivées partielles d'ordre 2) :

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \text{ noté } \frac{\partial^2 f}{\partial x^2} \text{ ou } f''_{xx} \ldots$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \text{ noté } \frac{\partial^2 f}{\partial x \partial y} \text{ ou } f''_{xy} \text{ ou } \partial_1 \partial_2 f$$

etc.

Dans le cas de l'exemple au-dessus, $\frac{\partial f}{\partial x} = 2xy + \sin y$, de sorte que

$$\frac{\partial^2 f}{\partial x^2}(x_0, y_0) = 2y_0, \quad \frac{\partial^2 f}{\partial y^2}(x_0, y_0) = -x_0 \sin y_0,$$

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = 2x_0 + \cos y_0.$$

Si tout se passe bien (ce sera le cas dans notre contexte) (*), l'ordre de dérivation partielle importe peu : $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$.

On appelle "équation aux dérivées partielles" une équation où la fonction inconnue intervient par ses dérivées partielles. Par exemple :

(*) Mathématiquement, une condition habituelle pour qu'il en soit ainsi est que <u>toutes</u> les (fonctions) <u>dérivées partielles</u> en jeu soient des fonctions <u>continues</u> de x, y, z. Ce qui veut dire ?

$$\frac{\partial f}{\partial t} - k^2\left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\right) = 0$$ (équation de la chaleur en 2D ; inconnue $f(t,x,y)$)

$$\frac{\partial^2 f}{\partial t^2} - c^2\left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}\right) = 0$$ (équation des ondes en 3D ; inconnue $f(t,x,y,z)$).

La fonction numérique $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$ qui intervient si souvent dans les équations aux dérivées partielles s'appelle le laplacien de f ; il se note Δf et se lit " delta f ". (La notation $\nabla^2 f$ est à proscrire ici).

- Gradient d'une fonction.

Lorsque toutes les dérivées partielles premières de f existent en $(x_0,y_0,z_0) \in U$, on appelle vecteur gradient le vecteur dont les composantes sont les (fonctions) dérivées partielles ; il se note $\overrightarrow{\nabla} f(x_0,y_0,z_0)$ (ou $\overrightarrow{\text{grad}} f(x_0,y_0,z_0)$) et se lit " nabla f " (*) :

$$\boxed{\overrightarrow{\nabla} f(x_0,y_0,z_0) = \begin{pmatrix} \frac{\partial f}{\partial x}(x_0,y_0,z_0) \\ \frac{\partial f}{\partial y}(x_0,y_0,z_0) \\ \frac{\partial f}{\partial z}(x_0,y_0,z_0) \end{pmatrix}.} \qquad (4)$$

Il s'agit bien d'un vecteur, d'où notre insistance à mettre une flèche.

Définition 2. On dit que f est différentiable (ou dérivable) en $(x_0,y_0,z_0) \in U$ si les dérivées partielles de f en (x_0,y_0,z_0) existent (et donc le vecteur $\overrightarrow{\nabla} f(x_0,y_0,z_0)$) et si

$$f(x_0+h, y_0+k, z_0+l) = f(x_0,y_0,z_0) + \left\langle \overrightarrow{\nabla} f(x_0,y_0,z_0), \begin{pmatrix} h \\ k \\ l \end{pmatrix} \right\rangle + \|(h,k,l)\|\, \varepsilon(h,k,l), \qquad (5)$$

où $\varepsilon(h,k,l)$ tend vers 0 quand (le vecteur) (h,k,l) tend vers $(0,0,0)$.

Il faut bien comprendre le sens du développement (5), dit développement du 1er ordre autour de (x_0,y_0,z_0) :

(*) Que je n'entende pas " delta f " pour $\overrightarrow{\nabla} f$!
Certains, au tout début, dénommaient ∇f par " atled "..., pourquoi diable ?
Nabla est le nom d'un ancien instrument de musique, en forme de triangle renversé.

$$\left.\begin{aligned} f(x_0+h, y_0+k, z_0+l) &= f(x_0,y_0,z_0) + \\ &\frac{\partial f}{\partial x}(x_0,y_0,z_0)\,h + \frac{\partial f}{\partial y}(x_0,y_0,z_0)\,k + \frac{\partial f}{\partial z}(x_0,y_0,z_0)\,k \end{aligned}\right\} = \begin{array}{l}\text{approximation}\\ \text{linéaire}\\ \text{du } 1^{\text{er}} \text{ ordre.}\end{array}$$

$$+ \text{ terme tendant vers } 0 \text{ plus vite que } \|(h,k,l)\|.$$

On démontre en Mathématiques le résultat suivant : soit U un ouvert de $\mathbb{R}^3$ (*) ; alors :

$\left(\begin{array}{l} f \text{ est différentiable en tout point } (x,y,z)\in U \text{ et l'application} \\ (x,y,z) \longmapsto \vec{\nabla} f(x,y,z) \text{ est continue } (**) \end{array}\right)$

$$\Longleftrightarrow$$

$\left(\begin{array}{l} \text{les trois dérivées partielles } \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \text{ existent en tout point de } U \\ \text{et sont des fonctions continues sur } U. \end{array}\right)$

On dit alors que f est <u>continûment</u> $\binom{**}{*}$ <u>différentiable sur U</u>, ou <u>de classe C^1 sur U</u> ; notation : <u>$f \in C^1(U)$</u>.

- Un mot des <u>notations différentielles</u>.

Si Δx désigne l'accroissement $(x+\Delta x) - x$, la dérivabilité de f en x (f est une fonction d'une seule variable) se traduit par

$$\Delta f = f(x+\Delta x) - f(x) = f'(x)\,\Delta x + o(|\Delta x|).$$

Moyennant le terme négligeable (mais qu'on ne néglige pas !) $o(|\Delta x|)$, Δf apparaît comme " linéaire en Δx ". A la limite, si on imagine un "accroissement infiniment petit" de la variable x, noté dx, on peut convenir que "l'accroissement infiniment petit" correspondant de f, noté df, vérifie $df = f'(x)\,dx$. Ceci est un langage utilisé couramment en Physique et Sciences de l'ingénieur, il donne une vision quelque peu intuitive de ce qu'est $f'(x)$.

(*) On dit que U est un ouvert (ou une partie ouverte) de $\mathbb{R}^3$ si, pour tout $(x_0,y_0,z_0)\in U$, il existe $r>0$ telle que la boule centrée en (x_0,y_0,z_0) et de rayon r soit contenue dans U. Autre notation fréquente : Ω.

(**) Si $G : U \to \mathbb{R}^3$ est à valeurs vectorielles (comme c'est le cas de $\vec{\nabla} f$), on dit que G est continue lorsque chaque fonction-composante de G est continue.

$\binom{**}{*}$ continûment s'écrit avec $\hat{u}$ et non ue ...

Pour les fonctions f de plusieurs variables x, y, z, au vecteur $\overrightarrow{\nabla f}(x,y,z)$ est associée une forme linéaire sur $\mathbb{R}^3$ (= application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}$):

$$(h,k,\ell) \longmapsto \langle \overrightarrow{\nabla f}(x,y,z), (h,k,\ell) \rangle .$$

On <u>note</u> cette forme linéaire comme ceci:

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz .$$

Ceci peut présenter des avantages pour le calcul "mécanique" des dérivées partielles (et donc des différentielles) des fonctions composées.

<u>Exemple 3.</u> (Dérivation en chaîne).

Schéma :

$$\begin{tikzcd} \Omega \subset \mathbb{R}^2 \arrow[r, "\vec{X}"] \arrow[dr, "f\circ\vec{X}"'] & U \subset \mathbb{R}^3 \arrow[d, "f"] \\ & \mathbb{R} \end{tikzcd} \qquad (u,v) \longmapsto \vec{X}(u,v) = \begin{pmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{pmatrix} ; \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \longrightarrow f(x,y,z)$$

Hypothèses: Ω, U ouverts ; $\vec{X}$ (c'est-à-dire x, y, z) et f des fonctions continûment différentiables.

Objectif : différentielle (et donc dérivées partielles) de la fonction composée $f\circ\vec{X} : (u,v) \longmapsto f(x(u,v), y(u,v), z(u,v))$.

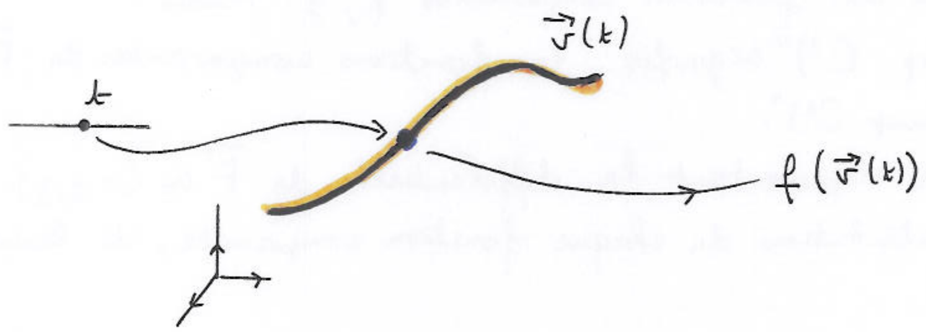
Résultats :

$$\frac{\partial (f\circ\vec{X})}{\partial u} = \frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial u} + \frac{\partial f}{\partial z}\cdot\frac{\partial z}{\partial u} ;$$

$$\frac{\partial (f\circ\vec{X})}{\partial v} = \frac{\partial f}{\partial v}\cdot\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial v} + \frac{\partial f}{\partial z}\cdot\frac{\partial z}{\partial v} . \qquad (6)$$

<u>Exemple 4.</u> (Dérivation le long d'un arc)

$$\begin{tikzcd} I \subset \mathbb{R} \arrow[r, "\vec{v}"] \arrow[dr, "f\circ\vec{v}"'] & \mathbb{R}^3 \arrow[d, "f"] \\ & \mathbb{R} \end{tikzcd} \qquad t \mapsto \vec{v}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} ; \quad \vec{v} \longrightarrow f(\vec{v})$$



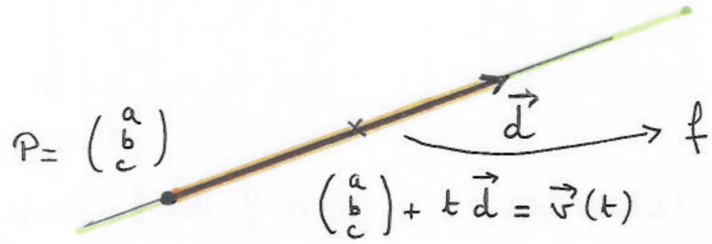
Résultats : si f et $\vec{v}$ sont différentiables (resp. C^1), il en est de même de la composée $t \longmapsto (f\circ\vec{v})(t) = f(\vec{v}(t))$; de plus

$$(f\circ\vec{v})'(t) = \langle \vec{\nabla} f(\vec{v}(t)), \vec{v}\,'(t) \rangle . \tag{7}$$
$$= \frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\cdot\frac{\partial z}{\partial t}$$

forme abrégée de $\frac{\partial f}{\partial x}(x(t), y(t), z(t)) . x'(t) + \dots$).

Cas particulier ; <u>la dérivée directionnelle</u>.

$$\vec{v}(t) = \underbrace{\begin{pmatrix} a \\ b \\ c \end{pmatrix}}_{\text{un point de } \mathbb{R}^3} + t \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} \longleftarrow \text{une direction } \vec{d} \text{ non nulle de } \mathbb{R}^3.$$



On regarde le comportement de f le long de la droite passant par P et dirigée par le vecteur $\vec{d}$:

$$\varphi(t) := f(P + t\vec{d})$$

Alors :

$$\varphi'(0) = \langle \vec{\nabla} f(P), \vec{d} \rangle ; \tag{8}$$

c'est <u>la dérivée directionnelle de f en P dans la direction $\vec{d}$</u>.

1.3 <u>Fonctions vectorielles de plusieurs variables.</u>

On cumule des difficultés : $\vec{F}$ est à valeurs <u>vectorielles</u> (dans $\mathbb{R}^2$ ou $\mathbb{R}^3$), donc avec des fonctions composantes, et ces fonctions composantes sont fonctions de <u>plusieurs variables</u> (2 ou 3) :

$$(x,y,z) \longmapsto \vec{F}(x,y,z) = \begin{pmatrix} f(x,y,z) \\ g(x,y,z) \end{pmatrix} \text{ par exemple.}$$

En fait, rien de nouveau : on traite $\vec{F}$ "étage par étage", c'est-à-dire en considérant chacune des fonctions composantes f, g. Ainsi :
"$\vec{F}$ différentiable (resp. C^1)" signifie "les fonctions composantes de $\vec{F}$ sont différentiables (resp. C^1)".
L'objet mathématique représentant la différentielle de $\vec{F}$ en (x,y,z) est l'assemblage des contributions de chaque fonction composante, de leurs gradients en (x,y,z).

On appelle <u>matrice jacobienne</u> de $\vec{F}$ en (x,y,z) la matrice suivante (2,3) :

$$J\vec{F}(x,y,z) := \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \end{bmatrix} \left(= \begin{bmatrix} \vec{\nabla} f(x,y,z)^T \\ \vec{\nabla} g(x,y,z)^T \end{bmatrix} \ (*) \right)$$

Dans le cas où $\vec{F}$ est à valeurs dans $\mathbb{R}^3$

$$(x,y,z) \longmapsto \vec{F}(x,y,z) = \begin{pmatrix} f(x,y,z) \\ g(x,y,z) \\ h(x,y,z) \end{pmatrix},$$

la matrice jacobienne de $\vec{F}$ en (x,y,z) est une matrice (3,3) :

$$J\vec{F}(x,y,z) := \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{bmatrix}.$$

Dans ce cas, on peut prendre le déterminant de cette matrice : cela s'appelle <u>le déterminant jacobien</u> (ou <u>le jacobien</u>) de $\vec{F}$ en (x,y,z). De même, on peut en prendre la trace : cela s'appelle <u>la divergence</u> de $\vec{F}$ en (x,y,z). En notations abrégées :

$$\boxed{\operatorname{div} \vec{F} = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z}.} \qquad (9)$$

Attention ! la divergence de $\vec{F}$ est un <u>scalaire</u>. Comme le suggère son nom, la divergence a une signification physique importante.

<u>Exemple 5</u>. Partant de $f : U \subset \mathbb{R}^3 \longrightarrow \mathbb{R}$ (fonction numérique), considérons $\vec{\nabla} f : U \subset \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ (à valeurs vectorielles). Alors, il est facile d'observer que

$$\operatorname{div}(\vec{\nabla} f) = \Delta f \qquad (10)$$

(la divergence du gradient est le laplacien).

(*) La transposition est là car le vecteur gradient $\vec{\nabla} f(x,y,z)$ est représenté (usuellement) en une colonne (ou une matrice à une colonne).

Exercice 6. (Changements de variables linéaires).
Soit $A \in \mathcal{M}_{2,3}(\mathbb{R})$ représentant (dans les bases canoniques) une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^3$; soit $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ de classe C^1.
Montrer que la fonction composée $(x,y) \longmapsto (f\circ A)(x,y) := f(A(x,y))$ a pour gradient

$$\overrightarrow{\nabla(f\circ A)}\,(x,y) = A^T \, \vec{\nabla} f \,(A(x,y)). \tag{11}$$

2. CALCUL DIFFÉRENTIEL ET DÉVELOPPEMENT DU SECOND ORDRE.

- On revient aux fonctions numériques de plusieurs variables

$$f : U \subset \mathbb{R}^2 \text{ ou } \mathbb{R}^3 \longrightarrow \mathbb{R}.$$

Définition 7. La fonction f, différentiable sur U (cf. p. 6) est dite deux fois différentiable en (x_0, y_0, z_0) si $\vec{F} = \vec{\nabla} f$ est différentiable en (x_0, y_0, z_0) (cf. § 1.3, p. 9).

De fait, on considérera des fonctions ayant plus de "régularité", les fonctions deux fois continûment différentiables sur U.

Définition 8. On dit que f est deux fois continûment différentiable sur U si toutes les dérivées partielles d'ordre deux de f existent et sont des fonctions continues (de (x,y,z)).
Notation pour cette classe de functions : $f \in C^2(U)$.

Une conséquence est le théorème de H. SCHWARZ (résultat admis) : l'ordre des dérivations partielles est sans conséquence sur le résultat ; par exemple

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \text{ soit } \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}. \tag{12}$$

Ceci va nous permettre de définir l'objet mathématique représentant la différentielle d'ordre deux, c'est la matrice hessienne (ou matrice des dérivées partielles secondes) de f

$$\underbrace{\begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\[2ex] \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{bmatrix}}_{\text{cas de 2 variables}} \qquad \left.\begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial x \partial z} \\[2ex] \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial^2 f}{\partial y \partial z} \\[2ex] \dfrac{\partial^2 f}{\partial z \partial x} & \dfrac{\partial^2 f}{\partial z \partial y} & \dfrac{\partial^2 f}{\partial z^2} \end{bmatrix}\right\} \text{cas de 3 variables.}$$

grâce au théorème de SCHWARZ(*), il s'agit de <u>matrices symétriques</u> ; on les note

$$Hf(x,y,z) \quad \text{ou bien} \quad \underline{\nabla^2 f(x,y,z)} \quad (\text{"nabla 2 de f"})$$

Ce dernier se lit "nabla 2 de f".

De même que $\nabla f(x_0,y_0,z_0)$ avait permis de définir une <u>forme linéaire</u> qui avait conduit à un développement du 1er ordre (cf. p.7 en haut), l'objet $\nabla^2 f(x_0,y_0,z_0)$ permet de définir une <u>forme quadratique</u> qui nous conduira à un développement du 2ème ordre.

Pour alléger l'écriture, on va s'en tenir aux fonctions de 2 variables : $(x,y) \in O \longmapsto f(x,y)$.

- <u>Formes quadratiques sur $\mathbb{R}^2$: a digest.</u>

Une <u>forme quadratique</u> q <u>sur $\mathbb{R}^2$</u> est une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie de la façon suivante :

$$(h,k) \in \mathbb{R}^2 \longmapsto q(h,k) = \langle M(h,k),(h,k) \rangle, \qquad (13)$$

où M est une matrice <u>symétrique</u> (2,2). Si $M = \begin{bmatrix} r & s \\ s & t \end{bmatrix}$, on détaille ce qui est en (13) en

$$\begin{bmatrix} r & s \\ s & t \end{bmatrix}\begin{pmatrix} h \\ k \end{pmatrix} = \begin{pmatrix} rh+sk \\ sh+tk \end{pmatrix} ; \left\langle \begin{pmatrix} rh+sk \\ sh+tk \end{pmatrix}, \begin{pmatrix} h \\ k \end{pmatrix} \right\rangle = \underline{rh^2 + tk^2 + 2s\,hk}.$$

attention au 2 ici

Autre manière d'écrire $q(h,k)$:

$$q(h,k) = [h,k] \times \begin{bmatrix} r & s \\ s & t \end{bmatrix} \times \begin{bmatrix} h \\ k \end{bmatrix}. \qquad (13)'$$

M est diagonalisable (dans $\mathbb{R}$) et a des valeurs propres <u>réelles</u> : il existe même U <u>orthogonale</u> telle que

$$U^T M U = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \qquad (14)$$

où λ_1 et λ_2 sont les 2 valeurs propres réelles de M.

Classification :

q (ou M) est dite <u>semi-définie positive</u> (resp. <u>définie positive</u>) lorsque $\lambda_1 \geq 0$ et $\lambda_2 \geq 0$ (resp. $\lambda_1 > 0$ et $\lambda_2 > 0$).

q (ou M) est dite semi-définie négative (resp. définie négative) lorsque $\lambda_1 \leq 0$ et $\lambda_2 \leq 0$ (resp. $\lambda_1 < 0$ et $\lambda_2 < 0$).

q (ou M) est dite <u>indéfinie</u> lorsque λ_1 et λ_2 sont de signe contraire.

Le calcul différentiel n'est pas très loin dans tout ça puisque :

(*) Ce SCHWARZ est le même que celui de l'inégalité (Ch. I, p.18) ; je ne veux plus voir de T dans son nom ...

$$\vec{\nabla} q(h,k) = M(h,k) = \begin{pmatrix} rh + sk \\ sh + tk \end{pmatrix} \text{ et } \nabla^2 q(h,k) = M = \begin{bmatrix} r & s \\ s & t \end{bmatrix} \qquad (15)$$

pour tout $(h,k) \in \mathbb{R}^2$.

THÉORÈME 9. (Développement du second ordre).
Soit $f : (x,y) \in U \subset \mathbb{R}^2 \longmapsto f(x,y)$ deux fois continûment différentiable sur U. On a alors le développement du second ordre suivant de f autour de $(x_0,y_0) \in U$:

$$f(x_0+h, y_0+k) = f(x_0,y_0) + \langle \vec{\nabla} f(x_0,y_0), (h,k) \rangle + \frac{1}{2} \langle M(h,k), (h,k) \rangle + (h^2+k^2)\, \varepsilon(h,k), \qquad (16)$$

où M est la matrice hessienne $\nabla^2 f(x_0,y_0)$ de f en (x_0,y_0), et $\varepsilon(h,k)$ tend vers 0 quand $(h,k) \to (0,0)$.

En format plus détaillé :

$$\left.\begin{aligned} f(x_0+h, y_0+k) = {} & f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)\, h + \frac{\partial f}{\partial y}(x_0,y_0)\, k \\ & + \frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(x_0,y_0)\, h^2 + \frac{\partial^2 f}{\partial y^2}(x_0,y_0)\, k^2 + 2\, \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)\, h k \right] \end{aligned}\right\} \text{approximation quadratique du second ordre.}$$

$$+ (h^2+k^2)\, \varepsilon(h,k).$$

Exercice 10. Tests sur le caractère de M.

Soit $M = \begin{bmatrix} r & s \\ s & t \end{bmatrix}$ Montrer que M est

- définie positive si et seulement si $r > 0$ et $rt - s^2 > 0$,
- définie négative si et seulement si $r < 0$ et $rt - s^2 > 0$,
- indéfinie si et seulement si $rt - s^2 < 0$.

Exercice 11. Développements exacts du second ordre d'une forme quadratique.

Soit $q(u) = \langle Mu, u \rangle$ une forme quadratique associée à la matrice symétrique $M \in \mathcal{M}_2(\mathbb{R})$. Ici $u \in \mathbb{R}^2$. Montrer :

$$\begin{aligned} q(u+v) &= q(u) + 2 \langle Mu, v \rangle + q(v) \\ &= q(u) + \langle \nabla q(u), v \rangle + \frac{1}{2} \langle \nabla^2 q(u) v, v \rangle . \end{aligned}$$

3. APPLICATION : CONDITIONS D'OPTIMALITÉ DANS LE PROBLÈME DE LA MINIMISATION D'UNE FONCTION DE 2 VARIABLES.

On ne s'adresse dans cette partie qu'à des fonctions numériques.

3.1 Illustration physique de la notion de gradient.

Considérons f fonction de 2 ou 3 variables, nous la supposons de classe C^1. Soit P un point et $\vec{d}$ une direction (de $\mathbb{R}^2$ par exemple). On a vu (p. 9) que $\varphi'(0) = \langle \vec{\nabla} f(P), \vec{d} \rangle$ lorsque $\varphi(t) = f(P + t\vec{d})$ (évaluation de f le long de la droite passant par P et dirigée par $\vec{d}$).
Question : dans quelle direction $\vec{d}$ cette dérivée est-elle maximale ? minimale ? La réponse est fournie par l'inégalité de SCHWARZ et les cas d'égalité qui vont avec (Chap. I, p. 18). Pour des raisons d'homogénéité, supposons $\vec{d}$ unitaire (i.e., $\|\vec{d}\| = 1$) ; évitons le cas trivial en supposant $\vec{\nabla} f(P) \neq \vec{0}$. On a alors :

- $-\|\vec{\nabla} f(P)\| \leq \langle \vec{\nabla} f(P), \vec{d} \rangle \leq \|\vec{\nabla} f(P)\|$;
- Egalité à gauche pour $\vec{d} = -\dfrac{\vec{\nabla} f(P)}{\|\vec{\nabla} f(P)\|}$; (17)
- Egalité à droite pour $\vec{d} = \dfrac{\vec{\nabla} f(P)}{\|\vec{\nabla} f(P)\|}$.

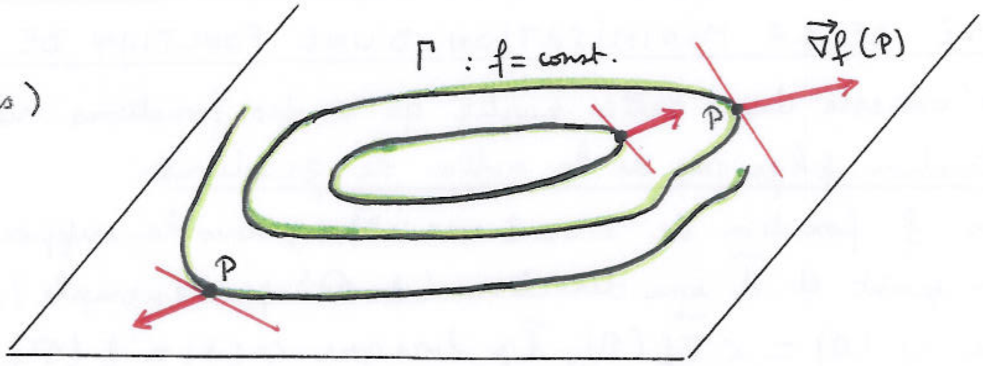
Ainsi, à partir du point P, $\dfrac{\vec{\nabla} f(P)}{\|\vec{\nabla} f(P)\|}$ est la direction unitaire de "plus grande montée" et $-\dfrac{\vec{\nabla} f(P)}{\|\vec{\nabla} f(P)\|}$ est la direction unitaire de "plus profonde descente".
Si $\vec{d}$ est orthogonal à $\vec{\nabla} f(P)$, $\varphi'(0) = \langle \vec{\nabla} f(P), \vec{d} \rangle = 0$, on dit que "$f$ est stationnaire dans cette direction".

Interprétation géométrique :

→ (2 variables). Si Γ est la courbe de niveau $\{(x,y) \mid f(x,y) = c\}$ (où c est un réel), si $P \in \Gamma$, et si $\vec{\nabla} f(P) \neq \vec{0}$, $\vec{\nabla} f(P)$ est un vecteur (dit) normal à Γ en P ($\vec{\nu} = \vec{\nabla} f(P) / \|\vec{\nabla} f(P)\|$ est le vecteur normal unitaire).

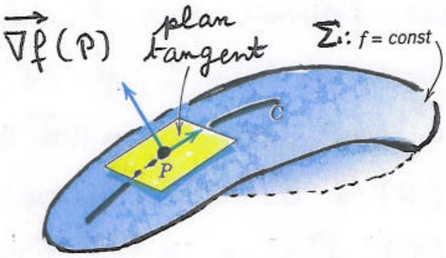
→ (3 variables). Si Σ est la surface de niveau $\{(x,y,z) \mid f(x,y,z) = c\}$, si $P \in \Sigma$, et si $\vec{\nabla} f(P) \neq \vec{0}$, $\vec{\nabla} f(P)$ est un vecteur normal à Σ en P.

(2 variables)



La droite passant par P et orthogonale à $\vec{\nabla} f(P)$ est la tangente à Γ en P.

(3 variables)



Le plan passant par P et orthogonal à $\vec{\nabla} f(P)$ est le plan tangent à Σ en P.

Exercice 12. Soit $f : (x,y) \longmapsto f(x,y)$ une fonction de 2 variables (de classe C^1), soit Σ la surface représentant son graphe, i.e.,

$$\Sigma = \{ (x, y, f(x,y)) \mid (x,y) \in \mathbb{R}^2 \}.$$

Vérifier que si $P = (x, y, f(x,y))$ est un point de Σ, le vecteur $(\vec{\nabla} f(P), -1)$ est normal à Σ en P
(Ind. Ecrire Σ comme une surface de niveau, $\{(x,y,z) \mid f(x,y) - z = 0\}$.)

3.2 La règle de FERMAT

Considérons une fonction f de 2 variables, nous la supposons de classe C^1.
On dit que $P = (a,b)$ est un minimiseur local de f (ou que f présente en (a,b) un minimum local) s'il existe $r > 0$ tel que

$$f(x,y) \geq f(a,b) \text{ dès que } \|(x,y) - (a,b)\| \leq r. \qquad (18)$$

Si l'inégalité au-dessus a lieu pour tout $(x,y) \in \mathbb{R}^2$ on parlera de $P = (a,b)$ comme d'un minimiseur global (ou absolu).
On définit un maximiseur local (resp. global) mutatis mutandis.

THÉORÈME 13. Règle de FERMAT (eh oui on est à Toulouse !).
Si $P = (a,b)$ est un minimiseur local de f, ou bien un maximiseur local de f, alors on a nécessairement

$$\vec{\nabla} f(P) = \vec{0}. \tag{19}$$

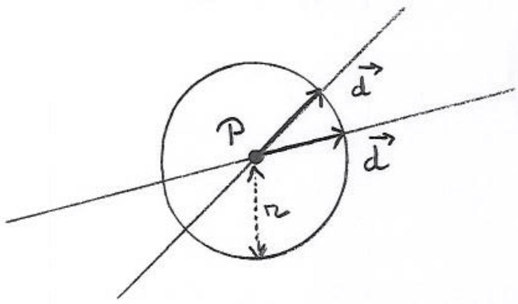
Démonstration. Soit $\vec{d}$ une direction unitaire de $\mathbb{R}^2$ et

$$\varphi : t \longmapsto \varphi(t) := f(P + t\vec{d}).$$

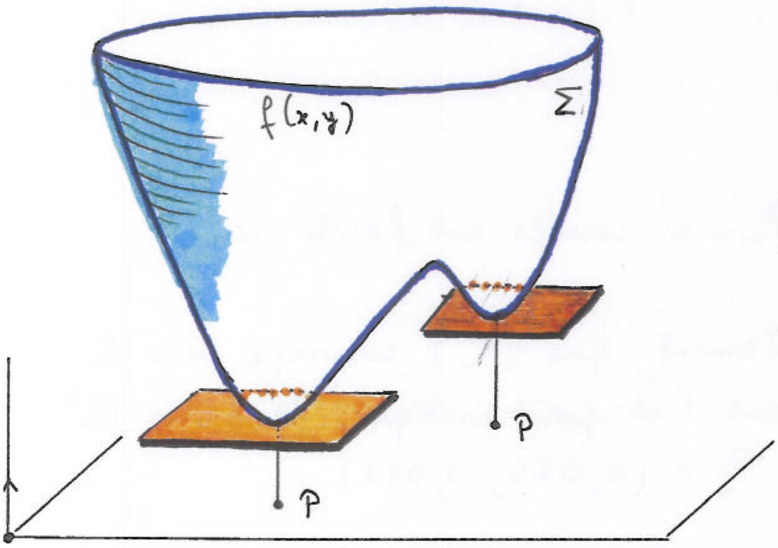
Alors, suite à la définition (18) de minimiseur local, on a :

$$\varphi(t) \geq \varphi(0) \text{ pour tout } t \in [-r, r]$$

(on est d'accord ?).



En conséquence, $\varphi'(0) = 0$.
Comme, $\varphi'(0) = \langle \vec{\nabla} f(P), \vec{d} \rangle$, il suffit de choisir $\vec{d} = \vec{\nabla} f(P) / \|\vec{\nabla} f(P)\|$ pour avoir $\|\vec{\nabla} f(P)\| = 0$, soit le résultat (19) annoncé. ∎



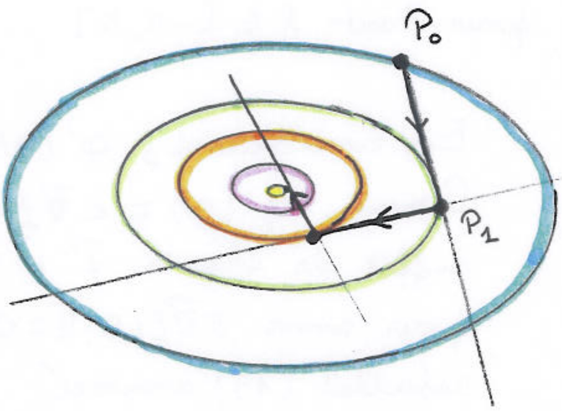
Les points $P = (a,b)$ en lesquels $\vec{\nabla} f(P) = \vec{0}$ sont appelés points stationnaires (ou critiques) de f. Pour de tels points P, les plans tangents au graphe Σ de f en $(P, f(P))$ sont horizontaux.

3.3. Un algorithme de minimisation : la méthode du gradient ou de plus profonde descente (de A. CAUCHY, 1847)

L'idée de l'algorithme est la suivante : à partir d'un point initial P_0, on construit récursivement une suite de points P_n, où P_{n+1} est déduit de P_n en se déplaçant le plus en profondeur possible dans la direction $-\vec{\nabla} f(P_n)$.

- Initialisation : P_0
- Passage de l'itéré $P_n = (x_n, y_n)$ à l'itéré $P_{n+1} = (x_{n+1}, y_{n+1})$:
  Soit $\varphi : t \geqslant 0 \longmapsto \varphi(t) := f[P_n - t\,\nabla f(P_n)]$; on choisit $t_n > 0$ minimisant φ, et on pose $P_{n+1} := P_n - t_n \nabla f(P_n)$.

Si tout se passe bien, la suite $(P_n)_{n\in\mathbb{N}}$ ainsi générée converge vers un point $\bar{P}$ minimiseur de f (ou au moins vers un point critique de f). C'est une convergence "en zig-zags", lente dans la pratique.



$P_0, P_1, \ldots, P_n, \ldots$

Exercice 14. Illustrer la méthode de plus profonde descente sur l'exemple suivant :

$$f(x,y) := x^2 + 3y^2$$
$$P_0 := (6,3)$$

Les calculs des t_n - et donc des P_n - successifs est facile ici.

1°) Déterminer $P_1, \ldots, P_6$.

2°) Vérifier, numériquement seulement, que (P_n) converge vers le point $\bar{P} = (0,0)$ unique minimiseur (et point critique) de f sur $\mathbb{R}^2$.

Rép. $P_1 = (3{,}484,\ -0{,}774), \ldots, P_6 = (0{,}065,\ 0{,}032)$.

3.4 Conditions d'optimalité d'ordre 2.

On se met dans le contexte du Théorème 9 (p. 13), et on se pose la question : si $\bar{P}$ est un point critique de f, comment décider (si possible) si $\bar{P}$ est un minimiseur local ou pas à l'aide de l'information du second ordre en $\bar{P}$, à savoir $\nabla^2 f(\bar{P})$?

Ce qui suit est la généralisation (aux fonctions de 2 variables) du résultat suivant sur les fonctions d'une variable (vu en L1) : si a est un minimiseur de f, alors $f'(a) = 0$ et $f''(a) \geqslant 0$.

THÉORÈME 15.

- **(CN)** Si $P=(a,b)$ est un minimiseur local de f, alors $\vec{\nabla} f(P)=\vec{0}$ et $\nabla^2 f(P) \in \mathcal{S}_2(\mathbb{R})$ est semi-définie positive.
- **(CS)** Si $\vec{\nabla} f(P)=\vec{0}$ et si $\nabla^2 f(P)$ est définie positive, alors P est un minimiseur local de f.

Démonstration. Tout le travail préparatoire a été fait avec le développement du second ordre (Théorème 9, p. 13).
(CN). A partir de P, regardons le comportement de f dans une direction $\vec{d} \in \mathbb{R}^2$ (que l'on prend unitaire, sans perte de généralité).
D'après le Théorème 9, on a :

$$f(P+t\vec{d}) = f(P) + t < \vec{\nabla} f(P), \vec{d} > + \frac{t^2}{2} < \nabla^2 f(P)\vec{d}, \vec{d} > + o(t^2)$$

(on est d'accord ?). (20)
Or $f(P+t\vec{d}) \geq f(P)$ pour t assez petit, disons $|t| \leq r$.
Comme $\vec{\nabla} f(P) = 0$, on divise par $t^2/2$ dans le développement (20) (c'est la raison de l'astuce du développement en t, à $\vec{d}$ fixée) :

$$\frac{f(P+t\vec{d}) - f(P)}{t^2/2} = < \nabla^2 f(P)\vec{d}, \vec{d} > + \varepsilon(t), \tag{21}$$

où $\varepsilon(t) \to 0$ quand $t \to 0$.
En passant à la limite, $t \to 0$, dans (21), on obtient

$$< \nabla^2 f(P)\vec{d}, \vec{d} > \geq 0.$$

Ceci étant démontré pour tout $\vec{d} \in \mathbb{R}^2$, on a bien la semi-définie positivité de $\nabla^2 f(P)$.
(CS). On part dans l'autre sens dans le développement du second ordre de f à partir de P.
Puisque $\nabla^2 f(P)$ est définie positive, il existe $\alpha > 0$ tel que $< \nabla^2 f(P)(h,k).(h,k) > \geq \alpha \|(h,k)\|^2$ $(= \alpha(h^2+k^2))$; on peut prendre pour α, par exemple, la plus petite valeur propre de $\nabla^2 f(P)$ (relire p. 12 et revoir l'Exercice 10).
D'après le développement du second ordre du Théorème 9 (p. 13), si $P=(x_0,y_0)$,

$$f(x_0+h, y_0+k) - f(x_0,y_0) = \frac{1}{2} < \nabla^2 f(P)(h,k), (h,k) > + (h^2+k^2)\ \varepsilon(h,k)$$

$$\geqslant \left[\frac{\alpha}{2} + \varepsilon(h,k)\right].(h^2+k^2)$$

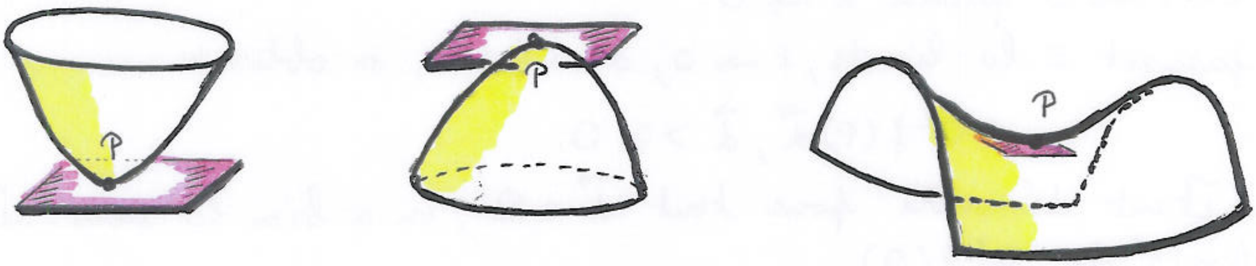
$$\geqslant \frac{\alpha}{4}(h^2+k^2) \text{ pour } \|(h,k)\| \text{ assez petit,}$$

disons $\|(h,k)\| \leq r$.

Ainsi, si $\|(h,k)\| \leq r$, $f(x_0+h, y_0+k) - f(x_0,y_0) \geqslant 0$. ■

Conséquence : classification des points critiques de f dans le cas où $\nabla^2 f(P)$ est inversible.

- Si $\nabla^2 f(P)$ est définie positive, f a un comportement autour du point critique P qui est illustré en (A) ci-dessous (curvette de fond P).
- Si $\nabla^2 f(P)$ est définie négative, f a un comportement autour de P qui est illustré en (B) ci-dessous (calotte de sommet P).
- Si $\nabla^2 f(P)$ est indéfinie, f a un comportement autour de P qui est illustré en (C) ci-dessous : on dit que P est un point-selle ou un col.



Vous voulez en savoir plus ? Lire "L'Optimisation" dans la collection "Que sais-je ?" (1996) : un excellent petit livre par un excellent auteur...