CAHIERS D'ENSEIGNEMENT

Extrema et Différentielles

Michel Karatchentzeff

VERSION 2.0 — MARS 2004
FASCICULE III
TAP 2

Ce fascicule correspond à une partie d'un cours donné de 1984 à 2004 à l'université Pierre et Marie Curie de Paris dans le cadre de la maîtrise de Physique et Applications.

Éditions successives

- version 1.0 année 1998
- version 1.1 année 1999
- version 1.2 année 2001
- version 1.3 année 2002
- version 2.0 année 2004

CHAPITRE 3

Extrema et Différentielles

Dans ce chapitre, les fonctions sont des fonctions à valeurs réelles d'une ou plusieurs variables réelles. La notion d'extremum associée à celles de continuité et de dérivabilité conduira naturellement à la notion de différentielle qui sera généralisée à d'autres entités que les fonctions usuelles. Par la suite, la différentielle se montrera être l'outil mathématique adapté au calcul des variations.

Commençons par rappeler quelques définitions.

- *Une fonction f à valeurs réelles admet un* maximum *en un point x_0 de son domaine de définition dans $\mathbb{R}^N$ s'il existe un voisinage de x_0 tel que pour tout point x de ce voisinage, on ait :*

$$f(x) - f(x_0) \leq 0$$

- *Une fonction d'une ou plusieurs variables réelles est* extrémale *en un point si, en ce point, elle possède un maximum ou un minimum, relatif ou absolu.*
- *Si la fonction possède plusieurs maxima, ceux (ou celui) pour lesquels la valeur de la fonction est la plus grande sont dits* maxima absolus *; les autres sont appelés* maxima relatifs.

Ces définitions s'étendent tout naturellement aux notions de minima relatifs ou absolus ; mais il faut bien remarquer que si l'on veut généraliser cette définition, on ne pourra le faire qu'en choisissant x_0 dans un espace topologique, l'espace d'arrivée devant, lui, posséder une relation d'ordre total, ce qui, pour un physicien, l'empêche d'être autre chose que $\mathbb{R}$.

3.1. Extrema des fonctions d'une variable réelle.

Dans le cas des fonctions d'une variable réelle, les définitions précédentes peuvent être affinées par les quelques lemmes suivants :

- *Si une fonction f admet un maximum (resp. un minimum) en un point x_0 appartenant à l'intérieur du domaine de définition de f, et si $f'(x_0)$ existe, alors :*

$$f'(x_0) = 0$$

Démonstration : par définition du maximum pour tout x suffisamment voisin de de x_0, $f(x) - f(x_0) \leq 0$, ce qui permet d'écrire :

$$\begin{array}{lcl} \text{si } x < x_0 & \qquad & \text{si } x > x_0 \\ \dfrac{f(x)-f(x_0)}{x-x_0} \geq 0 & & \dfrac{f(x)-f(x_0)}{x-x_0} \leq 0 \\ f'(x_0^-) \geq 0 & & f'(x_0^+) \leq 0 \end{array}$$

La dernière ligne s'obtient par passage à la limite de la précédente, puisque, f étant dérivable, les dérivées à gauche et à droite existent ; et comme de plus, ces dernières sont égales chacune à $f'(x_0)$

$$f'(x_0) \geq 0 \qquad\qquad f'(x_0) \leq 0$$

$$f'(x_0) = 0$$

• La réciproque est évidemment fausse : ce n'est pas parce que

$$f'(x_0) = 0$$

que f possède un extremum en ce point, comme le montre

$$f(x) = x^3$$

à l'origine.

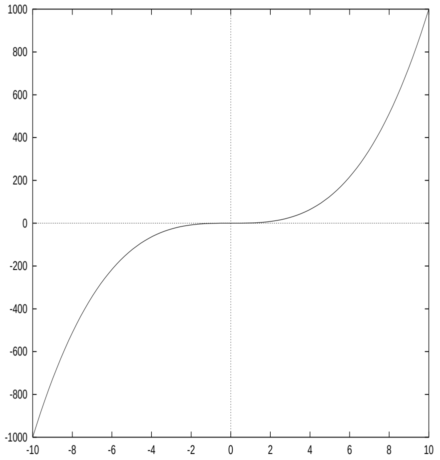


Figure 3.1. Exemple de fonction dont la dérivée s'annule en un point (l'origine) et qui n'est pas extrémale en ce point.

• Une fonction peut avoir un extremum sans qu'il y ait une dérivée en ce point comme le montrent les exemples des figures suivantes

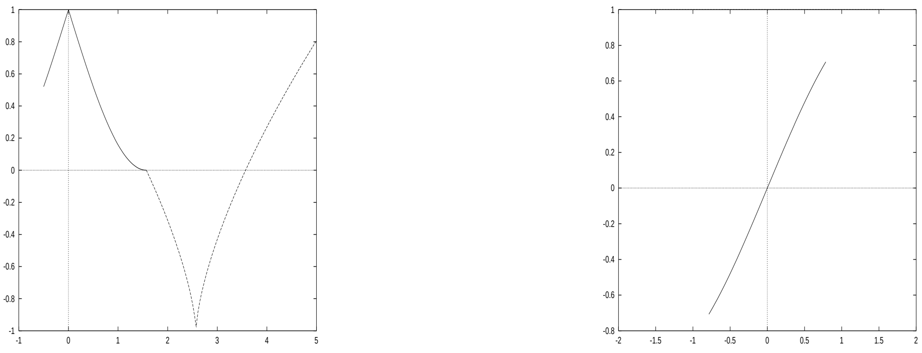


Figure 3.2. Exemples d'extrema en des points où la fonction n'est pas dérivable, ou dont la dérivée est infinie (dérivée discontinue ou infinie, bornes).

ou celui de la fonction bien “classique” (pour les mathématiciens) :

$$f(x) = \begin{cases} 1, & \text{si } x \notin \mathbb{Q} \\ 0, & \text{si } x \in \mathbb{Q}. \end{cases}$$

- Le comportement d'une fonction dans le voisinage d'un point x_0 où sa dérivée s'annule et où $f''(x_0)$ est différent de 0 peut être précisé. Mais auparavant, il n'est pas inutile de rappeler les théorèmes relatifs au développement d'une fonction.

Théorème de Taylor-Lagrange :

Soit f une fonction réelle d'une variable réelle, définie et dérivable jusqu'à l'ordre $n-1$ sur l'intervalle $[a,b]$;

si $f^{(n)}$ existe sur $]a,b[$, alors il existe un nombre réel c appartenant à $]a,b[$ tel que :

$$f(b) = f(a) + \frac{b-a}{1!}f'(a) + \cdots + \frac{(b-a)^{n-1}}{(n-1)!}f^{(n-1)}(a) + \frac{(b-a)^n}{n!}f^{(n)}(c)$$

Il n'y a aucune notion de convergence dans cette formule qui n'est qu'une généralisation de la formule de la moyenne.

Théorème de Taylor-Young :

Soit f une fonction réelle d'une variable réelle définie dans le voisinage d'un réel a ;

si f possède en a une dérivée n-ème finie,

alors il existe un nombre réel α positif et une fonction réelle φ_n tels que pour tout nombre réel h, avec $|h| < \alpha$, on ait :

$$f(a+h) = f(a) + \frac{h}{1!}f'(a) + \cdots + \frac{h^n}{n!}f^{(n)}(a) + \varphi_n(h)\frac{h^n}{n!}$$

où $\lim_{h\to 0} \varphi_n(h) = 0$.

Autrement dit $\varphi_n(h)\dfrac{h^n}{n!}$ est un $o(h^n)$ au sens de Landau.

C'est ce dernier théorème que nous allons appliquer avec les valeurs $a = x_0$, $n = 2$ et $f'(x_0) = 0$ pour étudier le comportement d'une fonction dans le voisinage d'un point où sa dérivée première s'annule sans que la dérivée seconde le fasse.

Avec ces conditions, le théorème de Taylor-Young affirme qu'il existe $\alpha > 0$ et φ_2 tels que pour tout h, $|h| < \alpha$:

$$f(x_0+h) = f(x_0) + hf'(x_0) + \frac{h^2}{2}f''(x_0) + \frac{h^2}{2}\varphi_2(h)$$

Soit :

$$f(x_0+h) - f(x_0) = \frac{h^2}{2}f''(x_0) + \frac{h^2}{2}\varphi_2(h)$$

Supposons maintenant que l'extremum en x_0 soit un maximum ; par définition, il existe donc un nombre réel β tel que, pour tout $|h| < \beta$, nous ayons :

$$f(x_0 + h) - f(x_0) < 0$$

et donc, si $\gamma = \inf(\alpha, \beta)$

$$\frac{h^2}{2} f''(x_0) + \frac{h^2}{2} \varphi_2(h) < 0$$

pour tout $|h| < \gamma$.

Comme $h^2 > 0$ et que $\varphi_2(h)$ tend vers 0 avec h, il s'ensuit[1] par passage à la limite que $f''(x_0) \leq 0$.

En conclusion, s'il y a un maximum en x_0, si la dérivée seconde existe, alors elle est négative en ce point et la réciproque de cette propriété est évidente.

L'étude du minimum se fait de façon analogue et est laissée en exercice. L'ensemble de ces résultats peut se résumer dans le tableau suivant où l'on a regroupé les différents comportements dans un voisinage de x_0 :

x		x_0			x_0	
$f''(x)$		$-$			$+$	
$f'(x)$	$\searrow$ $(+)$	0	$\searrow$ $(-)$	$\nearrow$ $(-)$	0	$\nearrow$ $(+)$
$f(x)$	$\nearrow$	Max	$\searrow$	$\searrow$	Min	$\nearrow$

Les propriétés démontrées ci-dessus ne sont évidemment pas exhaustives ; elles pourraient, par exemple, être complétées par des études du comportement d'une fonction f au voisinage d'un point x_0 où non seulement $f'(x_0) = 0$, mais encore $f''(x_0) = 0$ sans que la dérivée troisième ne s'annule, etc. Ces cas, qui relèvent de la "théorie des courbes", peuvent être traités sans difficulté et sont laissés au lecteur à titre d'exercice.

3.2. Les fonctions de plusieurs variables.

Une partie seulement des résultats précédents s'étend aux fonctions de plusieurs variables réelles.

---

[1] Supposons en effet que, contrairement au résultat annoncé, $f''(x_0) = k > 0$ et écrivons que $\varphi_2(h)$ tend vers 0 avec h :

$$\forall\ \varepsilon,\ \exists\ \beta,\ |h| < \beta \Rightarrow -\varepsilon < \varphi_2(h) < \varepsilon$$

Reportée dans la définition de maximum, l'inégalité ci-dessus permet d'écrire :

$$0 > f''(x_0) + \varphi_2(h) = k + \varphi_2(h) > k - \varepsilon$$

Comme ceci doit être vrai pour tout ε, il suffit de choisir, par exemple, $\varepsilon = k/2$ pour constater que nous aboutissons à une absurdité. La véracité de la proposition en découle.

1. *Pour qu'une fonction f définie sur un voisinage de a et possédant des dérivées partielles ait, par exemple, un maximum en a, il faut que toutes les dérivées partielles s'annulent en a.*

Démontrons cette propriété dans le cas $n = 3$ en posant $a = (a_1, a_2, a_3)$

- f maximum en a $\iff$ $\exists \alpha \in \mathbb{R}^{*+}$ tel que
  $|x_1 - a_1| < \alpha$, $|x_2 - a_2| < \alpha$, $|x_3 - a_3| < \alpha$,
  $\implies f(x) - f(a) < 0$ pour $x \neq a$
- Ce résultat, vrai en particulier pour $x_2 = a_2$ et $x_3 = a_3$, s'écrit dans ce cas

$$f(x_1, a_2, a_3) - f(a_1, a_2, a_3) < 0$$

et nous ramène au cas d'une fonction dérivable d'une seule variable ayant un maximum en a_1. D'après les résultats du paragraphe précédent, $f'_{x_1}(a_1, a_2, a_3) = 0$

- Il ne reste plus qu'à répéter ce raisonnement pour chacune des autres dérivées partielles pour achever la démonstration.

2. La propriété précédente est une condition nécessaire, mais elle n'est pas suffisante comme le montre le contre-exemple suivant :

Soit la fonction définie par $f(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_3^2$. À l'origine, on a toujours :

$$f'_{x_1}(0,0,0) = f'_{x_2}(0,0,0) = f'_{x_3}(0,0,0) = 0$$

Pour tout nombre ρ strictement positif, il sera donc possible de trouver :

- un point $(\rho, 0, 0)$ tel que $f(\rho, 0, 0) = \rho^2 > 0 = f(0,0,0)$;
- un point $(0, 0, \rho)$ tel que $f(0, 0, \rho) = -\rho^2 < 0 = f(0,0,0)$.

L'origine n'est donc pas un maximum puisque tout voisinage de 0 contient une sphère ouverte dans laquelle nous pouvons trouver des points x tels que $f(x) - f(0) > 0$ et d'autres point x tels que $f(x) - f(0) < 0$.

3. Les formules de Taylor :

- La formule de Taylor-Lagrange sera présentée dans le cas $N = 2$.

Soit donc f une application de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ définie dans un voisinage ouvert V d'un point u de $\mathbb{R}^2$.

Soit v un autre point de $\mathbb{R}^2$ choisi de façon que $u + v$ appartienne également à V.

Nous supposerons de plus que toutes les dérivées partielles de f existent jusqu'à l'ordre $n-1$ et sont continues dans V dans lequel les dérivées n-èmes sont simplement supposées exister.

Une base étant choisie, $u = (a, b)$ et $v = (h, k)$, il existe un nombre réel θ $(0 < \theta < 1)$ tel que

$$\begin{aligned} f(a+h, b+k) = {} & f(a,b) \\ & + [hf'_x(a,b) + kf'_y(a,b)] \\ & + \cdots \\ & + \frac{1}{(n-1)!}[hf'_x(a,b) + kf'_y(a,b)]^{[n-1]} \\ & + \frac{1}{n!}[hf'_x(a+\theta h, b+\theta k) + kf'_y(a+\theta h, b+\theta k)]^{[n]} \end{aligned}$$

Exemple $n = 2$:

$$\begin{aligned} f(a+h,b+k) &= f(a,b) + hf'_x(a,b) + kf'_y(a,b) \\ &\quad + \frac{1}{2}[h^2 f^{(2)}_{x^2}(a+\theta h, b+\theta k) + 2hk f^{(2)}_{xy}(a+\theta h, b+\theta k) + k^2 f^{(2)}_{y^2}(a+\theta h, b+\theta k)] \end{aligned}$$

Il ne s'agit pas d'un développement limité : l'existence des dérivées partielles d'ordre 2 en (a,b) n'est pas suffisante pour permettre de montrer que f est, par exemple différentiable s'il n'est pas tenu compte de l'hypothèse de continuité de f'_x et de f'_y.
Sans une hypothèse du genre f'' bornée (par exemple), il n'est pas possible de montrer que

$$\frac{1}{2}[h^2 f^{(2)}_{x^2} + 2hk f^{(2)}_{xy} + k^2 f^{(2)}_{y^2}] = \varepsilon(y)\|y\|$$

où $\varepsilon(y) \longrightarrow 0$ avec y. En effet, il est toujours possible d'écrire

$$\begin{aligned} &\frac{1}{2}[h^2 f^{(2)}_{x^2}(a+\theta h, b+\theta k) + 2hk f^{(2)}_{xy}(a+\theta h, b+\theta k) + k^2 f^{(2)}_{y^2}(a+\theta h, b+\theta k)] \\ &\quad = (|h|+|k|)\cdot \\ &\qquad \frac{1}{2}\frac{[h^2 f^{(2)}_{x^2}(a+\theta h, b+\theta k) + 2hk f^{(2)}_{xy}(a+\theta h, b+\theta k) + k^2 f^{(2)}_{y^2}(a+\theta h, b+\theta k)]}{(|h|+|k|)} \\ &\qquad \equiv \|y\|\varepsilon(y) \end{aligned}$$

et si les $f''_{..}$ sont toutes bornées par M

$$\varepsilon(y) \le \frac{1}{2}M\frac{h^2+2hk+k^2}{\|y\|} \le \frac{1}{2}M\frac{(h+k)^2}{\|y\|} \le \frac{1}{2}M\|y\|$$

quantité qui tend bien vers 0 avec y, et donc

$$f(x+y) = f(x) + [hf'_x(x) + kf'_y(x)] + \varepsilon(y)\|y\|$$

L'application entre crochets étant de plus linéaire et continue, la proposition est démontrée ; mais il n'y avait pas besoin de tout cela, puisque, comme nous l'avions démontré précédemment, les dérivées partielles premières étant continues, assurent à elles-seules l'existence de la différentielle. L'existence des dérivées partielles d'ordre 2 est donc superflue ; si elles existent le "reste" est un $\mathcal{O}(\|y\|^2)$ au lieu d'être un $o(\|y\|)$.

- Nous ne traiterons pas le cas de la formule de Taylor-Young à plusieurs variables qui suppose uniquement l'existence de la différentielle d'ordre n en x (et donc des différentielles précédentes et de leur continuité) car elle est assez lourde à exposer.

4. Le développement de Taylor-Lagrange permet d'étudier le comportement d'une fonction réelle de plusieurs variables au voisinage d'un extremum. Esquissons la méthode.

Plaçons nous dans les cas $N = 2$ et supposons que les dérivées partielles existent jusqu'à l'ordre 3 et que l'extremum a lieu au point (a, b), ce qui impose :

$$f'_x(a,b) = f'_y(a,b) = 0$$

Définissons :

$$f"_{x^2} = r \qquad f"_{xy} = s \qquad f"_{y^2} = t$$

ce qui permet de simplifier l'écriture de la formule de Taylor-Lagrange à l'ordre 3 en :

$$f(a+h,b+k)-f(a,b) = \frac{1}{2}(h^2r+2hks+k^2t)+\frac{1}{6}[hf'_x(a+\theta h,b+\theta k)+kf'_y(a+\theta h,b+\theta k)]^{[3]}$$

Le passage en coordonnées polaires $a + h = a + \rho\cos\alpha$, $b + k = b + \rho\sin\alpha$ transforme alors cette équation en :

$$\begin{aligned} f(a+h,b+k) - f(a,b) &= \frac{\rho^2}{2}(r\cos^2\alpha + 2s\sin\alpha\cos\alpha + t\sin^2\alpha) + \frac{\rho^3}{6}[\cdots]^{[3]} \\ &= \frac{\rho^2}{2}H + \frac{\rho^3}{6}K \end{aligned}$$

Lorsque ρ est petit et que le terme prépondérant est le terme en ρ^2, $f(a+h,b+k)-f(a,b)$ a le signe de

$$H = \cos^2\alpha[r + 2s\,\mathrm{tg}\,\alpha + t\,\mathrm{tg}^2\,\alpha]$$

Le principe de la discussion s'en déduit :
– si $s^2 - rt < 0$, le signe de H est le même que celui de r et celui de t ; c'est un minimum si r est positif, un maximum dans le cas contraire.
– si $s^2 - rt > 0$, dans tout cercle de centre (a, b), on trouve des points rendant H positif et d'autres rendant H négatif ; il n'y a donc pas d'extremum.
– le cas $s^2 - rt = 0$ nécessite une discussion légèrement plus délicate qui n'a pas sa place ici.

5. Mais une fonction de plusieurs variables n'a pas toujours des propriétés analogues à celle d'une fonction d'une variable, notamment quant à la continuité et à la dérivabilité. En voici deux exemples.

• *Une fonction continue de plusieurs variables est continue par rapport à chacune des variables, mais la réciproque est fausse.*

Considérons la fonction définie par :

$$f(x_1,x_2) = \begin{cases} \dfrac{x_1x_2}{x_1^2+x_2^2} & \text{si } (x_1,x_2) \neq (0,0) \\ 0 & \text{à l'origine} \end{cases}$$

Par construction, $f(0, x_2) = 0$ si $x_2 \neq 0$ et si $x_2 = 0$. C'est donc une fonction constante ; elle est donc continue. Il en est de même de $f(x_1, 0)$.

La fonction f est donc continue par rapport à chacune des variables (on dit qu'elle est séparément continue). Par contre, $f(x_1, x_1) = 1/2$ pour tout $x_1 \neq 0$.

Il n'y a donc pas continuité[2] si on arrive à l'origine en suivant la droite $x_1 = x_2$.

• Une autre propriété "manquante" est la suivante : si une fonction d'une variable est dérivable, elle est continue ; cette propriété devient fausse dans le cas des fonctions de plusieurs variables : dès que le nombre des variables dépasse 2, l'existence des dérivées partielles n'entraîne pas la continuite pour les fonctions de plusieurs variables.

Pour s'en convaincre, il suffit de prendre à nouveau :

$$f(x_1, x_2) = \begin{cases} \dfrac{x_1 x_2}{x_1^2 + x_2^2} & \text{si } (x_1, x_2) \neq (0,0) \\ 0 & \text{à l'origine} \end{cases}$$

f'_{x_1} et f'_{x_2} existent pour tout $(x_1, x_2) \neq (0,0)$: il suffit de les calculer pour le prouver.

À l'origine, par définition de la dérivée partielle :

$$f'_{x_1}(0,0) = \lim_{x_1 \to 0} \frac{f(x_1, 0) - f(0,0)}{x_1} = \lim_{x_1 \longrightarrow 0} \frac{0}{x_1} = 0$$

La démonstration étant analogue pour f'_{x_2}, les dérivées partielles existent donc en tout point de $\mathbb{R}^2$; la fonction, n'étant pas continue à l'origine, illustre bien la propriété "manquante" que nous avions énoncée.

Pour faire apparaître la continuité, il en faut plus ; et c'est ce que montre le théorème suivant :

Théorème : *Si une fonction f définie dans un voisinage de a admet des dérivées partielles* continues*, alors f est continue.*

Démonstration : Plaçons nous dans le cas de la dimension 2 ; de la continuité des dérivées partielles, il est aisé de déduire qu'il existe un domaine commun dans lequel les $\frac{\partial f}{\partial x_i}$ $(i = 1, 2)$ sont continues. Plus précisément :

Pour tout $\eta > 0$, il existe ρ tel que pour tout h, $|h| < \rho$, on a :

$$\left|\frac{\partial f}{\partial x_1}(a+h) - \frac{\partial f}{\partial x_1}(a)\right| < \eta \quad \text{et} \quad \left|\frac{\partial f}{\partial x_2}(a+h) - \frac{\partial f}{\partial x_2}(a)\right| < \eta$$

Cela posé, une fonction continue étant bornée sur un domaine borné, avec les mêmes hypothèses, nous pouvons écrire que :

pour tout h tel que $|h| < \rho$, il est possible de trouver deux nombres M et N tels que :

$$\left|\frac{\partial f}{\partial x_1}(a+h)\right| < M \quad \text{et} \quad \left|\frac{\partial f}{\partial x_2}(a+h)\right| < N$$

[2] En effet, $\epsilon = 1/4$ étant choisi, pour tout $x_1 \neq 0$,

$$f(x_1, x_1) - f(0,0) = 1/2 > 1/4$$

Il reste à montrer que, avec les hypothèses ci-dessus, pour tout ϵ, il existe un ρ tel que pour tout h, $|h| < \rho$ tel que :

$$|f(a+h) - f(a)| < \epsilon$$

ϵ étant donné, les dérivées partielles sont bornées et ρ s'en déduit. Il est alors toujours possible d'écrire que, quelque soient h_1 et h_2,

$$\begin{aligned} f(a_1+h_1, a_2+h_2) - f(a_1, a_2) = f(a_1+h_1, a_2+h_2) - f(a_1+h_1, a_2) \\ + f(a_1+h_1, a_2) - f(a_1, a_2) \end{aligned}$$

Comme les dérivées partielles existent et sont continues, il est possible d'appliquer deux fois, et de façon indépendante, le théorème des accroissements finis. Il existe donc θ_1 et θ_2 appartenant à l'intervalle $]0, 1[$ tels que :

$$f(a_1+h_1, a_2+h_2) - f(a_1, a_2) = h_2 \frac{\partial f}{\partial a_2}(a_1+h_1, a_2+\theta_2 h_2) + h_1 \frac{\partial f}{\partial a_1}(a_1+\theta_1 h_1, a_2)$$

Il ne reste plus qu'à remarquer que

$$\|(h_1, \theta_2 h_2)\|\|\| \leq \|h \qquad \|(\theta_1 h_1, h_2)\|\|\| \leq \|h$$

pour écrire :

$$|f(a_1+h_1, a_2+h_2) - f(a_1, a_2)| \leq |h_2|N + |h_1|M \leq (M+N)\|h\|$$

A ce niveau, de deux choses l'une :
ou $(M+N)\|h\| \leq \epsilon$, et la démonstration est achevé ;
ou il suffit de réduire la valeur de $\|h\|$ (c'est-à-dire de ρ) pour obtenir le même résultat.

Si nous examinons soigneusement la démonstration précédente, nous pouvons constater que l'accroissement de la fonction se décompose en deux partie dont l'une est linéaire et l'autre un $\text{o}(h, k)$; la partie linéaire s'appelle le gradient de la fonction ; c'est par construction une forme linéaire qui est également continue puisque l'espace est de dimension finie. Cette propriété se généralise à d'autres êtres mathématiques que les fonctions de variables réelles et conduit à la notion de différentielle. C'est cette notion de différentielle que nous allons développer, qui est la généralisation "naturelle" de la notion de dérivée pour les fonctions réelles d'une seule variable réelle.

3.3. La notion de différentielle.

DÉFINITION :

E et F sont des espaces vectoriels normés sur $\mathbb{R}$ (ou sur $\mathbb{C}$),
$B(a, r)$ est une boule ouverte de E centrée en a et de rayon r,
$f: B(a, r) \longrightarrow F$ est une application de E dans F ;

f est dite différentiable *en a s'il existe une application linéaire et continue l de E dans F telle que, pour tout h de E satisfaisant à $a+h \in B(a,r)$:*

$$f(a+h) - f(a) = l(h) + \epsilon(h)\|h\|$$

où $\varepsilon(h) \in F$ peut dépendre de a et où $\lim_{h \longrightarrow_E 0} \varepsilon(h) = 0$.

L'application l, quand elle existe, dépend en général de a ; c'est par définition la *différentielle* de f en a ; elle est notée $d(f(a))$. L'espace vectoriel des applications linéaires et continues de E dans F sera noté $\mathcal{L}(E,F)$.

Remarques :

- Comme il a été dit dans la définition, ε peut dépendre de a.
- Les normes sont *a priori* distinctes dans E et dans F ; c'est pour ne pas alourdir les symboles qu'elles ont été écrites de manière identique (il n'y a que si $E = F$ que le problème se pose).
- Le symbole $\lim_{h \longrightarrow_E 0} \varepsilon(h) = 0$ signifie que :
  a) l'égalité a lieu dans F ;
  b) $\forall \varepsilon, \exists \eta$ tel que pour $\|h\|_E < \eta$, on a $\|\varepsilon(h)\|_F < \varepsilon$ (Attention ! Dans cette définition que nous utiliserons plus loin le même symbole ε est associé à deux quantités ayant des signification différentes, mais qui sont corrélées).
- La définition de la différentielle ne fait pas intervenir la dimension des espaces vectoriels E et F ;
- Il est possible de remplacer les normes dans E ou dans F par des normes équivalentes : la différentielle reste inchangée.

Propriétés de la différentielle :

- *La différentielle est unique ;* cette propriété se démontre par l'absurde en supposant qu'il en existe une autre :

$$f(a+h) - f(a) = l_1(h) + \varepsilon_1(h)\|h\|$$

Par différence, nous obtenons l'égalité

$$l(h) - l_1(h) = \big(\varepsilon_1(h) - \varepsilon(h)\big)\|h\|$$

qui est vraie pour tout h de E satisfaisant à $\|h\| < r$; il est en particulier possible de remplacer h par λh où $|\lambda| < 1$

$$l(\lambda h) - l_1(\lambda h) = \big(\varepsilon_1(\lambda h) - \varepsilon(\lambda h)\big)\|\lambda h\|$$

soit, en appliquant la linéarité et en supposant $\lambda \neq 0$

$$l(h) - l_1(h) = \big(\varepsilon_1(\lambda h) - \varepsilon(\lambda h)\big)\frac{|\lambda|}{\lambda}\|h\|$$

Il ne reste plus qu'à prendre la norme des deux membres pour obtenir

$$\|l(h) - l_1(h)\| = \|(\varepsilon_1(\lambda h) - \varepsilon(\lambda h))\| \, \|h\|$$

Si maintenant, nous imposons à λ de tendre vers 0, λh tendra vers 0 ; donc par continuité de la multiplication scalaire et par définition de la différentielle, $\varepsilon(\lambda h)$ et $\varepsilon_1(\lambda h)$ tendront vers 0 dans F.

Pour tout h de E satisfaisant à $\|h\| < r$, nous aurons donc

$$\|l(h) - l_1(h)\| = 0 \quad \text{et donc dans la même boule} \qquad l(h) = l_1(h)$$

par définition de la norme ; ce qui entraîne l'égalité correspondante dans l'espace des applications linéaires et continues de E dans F :

$$l = l_1$$

(vrai dans une boule donc partout, car l et l_1 sont linéaires).

- *Si f est différentiable en a, f est continue en a*

Cela résulte des définitions ; puisque l appartient à $\mathcal{L}(E, F)$, l est en particulier continue à l'origine et donc $\lim_{h \longrightarrow 0} l(h) = 0$.

Puisque l'addition est continue dans F, il est possible de passer à la limite dans la formule de définition de la différentielle, et donc d'écrire :

$$\lim_{h \longrightarrow 0} \big(f(a + h) - f(a)\big) = 0 \quad \text{soit} \quad \lim_{h \longrightarrow 0} f(a + h) = f(a)$$

ce qui est la définition de la continuité en a. La réciproque de cette propriété est fausse ; nous donnerons un contre-exemple plus tard.

- *Si f est une application constante, f est différentiable et sa différentielle est l'application nulle*, car la définition s'écrit :

$$f(a + h) - f(a) = 0 = l(h) + \varepsilon(h)\|h\|$$

$l = 0$ et $\varepsilon = 0$ satisfont cette définition, ce qui achève la démonstration puisque la différentielle est unique.

- *f et g sont deux applications de E dans F ; λ et μ sont des nombres réels. Si f et g sont différentiables en a élément de E, $\lambda f + \mu g$ est différentiable en a.*

La démonstration de cette propriété est laissée en exercice.

- *Soient a et h deux éléments de E ; si f est une application de E dans F différentiable en a, et si g est l'application définie pour tout x de E par*

$$g(x) = f(x - h)$$

alors g est différentiable en $a + h$ et

$$d(g(a + h)) = d(f(a))$$

La démonstration de cette propriété est laissée en exercice.

- *Si f est une application linéaire et continue de E dans F ($f \in \mathcal{L}(E, F)$), f est différentiable en tout point a de E et*

$$d(f(a)) = f$$

(cette propriété s'étend au cas où f est une application affine).

La démonstration se déduit de l'unicité et de la remarque que pour tout h de E

$$f(a + h) - f(a) = f(h) = f(h) + 0$$

Différentielle d'une application composée :

Théorème : *E_1, E_2 et E_3 sont des espaces vectoriels normés ; $a \in E_1$.*
Soit $f_1: E_1 \longrightarrow E_2$ une application différentiable en a ;
soit $f_2: B(f_1(a), r) \subset E_2 \longrightarrow E_3$ une application définie dans la boule ouverte de centre $f_1(a)$ et de rayon r, différentiable en $f_1(a)$,
alors
$f_2 \circ f_1: E_1 \longrightarrow E_3$ est différentiable en a et

$$d(f_2 \circ f_1(a)) = d(f_2(f_1(a))) \circ d(f_1(a))$$

Démonstration : Pour $i = 1, 2, 3$, $\|x\|_i$ sera la norme de x dans E_i.

1. Par hypothèse, pour tout $\varepsilon_1 > 0$, il existe $\eta_1 > 0$ tel que

$$z \equiv f_1(a + y) - f_1(a) = l_1(y) + \varepsilon_1(y)\|y\|_1 \quad \text{avec}^* \quad \|\varepsilon_1(y)\|_2 < \varepsilon_1 \quad \text{si} \quad \|y\|_1 < \eta_1$$

Comme l_1 est continue, il existe M_1, la norme de l_1, tel que

$$\|l_1(y)\|_2 \leq M_1\|y\|_1 \quad \text{et donc} \quad \|z\|_2 \leq (M_1 + \varepsilon_1)\|y\|_1 \quad \text{si} \quad \|y\|_1 \leq \eta_1$$

2. Toujours par hypothèse,

$$f_2(f_1(a) + z) - f_2(f_1(a)) = l_2(z) + \varepsilon_2(z)\|z\|_2 \quad \text{avec} \quad \|\varepsilon_2(z)\|_3 < \varepsilon_2 \quad \text{si} \quad \|z\|_2 < \eta_2$$

3. Si $\|y\|_1 < \inf(\eta_1, \frac{\eta_2}{M_1+\varepsilon_1}) \equiv \eta$, alors $\|z\|_2 < \eta_2$; en effet, de deux choses l'une :

– ou $\eta_1 = \inf(\eta_1, \dfrac{\eta_2}{M_1 + \varepsilon_1})$ et $\|y\|_1 < \eta_1$,
alors $\eta_1 < \dfrac{\eta_2}{M_1 + \varepsilon_1}$ et $M_1 + \varepsilon_1 < \dfrac{\eta_2}{\eta_1}$;

---

* Attention : ne pas confondre $\varepsilon_1(y)$, la valeur de la fonction, avec le nombre ε_1.

d'où d'après **1** :
$\|z\|_2 \leq (M_1+\varepsilon_1)\|y\|_1 < \dfrac{\eta_2}{\eta_1}\cdot\eta_1 = \eta_2$.
– ou $\dfrac{\eta_2}{M_1+\varepsilon_1} < \inf(\eta_1, \dfrac{\eta_2}{M_1+\varepsilon_1})$ et $\|y\|_1 < \dfrac{\eta_2}{M_1+\varepsilon_1}$,
donc $\|z\|_2 \leq (M_1+\varepsilon_1)\cdot\dfrac{\eta_2}{M_1+\varepsilon_1} = \eta_2$

4. Il ne reste plus qu'à remplacer z dans **2.** par sa valeur en fonction de y :

$$l_2(z) = l_2[l_1(y) + \varepsilon_1(y)\|y\|_1] = l_2(l_1(y)) + l_2(\varepsilon_1(y))\cdot\|y\|_1$$

d'après la linéarité de l_2 ; tandis que la continuité permet d'écrire

$$\Big\|l_2[\varepsilon_1(y)]\|y\|_{1_3}\Big\| \leq M_2\Big\|\varepsilon_1(y)\|y\|_{1_2}\Big\| < M_2\varepsilon_2\|y\|_1$$

Comme par définition $f_2(f_1(a)+z) = f_2(f_1(a+y))$, nous obtenons :

$$f_2(f_1(a+y)) - f_2(f_1(a)) = l_2(l_1(y)) + \varepsilon_3(y)\|y\|_1$$

où

$$\|\varepsilon_3(y)\|_3 < M_2\varepsilon_1 + \varepsilon_2(M_1+\varepsilon_1) \quad \text{si} \quad \|y\|_1 < \eta$$

C'est bien ce que nous voulions démontrer.

Cas des fonctions numériques sur un espace vectoriel de dimension finie :

- Dans le cas où l'application f est à valeurs réelles ($f\colon E \longrightarrow \mathbb{R}$), la différentielle est une forme linéaire et continue sur E ; c'est donc un élément du dual topologique de E.
- Soit N la dimension de E que nous supposons pour l'instant finie, et soit donc $e_1, \dots, e_N$ une base de E. Soit a un élément de E. Nous poserons

$$a = a_1e_1 + \cdots + a_Ne_N$$

et, puisque, la dimension est finie, toutes les normes sont équivalentes, nous choisirons

$$\|a\| = |a_1| + \cdots + |a_N|$$

Soit maintenant une fonction $f\colon E \longrightarrow \mathbb{R}$ différentiable en a. Par définition,

$$f(a+h) - f(a) = l(h) + \varepsilon(h)\|h\| \quad \text{avec} \quad \lim_{h\longrightarrow 0}\varepsilon(h) = 0$$

i étant donné, pour l'élément particulier $h = h_ie_i$, la définition ci-dessus s'écrit :

$$\begin{aligned} f(a+h_ie_i) - f(a) &= l(h_ie_i) + \varepsilon(h_ie_i)\|h_ie_i\| \\ &= h_il(e_i) + |h_i|\varepsilon(h_ie_i)\|e_i\| \end{aligned}$$

Si $h_i \neq 0$, la dernière égalité peut se transformer en

$$\frac{f(a + h_i e_i) - f(a)}{h_i} = l(e_i) + \frac{|h_i|}{h_i}\varepsilon(h_i e_i)\|e_i\|$$

Si $h_i \longrightarrow 0$, le membre de droite a pour limite $l(e_i)$. Donc le membre de gauche, qui converge, est par définition la dérivée par rapport à la variable x_i en a de la fonction $f(x_1, \ldots, x_N)$ de la seule variable x_i, les autres étant fixées.
C'est ce qu'on appelle la dérivée partielle au point a. Elle est notée selon les usages

$$\frac{\partial f}{\partial x_i}(a) \quad \text{ou} \quad f'_{x_i}(a) \quad \text{ou tout simplement} \quad \partial_{x_i} f(a)$$

En utilisant simplement la linéarité de l, nous pouvons écrire pour h quelconque :

$$f(a + h) - f(a) = h_1 l(e_1) + \cdots + h_N l(e_N) + \varepsilon(h)\|h\|$$

et donc obtenir la formule

$$d(f(a))(h) = l(h) = h_1 \frac{\partial f}{\partial x_1}(a) + \cdots + h_N \frac{\partial f}{\partial x_N}(a)$$

ce qui donne pour les fonctions numériques sur des espaces vectoriels de dimension finie à la fois une propriété (*si f est différentiable en a, f possède toutes ses dérivées partielles en a*), et un moyen de calculer explicitement leur différentielle.

Notations et définitions :

Choisissons maintenant la fonction particulière définie pour tout x de E par

$$f(x) = x_i$$

C'est la projection de x sur le i-ème vecteur de base ; c'est donc une forme linéaire et continue puisque les espaces en jeu sont de dimension finie. Tenant compte de

$$\frac{\partial f}{\partial x_k}(a) = 0 \quad \text{si} \quad k \neq i \quad \text{et} \quad \frac{\partial f}{\partial x_i}(a) = 1$$

la formule ci-dessus s'écrit en remplaçant $f(x)$ par x_i

$$d(f(a))h = h_i = dx_i h$$

soit encore

$$df(a) = dx_i$$

ce qui permet de récrire la formule générale sous la forme :

$$df(a) = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(a)\, dx_i$$

l'égalité ayant lieu maintenant dans $\mathcal{L}(E, F)$.

Avec les conditions ci-dessus sur les espaces associés à f, la différentielle de f s'appelle également *gradient*, ce qui s'écrit :

$$df(a) = \operatorname{grad} f(a)$$

L'égalité a évidemment toujours lieu dans $\mathcal{L}(E, F)$; l'intérêt principal de cette formule est qu'elle est indépendante de la notion de base.

Remarque :

À partir du moment où l'on sait calculer sous certaines conditions une différentielle, il est aisé de montrer qu'une fonction continue n'est pas nécessairement différentiable. La fonction

$$f(x_1, x_2) = \begin{cases} \dfrac{x_1 x_2}{\sqrt{x_1^2 + x_2^2}} & \text{si } (x_1, x_2) \neq (0,0) \\ 0 & \text{à l'origine} \end{cases}$$

est continue à l'origine.

Cette fonction, identiquement nulle pour $x = 0$ et pour $y = 0$, a donc des dérivées partielles nulles à l'origine.

D'après l'unicité de la différentielle et les résultats du paragraphe précédent, si cette fonction était différentiable, sa différentielle serait nulle à l'origine. Donc dans un voisinage de l'origine :

$$f(0 + h_1, 0 + h_2) - f(0,0) = f(h_1, h_2) = 0 + \frac{h_1 h_2}{\sqrt{h_1^2 + h_2^2}}$$

Le dernier terme s'interprête donc comme le reste qui doit être un $\mathcal{o}(h)$. Or si nous choisissons $h_2 = m h_1$, cette quantité s'écrit

$$\frac{|h_1| m}{\sqrt{1 + m^2}}$$

et, divisée par $\|h\|$, elle ne tend pas vers 0. Par suite, pour $m \neq 0$, le coefficient de $\|h\|$ ne tend pas vers 0 avec h. La fonction n'est donc pas différentiable.

conditions suffisantes de différentiabilité :

1. Pour $N = 1$, il y a équivalence entre dérivabilité et différentiabilité en ce sens que l'existence de l'une de ces notions entraîne celle de l'autre.

La démonstration de cette propriété commence par la remarque que :

$$f(a + h) - f(a) = f((a_1 + h_1)e_1) - f(a_1 e_1)$$

Supposons maintenant l'existence de la dérivée qui se traduit par

$$\frac{f((a_1 + h_1)e_1) - f(a_1 e_1)}{h_1} = \frac{\partial f}{\partial x_1}(a) + \varepsilon_1(h_1) \quad \text{avec} \quad \lim_{h_1 \longrightarrow 0} \varepsilon_1(h_1) = 0$$

Il est donc possible d'identifier cette forme avec la définition de la différentielle :

$$f(a+h) - f(a) = l(h) + \varepsilon(h)\|h\|$$

en posant

$$h = h_1 e_1 \quad , \quad \|h\| = |h| \quad , \quad l(h) = h_1 \frac{\partial f}{\partial x_1}(a) \quad \text{et} \quad \varepsilon(h) = \frac{h_1}{|h_1|}\varepsilon_1(h_1)$$

Avec ces définitions, $\lim_{h\longrightarrow 0} \varepsilon(h) = 0$, et f est donc différentiable en a.

La démonstration de la réciproque de cette propriété est évidente et est laissée au lecteur.

2. Si $N > 1$, l'existence des dérivées partielles n'est pas une condition suffisante de différentiabilité.

Nous avons vu par exemple que la fonction

$$f(x_1, x_2) = \begin{cases} \dfrac{x_1 x_2}{x_1^2 + x_2^2} & \text{si } (x_1, x_2) \neq (0,0) \\ 0 & \text{à l'origine} \end{cases}$$

admettait des dérivées partielles et qu'en particulier $f'_{x_1}(0) = f'_{x_2}(0) = 0$.

Donc si la différentielle existait à l'origine, elle serait nulle et

$$f(h) - f(0) = f(h) = \frac{h_1 h_2}{h_1^2 + h_2^2} \equiv \varepsilon(h)\|h\|$$

Comme $\|h\| = |h_1| + |h_2|$, $\varepsilon(h)$ devrait avoir pour forme :

$$\varepsilon(h) = \frac{h_1 h_2}{h_1^2 + h_2^2} \, \frac{1}{|h_1| + |h_2|}$$

Or pour $h_2 = h_1$, $\varepsilon(h) = 1/(4|hh_1|)$ et $f(h) = 1/4$.

Donc si h_1 tend vers 0, $f(h)$ tend vers une constante non nulle alors que $\varepsilon(h)$ tend vers l'infini. f n'est donc pas différentiable.

3. Pour $N > 1$, l'existence et la continuité de $\partial x_i f(a+h)$ pour $\|h\| < r$ et $i = 1, \ldots, N$ sont des conditions suffisantes pour que f soit différentiable en a lorsque E est un espace vectoriel réel.

Démonstration dans le cas $N = 4$. Posons :

$$\begin{aligned} y_1 &= h_1 e_1 + h_2 e_2 + h_3 e_3 + h_4 e_4 \\ y_2 &= \phantom{h_1 e_1 + {}} h_2 e_2 + h_3 e_3 + h_4 e_4 \\ y_3 &= \phantom{h_1 e_1 + h_2 e_2 + {}} h_3 e_3 + h_4 e_4 \\ y_4 &= \phantom{h_1 e_1 + h_2 e_2 + h_3 e_3 + {}} h_4 e_4 \end{aligned}$$

À l'aide des relations précédentes, l'accroissement de f peut s'écrire

$$\begin{aligned} f(a+h) - f(a) = \quad & f(a+h) - f(a+y_2) \\ & + f(a+y_2) - f(a+y_3) \\ & + f(a+y_3) - f(a+y_4) \\ & + f(a+y_4) - f(a) \end{aligned}$$

Les quatre fonctions de la variable réelle t définies par

$$g_i(t) \equiv f(a + y_{i+1} + te_i) \qquad (i = 0, \ldots, 3)$$

ont alors les propriétés suivantes :

- elles sont définies si
$$|t| + |h_{i+1}| + \cdots + |h_N| < r$$
- et donc, puisque
$$|h_1| + \cdots + |h_N| < r$$
elles sont définies si $|t| < h_i$
- elles sont donc continues et dérivables sur le même intervalle.

Le théorème des accroissements finis s'écrit pour chacune des fonctions

$$\begin{aligned} f(a+y_i) - f(a+y_{i+1}) &= g_i(h_i) - g_i(0) = h_i g_i'(\theta_i h_i) \\ &= h_i \frac{\partial f}{\partial x_i}(a + y_{i+1} + \theta_i h_i e_i) \quad \text{où} \quad 0 < \theta_i < 1 \end{aligned}$$

Or

$$\|y_{i+1} + \theta_i h_i e_i\| = \theta_i |h_i| + |h_{i+1}| + \cdots + |h_N| < r$$

et comme $\partial_{x_i} f$ est continue en a

$$\frac{\partial f}{\partial x_i}(a + y_{i+1} + \theta_i h_i e_i) - \frac{\partial f}{\partial x_i}(a) = \varepsilon_i(h)$$

et $|\varepsilon_i(h)| < \varepsilon$ si r est choisi assez petit pour que

$$\|h\| < r \Longrightarrow \left| \frac{\partial f}{\partial x_i}(a + y_{i+1} + \theta_i h_i e_i) - \frac{\partial f}{\partial x_i}(a) \right| = \left| \varepsilon_i(h) \right| < \varepsilon$$

Dans ces conditions

$$f(a+h) - f(a) = h_1 \frac{\partial f}{\partial x_1}(a) + \cdots + \frac{\partial f}{\partial x_N}(a) + \varepsilon(h)\|h\|$$

où $\varepsilon(h)\|h\| = h_1 \varepsilon_1(h) + \cdots + h_N \varepsilon_N(h)$ Donc $|\varepsilon(h)| \|h\| \leq \varepsilon(|h_1| + \cdots + |h_N|) = \varepsilon \|h\|$
En définitive $|\varepsilon(h)| < \varepsilon$ si $\|h\| < r$ et f est différentiable en a.

Remarque :

Toutes les différentielles ne s'écrivent pas commes celles des fonctions réelles de variables réelles. En voici un exemple important :

Soit $E = \mathbb{R}^N$ et soit ψ une forme bilinéaire symétrique définie sur $E \times E$. Supposons de plus qu'il existe un nombre réel M tel que pour tout x et y appartenant à E, $|\psi(x, y)| \leq M\|x\|\|y\|$ (ce qui revient à dire que ψ est continue).

Alors la forme quadratique définie pour tout x de E par $\phi(x) \equiv \psi(x, x)$ est différentiable en tout a de E. En effet

$$\begin{aligned}\phi(a+h) - \phi(a) &= \psi(a+h, a+h) - \psi(a, a) = \psi(a, a) + 2\psi(a, h) + \psi(h, h) - \psi(a, a) \\ &= 2\psi(a, h) + \psi(h, h)\end{aligned}$$

$\psi(a, h)$ est par hypothèse la valeur d'une forme linéaire et continue. Par ailleurs, comme $|\psi(h, h)| = |\phi(h)| \leq M\|h\|^2$, il est toujours possible d'écrire

$$\phi(y) = \varepsilon(y)\|y\| \quad \text{avec} \quad \varepsilon(y) \leq \|y\| < \varepsilon \quad \text{si} \quad \|y\| < \frac{\varepsilon}{M}$$

ϕ est donc différentiable et

$$d(\phi(a))(y) = 2\psi(a, y)$$

Dans ce cas, la différentielle est définie par ses valeurs ; contrairement aux fonctions de plusieurs variables réelles,il n'est pas possible, avec les notations employées, d'expliciter la définition dans l'espace des formes linéaires et continues sur E.

Un autre exemple sera donné par la suite avec la différentielle de l'intégrale d'action.

Théorème :

Soit f une fonction réelle d'une ou plusieurs variables réelles définie sur un voisinage d'un point a,
Si f est différentiable en a et si f admet un extremum en a,
alors $d(f(a))$ est la forme nulle.

Démonstration :

f extremum en a $\iff$ $f(a+h) - f(a)$ garde un signe constant si $\|h\| < r$ avec $h \neq 0$

f différentiable en a $\iff$ $f(a+h) - f(a) = l(h) + \varepsilon(h)\|h\|$ où $|\varepsilon(h)| < \varepsilon$ si $\|h\| < \eta$

en posant pour simplifier l'écriture $l = d(f(a))$.

Soit un élément h_0 de E vérifiant $\|h_0\| < r$ avec $h_0 \neq 0$. Posons $l(h_0) = \delta$. Pour tout réel λ, nous aurons :

$$l(\lambda h_0) + \varepsilon(\lambda h_0)\|\lambda h_0\| = \lambda\delta + |\lambda|\varepsilon(\lambda h_0)\|h_0\|$$

et si λ est choisi assez petit pour que

$$\|\lambda h_0\| < \eta \quad \text{entraîne} \quad |\varepsilon(\lambda h_0)| < \frac{\delta}{2\|h_0\|}$$

alors $f(a + \lambda h_0) - f(a)$ aura le signe de $\lambda\delta$.

Si λ est changé en $-\lambda$, le signe de $f(a+\lambda h_0) - f(a)$ ne doit pas changer, ce qui impose :

$$\forall \lambda, \quad \|h\| < r \Longrightarrow l(h) = 0$$

et comme l est continue, l est nulle.

Il est à noter que le raisonnement ci-dessus est valable si E est un espace vectoriel normé quelconque.

3.4. Le problème des extrema liés.

Introduction :

Ce problème apparaît lorsqu'on se propose de trouver les valeurs de $x_i, i = 1, \ldots, N+2$ pour lesquelles, par exemple, la fonction

$$f(x_1, \ldots, x_N, x_{N+1}, x_{N+2})$$

est extrêmale sachant que

$$\begin{aligned} \phi_1(x_1, \ldots, x_N, x_{N+1}, x_{N+2}) = 0 \\ \phi_2(x_1, \ldots, x_N, x_{N+1}, x_{N+2}) = 0 \end{aligned}$$

L'effet des deux relations est de diminuer le nombre des variables indépendantes et les extrema de f sont recherchés, non dans $\mathbb{R}^{N+2}$, mais sur la variété définie par les deux relations ci-dessus.

Supposons pour cela que les équations $\phi_1 = 0$ et $\phi_2 = 0$ soient résolubles par rapport à x_{N+1} et x_{N+2}, c'est-à-dire que, *localement*, le déterminant fonctionnel

$$\frac{D(\phi_1, \phi_2)}{D(x_{N+1}, x_{N+2})} \neq 0$$

Il est alors possible *localement* de calculer x_{N+1} et x_{N+2} en fonction des autres variables :

$$\begin{aligned} x_{N+1} &= x_{N+1}(x_1, \ldots, x_N) \\ x_{N+2} &= x_{N+2}(x_1, \ldots, x_N) \end{aligned}$$

Il ne reste plus qu'à reporter ces valeurs dans f qui ne dépendra plus que de N variables ; il suffira ensuite d'appliquer les méthodes classiques de recherche des extrema.

Exemple 1 :

Rechercher les extrema de

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

sachant que $x^2 + y^2 = 1$.

Il faut *dès le départ* tenir compte de la condition $x^2 + y^2 = 1$. En effet, si la courbe définie par $f(x, y)$ est considérée seule, son minimum est nul puisqu'elle est somme de deux quantités positives et elle n'a pas de maximum.

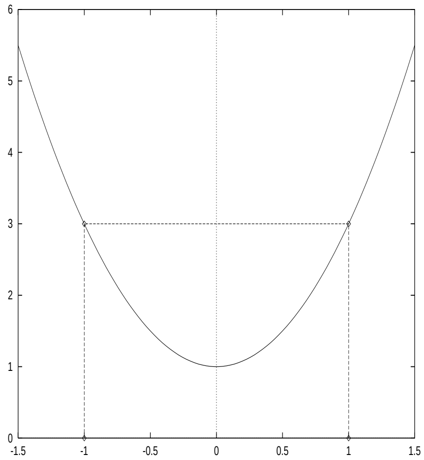


Figure 3.3. La parabole $2x^2 + 1$ et sa partie "utile".

Par contre, la condition $x^2 + y^2 = 1$ s'écrit $y = \pm\sqrt{1 - x^2}$ avec $|x| \leq 1$, ce qui, reporté dans f fournit

$$3x^2 + (1 - x^2) = 2x^2 + 1$$

Cette courbe est une parabole centrée sur $x = 0$ et de minimum 1.

Le minimum de f sachant que $x^2 + y^2 = 1$ est atteint pour $x = 0$ et donc $y = \pm 1$ d'après la condition.

Le maximum est atteint aux bornes du domaine en $x = \pm 1$ et donc $y = 0$ d'après la condition de contrainte.

Mais à part des cas aussi élémentaires, cette méthode ne peut s'appliquer facilement car il n'est pas toujours possible de formuler la dépendance explicite d'une variable en fonction des autres sans alourdir considérablement les calculs.

Exemple 2 :

Trouver les côtés d'un triangle rectangle dont le périmètre soit minimal pour une surface S *donnée.*

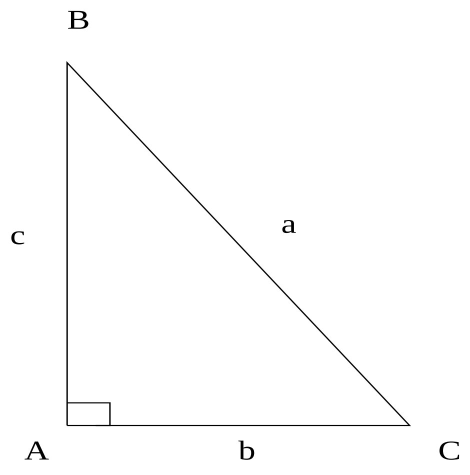


Figure 3.4. Construction d'un triangle.

Notations :

a désigne la longueur de l'hypothénuse ;

b et c celles des deux autres côtés.

Équations du problème :

minimiser

$$a + b + c$$

sachant que

$$bc = 2S \quad \text{et} \quad a^2 = b^2 + c^2$$

Méthode lourde[3] :

b et c sont extraits des deux dernières équations et reportés dans $a+b+c$ qui n'est alors plus fonction que de a et dont il ne reste plus qu'à chercher le minimum. Soit :

$$c = \frac{2S}{b} \qquad \longrightarrow \qquad b^4 - b^2a^2 + 4S^2 = 0 \qquad \longrightarrow \qquad \begin{cases} b_1^2 = \dfrac{a^2 + \sqrt{a^4 - 16S^2}}{2} \\ b_2^2 = \dfrac{a^2 - \sqrt{a^4 - 16S^2}}{2} \end{cases}$$

Pour que le problème ait une solution, il faut nécessairement que $a^4 - 4S^2 \geq 0$ et comme $b_1^2 + b_2^2 = a^2$, il est toujours possible de supposer que $b_1 = b$ et $b_2 = c$. Dans ces conditions, la fonction à minimiser s'écrit :

$$f(a) = a + \sqrt{\frac{a^2 + \sqrt{a^4 - 16S^2}}{2}} + \sqrt{\frac{a^2 - \sqrt{a^4 - 16S^2}}{2}}$$

La méthode de calcul utilisant directement la dérivée de cette fonction est assez lourde car elle fait intervenir des quantités du type

$$(\sqrt{u})' = \Bigg(\sqrt{\frac{a^2 + \sqrt{a^4 - 16S^2}}{2}}\Bigg)' = \frac{1}{\sqrt{u}}\Big(\frac{a^2 + \sqrt{a^4 - 16S^2}}{2}\Big)'$$

[3] Il est évident que, sans avoir besoin de faire de violents efforts intellectuels, le lecteur pourra trouver des solutions algébriques beaucoup plus simples et plus élégantes que celle qui est présentée ici. Il pourra alors constater que ces solutions dépendent explictement des conditions du problèmes et ne peuvent être appliquées qu'à ce problème. La méthode proposée ici est qualifée de "lourde" parce qu'elle est systématique et générale : on calcule à partir de chaque contrainte l'une des variables en fonction des autres ; on reporte dans l'expression à minimiser et on est ramené au calcul du minimum d'une fonction de plusieurs variables indépendantes.

Il est plus simple de remarquer que

$$\big[f(a)-a\big]^2 = a^2+4S \quad \text{soit} \quad f(a) = a+\sqrt{a^2+4S}$$

puisque $f(a)-a=b+c$ est nécessairement positif.

La fonction définie par $f(a)$ est donc croissante comme somme de deux fonctions croissantes (ce qui ne pouvait être affirmé pour la somme des trois fonctions constituant initialement f, puisque la troisième est visiblement décroissante).

Le minimum est donc atteint pour la plus petite valeur de l'intervalle de définition de f, c'est-à-dire pour $a^2-4S=0$, soit $a=2\sqrt{S}$. Comme cette valeur est celle qui annule le discriminant de l'équation définissant b (ou c) : $b^4-b^2a^2+4S=0$, la solution du problème est donnée par :

$$b=c=\sqrt{2S}$$

Exemple 3 :

Trouver les extrema de $f(x,y)=3x^2+xy+y^2$ *sachant que* $x^2+y^2=1$.

Dans ce cas, les diverses représentations des fonctions doivent être explicitées :
$y=\pm\sqrt{1-x^2}$ doit être reporté dans la fonction initiale, ce qui conduit à considérer :

$$g(x)\equiv f(x,\pm\sqrt{1-x^2}) = 2x^2+1\pm x\sqrt{1-x^2}$$

dont la dérivée s'écrit :

$$g'(x)=\frac{4x\sqrt{1-x^2}\pm(1-2x^2)}{\sqrt{1-x^2}}$$

et dont les zéros, solutions de l'équation :

$$20x^4-20x^2+1=0 \quad \text{ont pour valeurs} \quad x^2=\frac{5\pm2\sqrt{5}}{10}$$

Le plan des étapes ultérieures :

Il est facilement concevable que ce genre de technique conduit à des calculs lourds. Cependant, il est possible de montrer que, dans la mesure où une fonction est inversible, nombre de calculs sont inutiles et qu'il existe un moyen systématique d'effectuer les calculs nécessaires.

Dans ce qui suit, nous allons donc

- faire l'étude des cas où une fonction peut être définie implicitement à partir d'une relation donnée, car si ce n'était pas possible, il ne pourrait y avoir de solution : ce sera le "théorème des fonctions implicites." ;
- en déduire une méthode de résolution du problème posé : ce sera la "méthode des multiplicateurs de Lagrange".

3.5. Le théorème des fonctions implicites.

Position du problème :

L'idée directrice de ce théorème est de trouver les conditions sous lesquelles il est possible d'exprimer l'une des variables (z par exemple) liée à d'autres (x, y par exemple) par une relation donnée

$$\varphi(x, y, z) = 0 \tag{1}$$

en fonction de ces deux autres variables.

Mais comme il existe des fonctions qui n'ont pas de racines ($x^2 + y^2 + z^2 + 1 = 0$ par exemple), il n'est pas possible de songer, sans ajouter d'autres hypothèses, trouver un z qui pour x et y fixés serait solution de (1).

Mais si, par contre, on sait que pour $x = x_0$ et $y = y_0$ il existe un $z = z_0$ satisfaisant à (1), on peut concevoir que moyennant quelques hypothèses "mathématiques" (continuité et dérivabilité), on puisse trouver une solution $z = z(x, y)$ satisfaisant à (1) tant qu'on reste peu éloigné de (x_0, y_0) et qui prenne la valeur z_0 en (x_0, y_0).

On dit alors qu'une telle fonction est définie "implicitement" par (1), ce qui donne le nom du théorème qui précise ses conditions d'existence.

Énoncé :

Soit l'équation

$$\varphi(x, y, z) = 0 \tag{1}$$

Si l'équation (1) est vérifiée en (x_0, y_0, z_0),
si φ *et* φ'_z *sont continues au voisinage de* (x_0, y_0, z_0),
si $\varphi'_z(x_0, y_0, z_0) \neq 0$,
alors l'équation (1) admet pour des valeurs voisines de (x_0, y_0) *une solution et une seule voisine de* z_0. *Cette solution qui sera notée* $z(x, y)$ *est une fonction continue du couple* (x, y) *; elle prend la valeur* z_0 *au point* (x_0, y_0).

Démonstration :

La démonstration se fait en plusieurs étapes :

• Rien n'empêchant de supposer que $\varphi'_z(x_0, y_0, z_0) = 2m > 0$ (puisqu'il suffit de multiplier l'équation (1) par -1 si tel n'était pas le cas pour s'y ramener), la continuité entraîne alors qu'il existe trois nombres réels positifs h, k, l tels que les relations

$$|x - x_0| < h \quad , \quad |y - y_0| < k \quad , \quad |z - z_0| < l \tag{2}$$

impliquent

$$\varphi'_z(x, y, z) > m > 0$$

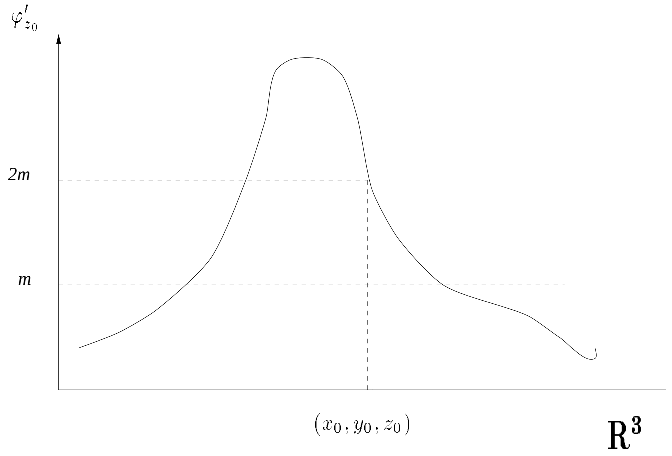


Figure 3.5. Choix et détermination du domaine de travail

• x et y sont ensuite fixés avec pour seule condition de vérifier les deux premières inégalités de (2) ; z quant-à lui variera de

$$z' = z_0 - l \quad \text{à} \quad z'' = z_0 + l$$

$\varphi(x, y, z)$ est alors une fonction croissante de z et varie de $\varphi(x, y, z')$ à $\varphi(x, y, z'')$.
Sa dérivée par rapport à z étant supérieure à m, φ croît plus vite que mz ; donc $\varphi(x, y, z) - mz$ est une fonction croissante de z.
En intégrant successivement de z' à z_0 et de z_0 à z'' l'inégalité

$$\varphi'_z(x, y, z) > m > 0$$

nous obtenons

$$\int_{z'=z_0-l}^{z_0} [\varphi'_z(x, y, z) - m]dz \geq 0 \quad \text{et} \quad \int_{z_0}^{z''=z_0+l} [\varphi'_z(x, y, z) - m]dz \geq 0$$

soit

$$\varphi(x, y, z_0) - \varphi(x, y, z') - ml \geq 0 \quad \text{et} \quad \varphi(x, y, z'') - \varphi(x, y, z_0) - ml \geq 0$$

inégalités qu'il est plus pratique d'écrire sous la forme

$$\varphi(x, y, z') \leq \varphi(x, y, z_0) - ml \quad \text{et} \quad \varphi(x, y, z'') \geq \varphi(x, y, z_0) + ml$$

• Comme $\varphi(x_0, y_0, z_0) = 0$ et que la fonction φ est continue, il existe un voisinage de (x_0, y_0, z_0), contenu dans le précédent tel que $|\varphi(x, y, z_0)| < ml$.
Dans ce voisinage, les inégalités ci-dessus montrent que

(3) $$\varphi(x, y, z') < 0 \quad \text{et} \quad \varphi(x, y, z'') > 0$$

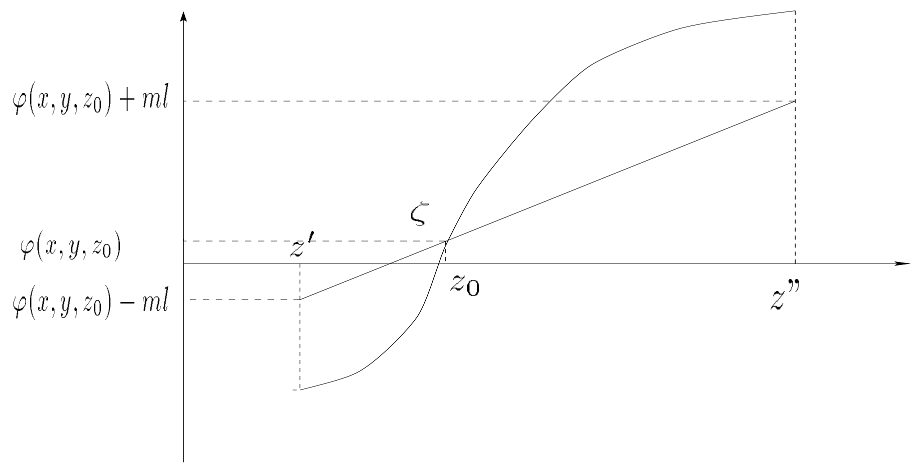


Figure 3.6. Structure finale de la démonstration.

- Conséquence :

φ en tant que fonction de z, pour x et y fixés satisfaisant à (2), est une fonction croissante, continue dans l'intervalle (z', z'') et satisfaisant à (3). Elle s'annule donc pour une valeur $z = \zeta$ et une seule (théorème de la valeur intermédiaire).

Donc pour tout x et y fixés satisfaisant à (2), il existe un z et un seul qui annule (1).

- Il reste à montrer la continuité de z dans ce domaine ; pour cela, il suffit de montrer cette continuité en z_0, par exemple, puisque rien ne distingue x_0, ou y_0 des autres points x ou y de la région.

Soit donc z la valeur telle que $\varphi(x, y, z) = 0$. De la relation $\varphi'_z(x, y, z) \geq m$, une intégration de z_0 à z conduit à l'inégalité :

$$|\varphi(x, y, z) - \varphi(x, y, z_0)| \geq m|z - z_0|$$

Comme $\varphi(x, y, z) = 0$ et comme φ est continue le premier membre tend vers 0 quand (x, y) tend vers (x_0, y_0) ; $z - z_0$ tend donc vers 0, ce qui achève la démonstration.

Autres formulations du théorème des fonctions implicites :

Cette démonstration s'étend à un nombre quelconque, mais fini, de variables ; mais il n'est pas démontré dans le cas des espaces vectoriels de dimension infinie.

Interprétation géométrique.

Elle se fait très simplement dans le cas d'une fonction de deux variables, et, pour la clarté de l'exposé, nous énoncerons à nouveau le théorème des fonctions implicites dans ce cas.

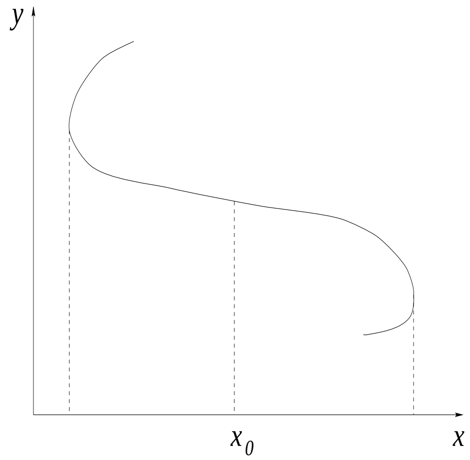


Figure 3.7. Les limites du domaine d'application du théorème des fonctions implicites.

La relation

$$\varphi(x, y) = 0$$

définit une variété à une dimension dans $\mathbb{R}^2$ *; c'est donc une courbe du plan.*

Soit (x_0, y_0) *un point de cette courbe ; on a :* $\varphi(x_0, y_0) = 0$.

Si $\varphi(x, y)$ *est continue dans un voisinage de* (x_0, y_0),
si $\varphi'_y(x, y)$ *existe et est continue dans le même voisinage,*
et si $\varphi'_y(x_0, y_0) \neq 0$,
alors il existe une application $\omega\colon \mathbb{R} \longrightarrow \mathbb{R}$ *;* $\quad x \longrightarrow \omega(x)$ *définie et continue dans un voisinage de* x_0 *telle que*

$$y_0 = \omega(x_0) \quad \text{et} \quad \varphi(x, \omega(x)) \equiv 0 \quad \text{dans ce voisinage.}$$

Si de plus, $\varphi'_x(x, y)$ *existe dans ce même voisinage et est continue, alors* ω *est dérivable dans un voisinage de* x_0 *où l'on a :*

$$\omega'(x) = -\frac{\varphi'_x(x, \omega(x))}{\varphi'_y(x, \omega(x))} \quad \text{avec} \quad \omega'(x_0) = -\frac{\varphi'_x(x_0, y_0)}{\varphi'_y(x_0, y_0)}$$

Remarque 1 :

Sous la forme précédente, la démonstration de la formule pour la dérivée est évidente : S'il existe un voisinage de x_0 tel que $\varphi(x, \omega(x)) = 0$ pour tout x de ce voisinage, le théorème de dérivation d'une fonction composée s'applique :

$$\varphi'_x(x, \omega(x)) + \varphi'_y(x, \omega(x))\omega'(x) = 0$$

La continuité en x_0 s'exprime alors par :

$$\varphi'_x(x_0, y_0) + \varphi'_y(x_0, y_0)\omega'(x_0) = 0$$

d'où la valeur de $\omega'(x_0)$ (si $\varphi'_y(x_0, y_0) \neq 0$).

Remarque 2 :

Dans la première partie du théorème, il est nécessaire de supposer que $\varphi(x, y)$ et $\varphi'_y(x, y)$ existent et sont continues ; φ'_x n'intervient pas.

Par contre l'introduction de l'existence et de la continuité de φ'_x entraîne l'existence et la continuité de la différentielle de φ.

Interprétation géométrique :

L'équation de la tangente au point courant (x_0, y_0) de la courbe $\varphi(x, y) = 0$ s'exprime à partir du gradient de φ par :

$$(x - x_0)\frac{\partial \varphi}{\partial x}(x_0, y_0) + (y - y_0)\frac{\partial \varphi}{\partial y}(x_0, y_0) = 0$$

La quantité $-\varphi'_{x_0}/\varphi'_{y_0}$ est le coefficient angulaire de cette droite.

Les points où φ'_{y_0} est nul sont les points où la tangente à la courbe est parallèle à l'axe des y.

Le théorème des fonctions implicites indique qu'il n'existe de fonction implicite que dans le voisinage de tout point dont la tangente n'est pas parallèle à l'axe des y.

Cette phrase peut être exprimée autrement : étant donné un point x_0, il ne doit y avoir qu'un point y_0 déterminé sur la courbe par une parallèle à $0y$ passant par x_0. C'est la notion même de fonction qui apparaît là.

Remarque 3 :

L'interprétation est la même en dimension 3, la tangente étant simplement remplacée par le plan tangent. On conçoit fort bien que cette interprétation s'étende pour les hypersurfaces à la dimension N.

Exemple :

Si la courbe choisie est un cercle : $\varphi(x, y) = x^2 + y^2 - 1 = 0$, la dérivée partielle $\varphi'_y = 2y$ s'annule pour $y = 0$ et donc aux points d'abscisses $x = \pm 1$.

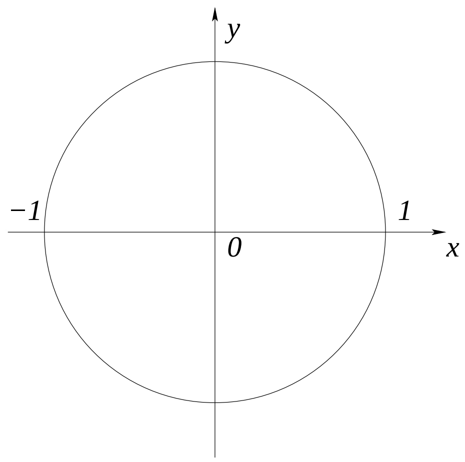


Figure 3.8. Les deux domaines du cercle.

Le cercle est donc naturellement décomposé en deux demi-cercles

$$y = +\sqrt{1-x^2} \quad \text{et} \quad y = -\sqrt{1-x^2}$$

qui tous deux correspondent à des fonctions parfaitement définies. Il est parfaitement évident qu'il n'est pas possible d'"inverser le cercle" dans le voisinage des deux points $(-1, 0)$ et $(1, 0)$.

3.6. Multiplicateur de Lagrange.

Exposé de la méthode :

Supposons que l'on recherche les extrema d'une fonction de deux variables réelles $f(x, y)$ assujetties à la condition

$$\varphi(x, y) = 0 \tag{4}$$

et supposons qu'il y ait bien un extremum relatif en (x_0, y_0).
Supposons enfin qu'en (x_0, y_0) φ satisfasse aux hypothèses du théorème des fonctions implicites ($\varphi(x_0, y_0) = 0$; $\varphi'_{y_0} \neq 0$; φ et φ'_{y_0} sont continues dans un voisinage de (x_0, y_0))
D'après ce théorème, (4) définit de façon unique au voisinage de $x = x_0$ une fonction $y = \omega(x)$ telle que $y_0 = \omega(x_0)$ et $\varphi(x, \omega(x)) = 0$ dans un voisinage de x_0 ; cette fonction est même dérivable dans le même voisinage si φ'_x existe et est continue, ce que nous supposerons également.
Si nous remplaçons dans f y par $\omega(x)$, nous pouvons alors affirmer que la fonction

$$x \longrightarrow f(x, \omega(x))$$

doit passer par un extremum en x_0, et donc que sa dérivée doit s'annuler en ce point, ce qui s'écrit

$$f'_x(x_0, \omega(x_0)) + f'_y(x_0, \omega(x_0))\omega'(x_0) = 0$$

Par ailleurs si nous effectuons la même opération dans (4), nous obtiendrons après dérivation :

$$\varphi'_x(x_0, \omega(x_0)) + \varphi'_y(x_0, \omega(x_0))\omega'(x_0) = 0$$

Multipliant alors cette dernière équation par λ et l'ajoutant à la première, nous obtiendrons :

$$(f'_{x_0} + \lambda\varphi'_{x_0}) + (f'_{y_0} + \lambda\varphi'_{y_0})\omega'(x_0) = 0$$

Si nous définissons λ par la relation

$$f'_{y_0} + \lambda\varphi'_{y_0} = 0$$

ce qui est toujours possible puisque $\varphi'_{y_0} \neq 0$ et ce qui permettra d'éviter d'utiliser $\omega'(x_0)$, il reste à satisfaire l'égalité

$$f'_{x_0} + \lambda\varphi'_{x_0} = 0$$

condition qui jointe à $\varphi(x_0, y_0) = 0$ permet de calculer x_0 et y_0.
En conclusion, les trois équations

$$\begin{aligned} &f'_{x_0} + \lambda\varphi'_{x_0} = 0 \\ &f'_{y_0} + \lambda\varphi'_{y_0} = 0 \\ &\varphi(x_0, y_0) = 0 \end{aligned}$$

permettent de calculer λ, x_0, y_0 si $\varphi'_{y_0} \neq 0$ et de montrer qu'il s'agit d'un extremum.
En effet, supposons ces éléments calculés, les équations permettent toujours d'écrire

$$f'_{x_0} + \lambda\varphi'_{x_0} + (f'_{y_0} + \lambda\varphi'_{y_0})\omega'(x_0) = 0$$

soit encore

$$f'_{x_0} + \omega'(x_0)f'_{y_0} + \lambda(\varphi'_{x_0} + \omega'(x_0)\varphi'_{y_0}) = 0$$

L'hypothèse de dérivabilité de φ en x_0 et y_0 implique la nullité du terme entre parenthèses et les termes restant montrent que f est également dérivable aux mêmes points ; f y est donc extrémale.

Remarques :

– Si $\varphi'_{y_0} = 0$ et $\varphi'_{x_0} \neq 0$, il est toujours possible de faire le même raisonnement en échangeant les rôles des dérivées partielles.

– Si $\varphi'_{y_0} = 0$ et $\varphi'_{x_0} = 0$, le théorème des fonctions implicites ne s'applique pas et la méthode ne peut fournir de résultat, ce qui ne prouve pas son inexistence.

– L'introduction de λ permet d'éviter d'avoir à calculer explicitement $\omega'(x_0)$. λ est dit "multiplicateur de Lagrange".

Bilan et généralisation :

Les résultats précédents se généralisent à un nombre fini de conditions ; dans le cas de deux conditions, ils se traduisent par le

Théorème :

Pour écrire qu'une fonction réelle f de n variables réelles est extremum sachant que

$$\varphi_1(x_1, \ldots, x_n) = 0 \quad \text{et} \quad \varphi_2(x_1, \ldots, x_n) = 0$$

où φ_1 et φ_2 sont aussi des fonctions réelles de n variables réelles, on annule les n dérivées partielles de

$$f + \lambda\varphi_1 + \mu\varphi_2$$

où λ et μ sont des constantes.

On obtient ainsi n équations qui, jointes à $\varphi_1 = 0$ et $\varphi_2 = 0$ permettent en général de calculer

$$x_1, \ldots, x_n, \lambda, \mu$$

λ et μ s'appellent multiplicateurs de Lagrange.

Interprétation en termes de différentielles :

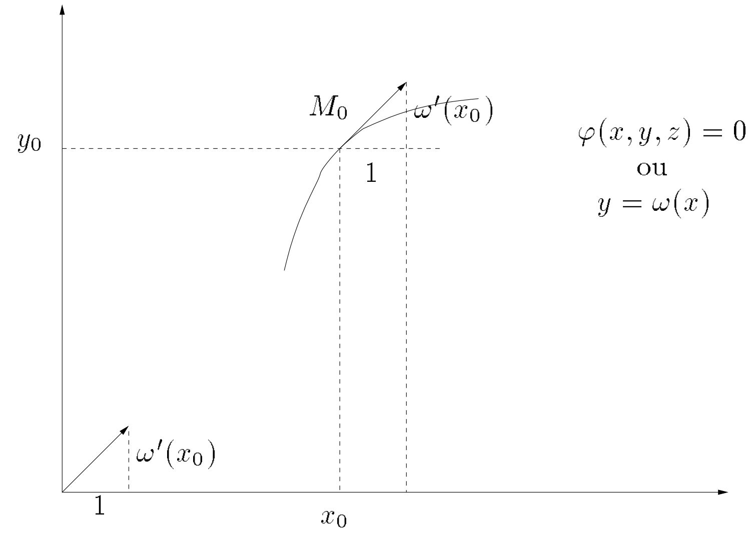


Figure 3.9. Interprétation géométrique des relations différentielles.

Reprenons les résultats du sous-paragraphe précédent en introduisant dans l'interprétation de ce qui a été fait la notion de différentielle.

Nous avons recherché les extrema d'une "fonction" $f(x, y)$ en un point (x_0, y_0) sachant que x et y devaient satisfaire à une relation $\varphi(x, y) = 0$ et que, de plus, au voisinage de ce point (x_0, y_0), y devait s'exprimer comme une fonction de ω de x.

De l'hypothèse $\varphi(x, y) = 0$ et de l'existence et la continuité de toutes les dérivées partielles de φ, nous avons déduit successivement les deux relations :

$$\left(\frac{\partial \varphi}{\partial x}\right)_0 + \left(\frac{\partial \varphi}{\partial y}\right)_0 \omega'(x_0) = 0$$
$$\left(\frac{\partial f}{\partial x}\right)_0 + \left(\frac{\partial f}{\partial y}\right)_0 \omega'(x_0) = 0$$

la seconde étant une conséquence de la première. Autrement dit,

$$\left(\frac{\partial \varphi}{\partial x}\right)_0 + \left(\frac{\partial \varphi}{\partial y}\right)_0 \omega'(x_0) = 0 \qquad \Longrightarrow \qquad \left(\frac{\partial f}{\partial x}\right)_0 + \left(\frac{\partial f}{\partial y}\right)_0 \omega'(x_0) = 0$$

En termes de différentielles, la différentielle de φ s'annule en tous les points du plan tels que $\varphi(x, y) = 0$ et, en particulier, sur ceux où y peut être défini comme une fonction implicite de x. Comme il est donc toujours possible d'écrire :

$$\left(\frac{\partial \varphi}{\partial x}\right)_0 + \left(\frac{\partial \varphi}{\partial y}\right)_0 \omega'(x_0) = 0 = \left(\left(\frac{\partial \varphi}{\partial x}\right)_0, \left(\frac{\partial \varphi}{\partial y}\right)_0\right) \begin{pmatrix} 1 \\ \omega'(x_0) \end{pmatrix} \quad ,$$

et le vecteur

$$v = \begin{pmatrix} 1 \\ \omega'(x_0) \end{pmatrix}$$

appartient au noyau de la forme linéaire dont les composantes dans la base canonique s'écrivent $\left(\left(\dfrac{\partial\varphi}{\partial x}\right)_0, \left(\dfrac{\partial\varphi}{\partial y}\right)_0\right)$.
Ce vecteur v, issu de l'origine, est parallèle au vecteur tangent en (x_0, y_0) à la courbe $\varphi(x, y) = 0$. Comme il n'est jamais nul, il engendre un espace vectoriel : la droite menée par l'origine parallèle à la tangente à la courbe en (x_0, y_0). Cet espace vectoriel (la droite) est donc le noyau de la forme linéaire définie ci-dessus[4].

Tout élément h de ce noyau étant proportionnel au vecteur v, les deux relations initiales peuvent également s'écrire :

$$\begin{aligned}\left(\frac{\partial f}{\partial x}\right)_0 h_1 + \left(\frac{\partial f}{\partial y}\right)_0 h_2 = 0\\ \left(\frac{\partial\varphi}{\partial x}\right)_0 h_1 + \left(\frac{\partial\varphi}{\partial y}\right)_0 h_2 = 0\end{aligned}$$

La relation logique, quant-à elle, prend la forme

$$\forall\alpha\in\mathbb{R}, h = \alpha v = \begin{pmatrix} h_1\\ h_2\end{pmatrix},\ \left(\frac{\partial\varphi}{\partial x}\right)_0 h_1 + \left(\frac{\partial\varphi}{\partial y}\right)_0 h_2 = 0 \quad\Longrightarrow\quad \left(\frac{\partial f}{\partial x}\right)_0 h_1 + \left(\frac{\partial f}{\partial y}\right)_0 h_2 = 0$$

En reprenant les notations différentielles, les hypothèses de la théorie des multiplicateurs de Lagrange s'écrivent indifféremment :

$$\mathrm{Ker}\ (\mathrm{grad}\,\varphi)_0 \subset \mathrm{Ker}\ (\mathrm{grad}\,f)_0$$

ou

$$\text{pour tout } h \in \mathbb{R}^2 \text{ tel que, } (d\varphi)_0 h = 0 \text{ alors } (df)_0 h = 0$$

Dans le sous-paragraphe précédent, nous avons également montré que :

$$\begin{aligned}f'_{x_0} + \lambda\varphi'_{x_0} = 0\\ f'_{y_0} + \lambda\varphi'_{y_0} = 0\end{aligned}$$

c'est-à-dire que nous avons établi la proportionnalité entre les composantes des dérivées partielles. En "langage de différentielles", cela s'écrit

$$(df)_0 = -\lambda(d\varphi)_0$$

Autrement dit, les différentielles sont proportionnelles, et nous avons également vu que l'on pouvait établir la réciproque de cette propriété.

En résumé, nous avons donc effectué deux étapes :

---

[4] Rappelons que le noyau d'une application linéaire g d'un espace vectoriel E dans un espace vectoriel F est l'ensemble des t appartenant à E tels que $g(t) = 0$.

– la première relève de l'analyse et peut se schématiser par l'équivalence

$$\begin{array}{c} \text{extremum de } f \\ \text{sachant que } \varphi = 0 \end{array} \iff \begin{array}{c} \text{sur l'extremum} \\ \forall h,\ (d\varphi)_0 h = 0 \Longrightarrow (df)_0 h = 0 \end{array}$$

et conduit donc à une relation entre formes linéaires.

– La seconde étape relève, elle, de l'algèbre : $L_1 = (d\varphi)_0$ et $L = (df)_0$ étant deux formes linéaires, nous avons montré l'équivalence :

$$\Big(\text{Ker } L_1 \subset \text{Ker } L\Big) \iff \Big(\exists \lambda \in \mathbb{R},\ L = -\lambda L_1\Big)$$

ce qui s'écrit encore

$$\Big(\forall h, L_1 h = 0 \Longrightarrow Lh = 0\Big) \iff \Big(\exists \lambda \in \mathbb{R},\ L = -\lambda L_1\Big)$$

Nous admettrons que ces deux étapes se généralisent, ce qui conduit au théorème suivant :

Théorème :

Soit E un espace vectoriel réel de dimension quelconque, et f une application différentiable de E dans $\mathbb{R}$.

Si f admet en a un extremum quand x satisfait à

$$C_1(x) = \cdots = C_k(x) = 0$$

les C_i étant différentiables,
alors, pour tout h de E tel que :

$$(dC_1(a))h = \cdots = (dC_k(a))h = 0 \text{ on a } (df(a))h = 0$$

où $(dC_i(a))h$ désigne la valeur sur h de la différentielle de l'application C_i calculée au point a.

De plus, L et L_i ($i = 1, \dots k$) étant des formes linéaires sur un espace vectoriel E de dimension quelconque, on a alors l'équivalence

$$\Big[\forall h \in E, \Big(L_1(h) = \cdots = L_k(h) = 0\Big) \Longrightarrow L(h) = 0\Big] \iff \Big(\exists \lambda_i \in \mathbb{R},\ L = \lambda_1 L_1 + \cdots + \lambda_k L_k\Big)$$

Autrement dit, $(df(a))$ s'exprime comme une combinaison linéaire des $(dC_i(a))$ dont les coefficients sont les multiplicateurs de Lagrange.

Exemples :

Certains des exercices qui suivent ont déjà été traités par une autre méthode ; ils sont repris dans le cadre de celle des multiplicateurs de Lagrange.

• *Chercher les extrema de* $f(x,y) = 3x^2 + y^2$ *sachant que* $\varphi(x,y) = x^2 + y^2 = 1$
L'annulation des deux dérivées partielles de

$$f(x,y) + \lambda\varphi(x,y) = 3x^2 + y^2 + \lambda(x^2 + y^2 - 1)$$

conduit aux équations

$$\begin{aligned} 6x + 2\lambda x &= 0 \\ 2y + 2\lambda y &= 0 \end{aligned}$$

Le système d'équations à résoudre s'écrit donc

$$\begin{aligned} 3x + \lambda x &= 0 \\ y + \lambda y &= 0 \\ x^2 + y^2 &= 1 \end{aligned}$$

La première équation a deux solutions : $x = 0$ et $\lambda = -3$

$x = 0$, reportée dans les deux autres équations fournit comme solutions $y = \pm 1$ et $\lambda = -1$.

$\lambda = -3$ fournit, de son côté, $y = 0$ et $x = \pm 1$.

En définitive, à $\lambda = -1$, $x = 0$, $y = \pm 1$ correspond le minimum 1 de f ; alors que $\lambda = -3$, $x = \pm 1$, $y = 0$ correspondent au maximum 3. Ce sont bien les solutions déjà trouvées.

Sur cet exemple, il est facile de voir la "dualité" des problèmes

- extrema de $3x^2 + y^2$ sachant que $x^2 + y^2 = 1$;
- extrema de $x^2 + y^2$ sachant que $3x^2 + y^2 = 1$.

• *Chercher les extrema de* $f(x,y) = 3x^2 + xy + y^2$ *sachant que* $\varphi(x,y) = x^2 + y^2 = 1$
A nouveau, l'annulation des dérivées partielles de

$$f(x,y) + \lambda\varphi(x,y) = 3x^2 + xy + y^2 + \lambda(x^2 + y^2 - 1)$$

conduit au système d'équations à résoudre

$$\begin{aligned} 2(3+\lambda)x + y &= 0 \\ x + 2(1+\lambda)y &= 0 \\ x^2 + y^2 &= 1 \end{aligned}$$

Si nous éliminons λ entre les deux premières équations, nous obtenons le système :

$$\begin{aligned} 4xy + y^2 - x^2 &= 0 \\ x^2 + y^2 &= 1 \end{aligned}$$

Les solutions en y de la première équation sont maintenant calculables en fonction de x et valent

$$y = x(-2 \pm \sqrt{5})$$

Il ne reste plus qu'à les reporter dans la seconde pour obtenir les solutions en x

$$x^2 = \frac{5 \mp 2\sqrt{5}}{10} = \begin{cases} 0.052\ldots \\ 0.947\ldots \end{cases}$$

d'où nous pouvons déduire x, puis y, et éventuellement λ.

La systématique du calcul est évidente ; sur cet exemple, on voit à nouveau que l'on peut intervertir les rôles de $f(x,y) = 3x^2 + xy + y^2$ et de $\varphi(x,y) = x^2 + y^2 = 1$ dans le calcul des extrema.

• *Trouver les côtés d'un triangle rectangle dont le périmètre soit minimal pour une surface S donnée.*

Nous gardons les mêmes notations que dans l'exemple introductif. La quantité à minimiser est soumise à deux conditions :

– le triangle est rectangle : $a^2 - b^2 - c^2 = 0$;

– sa surface est constante : $bc - 2S = 0$.

il y a donc deux multiplicateurs de Lagrange et la fonction associée s'écrit :

$$F(a,b,c) = a + b + c + \lambda(bc - 2S) + \mu(a^2 - b^2 - c^2) = 0$$

Le système d'équations contenant la solution du problème est donc constitué des deux équations correspondant aux conditions ci-dessus et aux trois équations correspondant à l'annulation des dérivées partielles de la fonction F :

$$\begin{aligned} \partial_a F(a,b,c) &= 1 + 2\mu a = 0 \\ \partial_b F(a,b,c) &= 1 + \lambda c - 2\mu b = 0 \\ \partial_c F(a,b,c) &= 1 + \lambda b - 2\mu c = 0 \end{aligned}$$

La première des équations donne la valeur de μ qui, reportée dans les suivantes fournit :

$$\begin{aligned} 1 + \lambda c + \frac{b}{a} &= 0 \\ 1 + \lambda b + \frac{c}{a} &= 0 \end{aligned}$$

Il suffit d'éliminer λ pour obtenir

$$(b - c)\Big(\frac{a+b+c}{a}\Big) = 0$$

et comme le périmètre est toujours différent de 0, il en résulte que $b = c$.

Il ne reste plus qu'à reporter cette égalité dans les équations de conditions pour obtenir

$$b = \sqrt{2S} \quad \text{puis} \quad a = 2\sqrt{S}$$

Il est à remarquer que nous aurions aussi bien pu dès le départ imposer que $a = \sqrt{b^2+c^2}$; dans ce cas, il ne reste plus qu'une condition (celle de la surface) et la fonction à considérer s'écrit :

$$G(b,c) = \sqrt{b^2+c^2} + b + c + \lambda(bc - 2S)$$

L'annulation de ses dérivées partielles fournit le système

$$\begin{aligned} \frac{b}{\sqrt{b^2+c^2}} + 1 + \lambda c = 0 \\ \frac{c}{\sqrt{b^2+c^2}} + 1 + \lambda b = 0 \end{aligned}$$

L'élimination de λ conduit alors à :

$$(b-c)\Big[\frac{b+c}{\sqrt{b^2+c^2}} + 1\Big] = 0$$

Le terme entre crochets, qui fait intervenir le périmètre, est nécessairement strictement positif, et la seule solution possible de cette équation est :

$$b = c$$

Reportée dans la condition pour la surface, elle redonne (heureusement !) la même valeur de b et donc de a.

- *Calculer la plus courte distance d'un point à un plan.*

Nous nous plaçons dans un système de coordonnées rectangulaires dans lequel le point a pour coordonnées (a, b, c) et le plan pour équation

$$Ax + By + Cz + D = 0$$

et où (x, y, z) sont les coordonnées d'un point quelconque du plan.
Le problème consiste à minimiser (le carré de) la distance entre deux points

$$f(x,y,z) = (x-a)^2 + (y-b)^2 + (z-c)^2$$

sachant que

$$\varphi(x,y,z) = Ax + By + Cz + D = 0$$

Il n'y a donc qu'une condition et la fonction à introduire s'écrit :

$$\Phi(x,y,z) = (x-a)^2 + (y-b)^2 + (z-c)^2 + \lambda\big(Ax + By + Cz + D\big)$$

Les solutions cherchées sont celle du système construit à partir de l'annulation des dérivées partielles de la fonction ci-dessus et de l'équation du plan, soit :

$$\begin{aligned} x &= a - \frac{1}{2}\lambda A \\ y &= b - \frac{1}{2}\lambda B \\ x &= c - \frac{1}{2}\lambda C \\ Ax + By &+ Cz + D = 0 \end{aligned}$$

En reportant les trois premières équations dans la dernière, nous obtenons

$$\lambda = \frac{2(Aa + Bb + Cc + D)}{A^2 + B^2 + C^2}$$

et donc x, y, z, ce qui achève la démonstration.

• *Les axes étant choisis orthogonaux, trouver la distance minimum entre le point de coordonnées* $(-1, 0)$ *et la cubique d'équation* $y^2 - x^3 = 0$.

C'est un exemple pour lequel la méthode des multiplicateurs de Lagrange ne permet pas de calculer la solution alors que cette dernière existe.

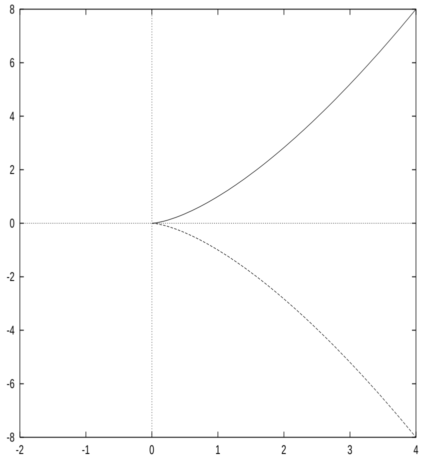


Figure 3.10. La courbe $y^2 = x^3$; l'origine, qui est le point le plus proche de cette courbe de $(-1, 0)$, ne peut être déterminée par la méthode puisque c'est un point singulier.

En effet, géométriquement la cubique a l'allure ci-contre et il est évident que le point de la cubique le plus proche de $(-1, 0)$ est l'origine.

Or, la contrainte correspondant à ce problème s'écrit

$$\varphi(x, y) = y^2 - x^3 = 0$$

et cette fonction n'est pas inversible à l'origine puisque

$$\frac{\partial\varphi}{\partial x}(0, 0) = \frac{\partial\varphi}{\partial y}(0, 0) = 0$$

La méthode des multiplicateurs de Lagrange qui consiste à minimiser

$$f(x, y) = (x + 1)^2 + y^2$$

sachant que

$$\varphi(x, y) = y^2 - x^3 = 0$$

conduit au système

$$\begin{aligned} f'_x + \lambda\varphi'_x &= 0 = 2(x + 1) - 3\lambda x^2 \\ f'_y + \lambda\varphi'_y &= 0 = 2y + 2\lambda y \\ \varphi(x, y) &= 0 = y^2 - x^3 \end{aligned}$$

Il est alors évident que la solution du problème $x = 0$, $y = 0$ ne satisfait pas à ce système puisqu'elle conduit à une absurdité dans le cas de la première équation.

C'est tout à fait l'analogue de la situation que l'on rencontre dans le calcul des extrema des fonctions. Ces derniers peuvent être obtenus pour les valeurs annulant dérivées ou différentielles ; mais ils peuvent très bien se trouver en des points où les fonctions ne sont ni dérivables ni différentiables.

3.7. Compléments sous forme d'exercices.

Exercice 1. Extrema d'une fonction de deux variables.

On rappelle qu'une fonction réelle d'une ou plusieurs variables réelles atteint ses bornes dans tout ensemble compact (fermé borné de $\mathbb{R}^N$). Comme les méthodes d'étude des extrema varient selon que l'on se trouve à l'intérieur ou sur la frontière de l'ensemble de définition de la fonction, on suivra pour cette étude les étapes suivantes :

1. recherche de l'ensemble des extrema de la fonction ;
2. restriction de l'étude à ceux situés à l'intérieur du domaine ;
3. étude des extrema sur la frontière où la méthode des multiplicateurs de Lagrange s'applique souvent facilement.

On applique ce qui précède à l'étude des extrema de la fonction définie par :

$$f(x, y) = x^2 + y^2 - xy + x + y$$

dans le domaine $x \leq 0, \quad y \leq 0, \quad x + y \geq -3$

1. Après avoir tracé le domaine de définition de f, trouver le ou les points du plan où f est stationnaire ($f'_x(x, y) = f'_y(x, y) = 0$), puis vérifier s'ils sont ou non dans l'ensemble de définition de la fonction.

2. Le type de l'extremum (maximum, minimum,...) s'obtient en étudiant le comportement de f au voisinage du point où f est stationnaire.

2.a. Donner le développement de f au voisinage de chaque extremum (si (x_0, y_0) sont les coordonnées d'un extremum, on posera : $x = x_0 + u$, $y = y_0 + v$, et on se rappellera que f est une fonction algébrique).

2.b. En utilisant simplement les définitions, déduire du résultat précédent le type de l'extremum.

3. Reste l'étude des extrema sur la frontière. Pour cela,

3.a. paramétrer chaque segment de la frontière ;

3.b. puis faire l'étude de l'extremum correspondant, s'il existe ;

3.c. enfin vérifier, sur le segment de votre choix, que la méthode des multiplicateurs de Lagrange redonne le même résultat.

Exercice 2.

On considère la fonction réelle de deux variables réelles définie par

$$f(x, y) = x^4 + y^4 - 2(x - y)^2$$

1. En les justifiant, donner les équations satisfaites par les valeurs de x et de y qui rendent f extrémale.

2. Résoudre ces équations.

3. En utilisant le développement de Taylor de f, trouver parmi les solutions celles qui correspondent à des extrema et préciser la nature de ces derniers.

Exercice 3.

On se place toujours $\mathbb{R}^3$ muni de sa base canonique. Si

$$f(x, y, z) = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}$$

représente la valeur de la distance d'un point de $\mathbb{R}^3$ de coordonnées (x, y, z) à un point A de coordonnées (a, b, c). Les valeurs de x, y, z prises dans un ensemble donné et qui rendent extremum f sont alors les mêmes que celles qui rendent extremum f^2.

1. Pourquoi ?

Le but de cet exercice est de généraliser ce résultat lorsque x, y, z sont de plus soumis à une contrainte

$$K(x, y, z) = 0$$

2. Écrire les équations donnant les *extrema* de la fonction f ci-dessus lorsque x, y, z satisfont à la contrainte $K(x, y, z) = 0$. On appellera λ le multiplicateur de Lagrange correspondant.

3. Éliminer λ^2 dans le système de la question précédente. En déduire le système satisfait par les variables x, y, z. Discuter de l'indépendance des équations du système obtenu.

4. On reprend le même problème que celui de la question **2.** mais avec la fonction

$$g(x, y, z) = (f(x, y, z))^2$$

Écrire les équations d'Euler-Lagrange correspondant à la même condition $K(x, y, z) = 0$. Le multiplicateur de Lagrange sera maintenant noté μ.

5. Éliminer μ et montrer que x, y, z satisfont au même système d'équations que celui obtenu dans la question **3.**

Exercice 4.

L'espace $\mathbb{R}^3$ étant rapporté à sa base canonique, **en utilisant les multiplicateurs de Lagrange**, trouver le point du plan

$$2x + y + 3z + 6 = 0$$

qui est le plus proche du point de coordonnées $(0, 1, 0)$. (Justifier les équations du problème, ainsi que la conduite des calculs).

Exercice 5.

On cherche un quadrilatère de périmètre maximum, inscriptible dans un demi-cercle de rayon unité et dont l'un des côtés est le diamètre.

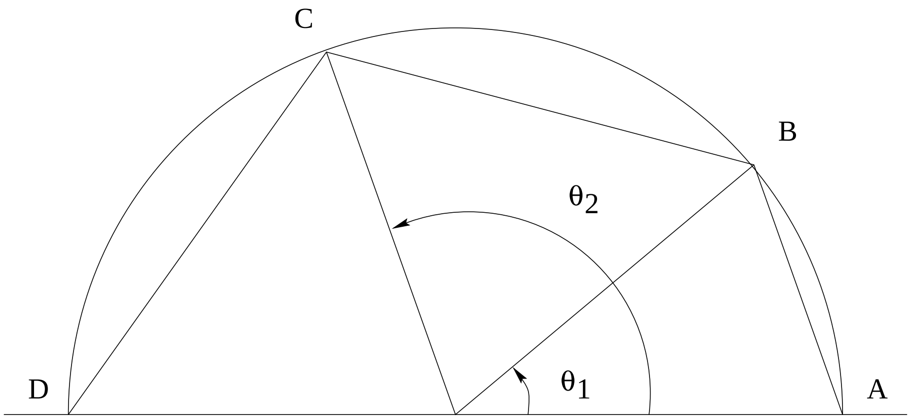


Figure . Les conventions de l'exercice 5.

1. Quels sont les théorèmes qui permettent d'affirmer que ce quadrilatère existe ?

2. Montrer que $AB = 2\sin(\dfrac{\theta_1}{2})$.

3. Écrire et justifier les équations du problème.

4. Les résoudre.

5. Dans un système de coordonnées cartésiennes dont l'origine est le centre du demi-cercle, le diamètre DA l'axe des x, on pose $A = (1, 0)$, $B = (x_1, y_1)$, $C = (x_2, y_2)$ et $D = (-1, 0)$. Expliquer pourquoi l'utilisation des multiplicateurs de Lagrange dans un tel problème aurait conduit à une résolution moins pratique que celle utilisée.

Exercice 6.

Le but de cet exercice est de retrouver d'une autre façon la définition usuelle des éléments propres à partir de celle du cours où l'on a montré que, H étant un opérateur symétrique défini-positif sur $\mathbb{R}^N$, sa plus grande valeur propre est définie par le maximum de (Hu, u) lorsque u parcourt le cercle unité $(u, u) = 1$ de $\mathbb{R}^N$ et que les vecteurs propres correspondants sont précisément ceux qui permettent d'atteindre ce maximum.

Pour simplifier les écritures, on se placera dans un espace vectoriel réel de dimension 2 muni du produit scalaire usuel ; on supposera que l'opérateur H possède de plus des valeurs propres distinctes et pour unifier les notations, on conviendra que H et u s'écrivent dans la base canonique sous la forme :

$$H = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \; ; \quad u = \begin{pmatrix} x \\ y \end{pmatrix}$$

1. En utilisant la méthode des multiplicateurs de Lagrange, écrire les équations correspondant à cette définition.

2. Retrouver à partir de ces équations la définition usuelle des éléments propres de H.

Exercice 7.

Dans le plan $3x - 2z = 0$, trouver le point dont la somme des carrés des distances aux points $A(1,1,1)$ et $B(2,3,4)$ est minimale.

Pour ce faire, on utilisera la méthode des multiplicateurs de Lagrange pour laquelle on posera soigneusement le problème avant d'en définir les équations correspondantes, puis de les résoudre.

Exercice 8.

On recherche les extrema de la fonction réelle de deux variables réelles

$$\sin^2 x + \sin^2 y$$

sachant que x et y sont soumises à la condition

$$\sin x + \sin y = 1$$

1. En utilisant la méthode des multiplicateurs de Lagrange, écrire les équations du problème.

2. Les résoudre.

Exercice 9.

On pourra fortement réduire les calculs de cet exercice en utilisant des propriétés de symétrie par permutation circulaire.

Dans un espace vectoriel $\mathbb{E}$ de dimension 3 muni de la norme euclidienne et rapporté à un système orthonormé, on considère les surfaces d'équations

$$x_1^2 + y_1^2 + z_1^2 = R^2 \qquad (S)$$

$$ax_2 + by_2 + cz_2 + d = 0 \qquad (P)$$

a, b, c sont des nombres réels ; d et R sont des nombres réels positifs. x_1, y_1, z_1 (resp. x_2, y_2, z_2) sont les coordonnées dans la base définie ci-dessus du point courant M_1 (resp. M_2) de S (resp. P).

1. Donner l'interprétation géométrique du triplet (a, b, c).

2. On recherche, si elle existe, la plus courte distance entre les surfaces S et P en minimisant la distance $M_1M_2^2$ et en imposant les conditions $M_1 \in S$ et $M_2 \in P$.
Écrire les équations du problème. On appellera λ (resp. μ) le multiplicateur de Lagrange associé à la surface S (resp. P).

3. En supposant que les points M_1 et M_2 sont solution du système précédent, déduire de ce même système l'interprétation géométrique de la droite M_1M_2.

4. Pour obtenir explicitement l'ensemble des solutions du système, on procédera de la manière suivante :

– calculer x_1, y_1, z_1 et x_2, y_2, z_2 en fonction de λ et μ. On pourra simplifier les écritures en posant

$$n = \sqrt{a^2 + b^2 + c^2}$$

– en déduire les relations auxquelles doivent satisfaire λ et μ, puis calculer ces derniers.

– donner enfin explicitement toutes les solutions du système.

5. Interpréter géométriquement les solutions obtenues. En déduire la solution du problème initial.

6. Calculer explicitement à partir des résultats précédents la valeur de la plus courte distance.

Exercice 10.

En optique, l'indice d'un milieu est le rapport de la vitesse de phase de la lumière dans le vide à celle de la vitesse de phase dans le milieu ($n = c/v$) et on appelle *chemin optique* le produit de la distance parcourue par un rayon lumineux dans un milieu donné par l'indice de ce milieu.

Le principe de Fermat stipule que le trajet suivi par les rayons lumineux correspond au chemin optique minimum.

On considère un système constitué de deux milieux repérés par les chiffres 1 et 2, d'indices différents et séparés par un plan et soit A_1 et A_2 deux points respectivement dans chacun des milieux et dont les projections sur le plan sont A' et A''.

On se place dans le plan défini par les points $A_1,\ A',\ A_2,\ A''$. On remarquera que A_1 et A_2 étant fixés, il en est de même de A' et A'' et on posera :

$$A_1A' = a_1 \qquad A_2A'' = a_2 \qquad A'A'' = d$$

Le chemin optique reliant A_1 et A_2 et passant par un point quelconque M de la surface de séparation se définit alors par :

$$\frac{1}{c}\ L(A_1MA_2) = \frac{A_1M}{v_1} + \frac{A_2M}{v_2} \quad \text{sachant que} \quad A'A'' = d = \text{constante.}$$

1. Soit M un point quelconque situé entre A' et A'' et soit M' le symétrique de M par rapport à A'. Montrer que $L(A_1M'A_2) \geq L(A_1MA_2)$; en déduire que le point M correspondant au minimum du chemin optique se trouve nécessairement entre A' et A''.

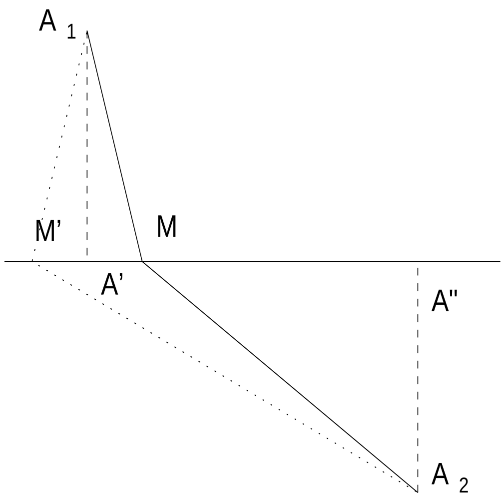


figure 10–a

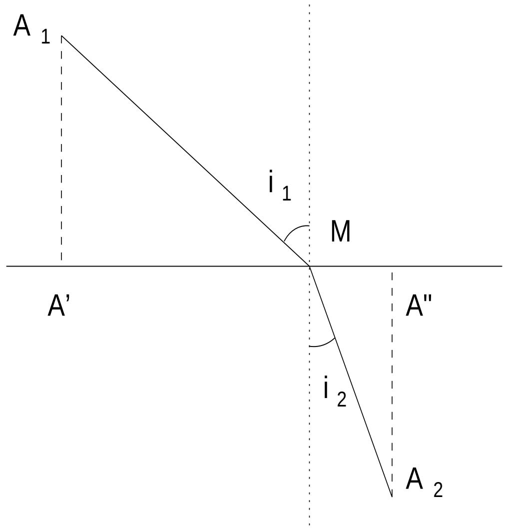


figure 10–b

2. En utilisant les multiplicateurs de Lagrange, écrire les équations du problème en fonction des angles i_1 et i_2.

3. Les résoudre et retrouver les lois de Snell-Descartes de la réfraction.

4. Récrire les équations du problème lorsque le rayon devient tangent à la surface de séparation (A_1M se confond ave $A'M$).

5. Les résoudre.

Exercice 11.

1. p, q, r étant trois nombres réels positifs, a un nombre réel positif, trouver, en utilisant la méthode des multiplicateurs de Lagrange, les valeurs de x, y, z du domaine $x \geq 0$, $y \geq 0$, $z \geq 0$ qui rendent maximum la quantité

$$x^p y^q z^r$$

sachant que $x + y + z = a$.

2. Sans refaire la démonstration, énoncer le résultat précédent pour un nombre quelconque N de variables.

3. Il est très facile de construire un entonnoir (c'est un cône droit de base circulaire) en découpant dans un cercle de rayon R un secteur circulaire d'angle α et en faisant coïncider les deux rayons externes restant.

On utilisera pour les calculs les notations de la figure 11. Exprimer le volume du cône en fonction de R et de α.

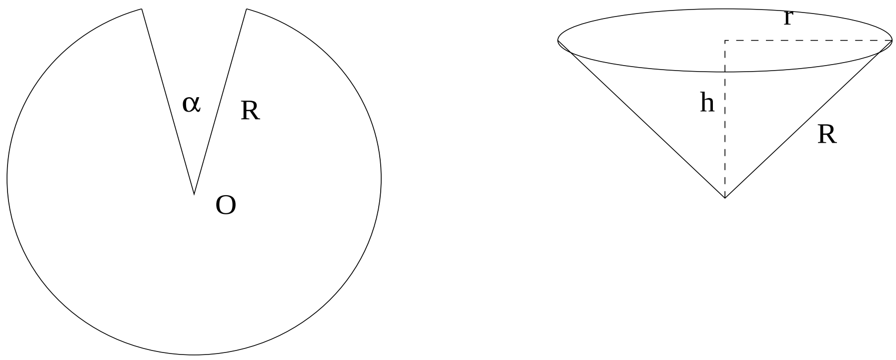
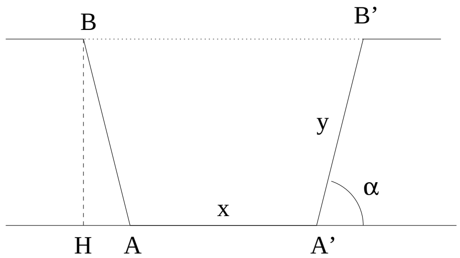


Figure de l'exercice 11. Fabrication d'un entonnoir avec un secteur circulaire.

4. En utilisant exclusivement le résultat de la question 2, calculer l'angle α du secteur que l'on doit découper pour obtenir un entonnoir de volume maximum.

Exercice 12.

Un canal, conçu pour un écoulement d'eau, peut être caractérisé par sa section puisque cette dernière est reliée au débit. On peut imposer à la section d'avoir la forme d'un trapèze isocèle que l'on schématise sur la figure suivante :



La grande base BB' simule le niveau permanent (ou maximal) de l'eau ; la petite base $AA' = y$ correspond au fond du canal et les côtés $\overrightarrow{AB} = \overrightarrow{A'B'} = x$ représentent les berges.

Si l'on désire minimiser les infiltrations, il est évident qu'il faut minimiser la surface de contact entre le sol et l'eau qui sera caractérisée par la quantité :

$$L(x, y) = 2x + y$$

Les paramètres importants sont donc alors : : x, y et l'angle $\alpha = (AA', A'B')$ que forme la berge avec le fond.

1. Exprimer la surface S de la section droite du canal (c'est-à-dire la surface du trapèze $AA'B'B$) en fonction des paramètres x, y et α.

2. La minimisation des infiltrations revient donc à minimiser la quantité L (caractéristique des infiltrations) pour une quantité S (caractéristique du débit) donnée.

En expliquant soigneusement comment s'applique la méthode des multiplicateurs de Lagrange, écrire les équations du problème.

3. Montrer qu'une simple élimination des autres paramètres permet de calculer l'angle α ; donner sa valeur ; qu'a-t-elle de remarquable ?

4. Calculer ensuite x, y et le multiplicateur de Lagrange en fonction de S.

Exercice 13.

On considère une caisse parallélépipédique ouverte (pas de couvercle) dont l'épaisseur des parois est δ, dont le volume *intérieur* est V et dont les dimensions *extérieures* sont x, y, z, z étant sa hauteur.

Cette caisse est représentée sur la figure 1 qui suit ; son volume intérieur est :

$$(x - 2\delta)(y - 2\delta)(z - \delta) = V$$

et l'aire de ses panneaux est :

$$xy + 2(x + y - 2\delta)(z - \delta)$$

On désire fabriquer cette caisse en utilisant le minimum de matériaux.

1. Écrire en les justifiant les équations du problème.

2. Les résoudre pour obtenir x, y, z.

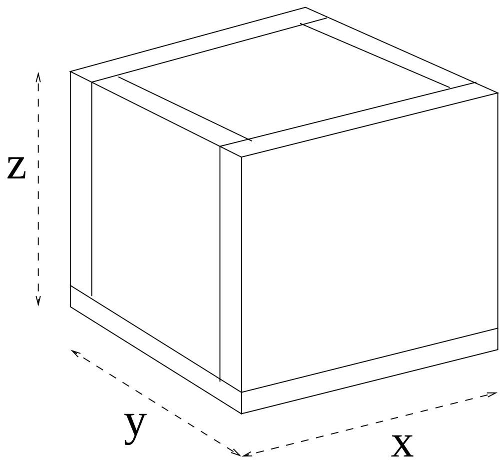


Figure de l'exercice 13. Mode de construction de la caisse.

Exercice 14.

Déterminer les dimensions d'une baignoire rectangulaire de volume V pour que sa surface soit minimum. On suppose que les côtés de la baignoire sont des portions de plan et on rappelle qu'une telle baignoire ne peut avoir que cinq côtés.

Exercice 15.

On cherche un point d'une ellipse où la tangente à cette dernière forme avec les axes de coordonnées un triangle de superficie minimum.

On se place donc dans $\mathbb{R}^2$, et, pour simplifier, on suppose que l'ellipse est rapportée à ses axes principaux. Son équation est donc de la forme :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

et les conventions utilisées sont reportées sur la figure suivante :

1. Écrire l'équation de la tangente en un point (x_0, y_0) quelconque de l'ellipse.

2. En déduire les coordonnées des points d'intersection A et B de cette tangente avec les axes de coordonnées, puis la valeur de la surface du triangle OAB.

3. Donner l'expression algébrique du problème géométrique initial.

4. Écrire les équations du problème.

5. Les résoudre et donner les coordonnées du ou des points solution.

6. Donner l'énoncé du même problème généralisé dans un espace à N dimensions.

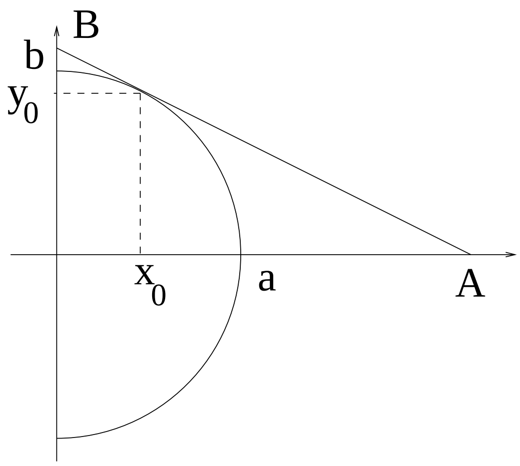


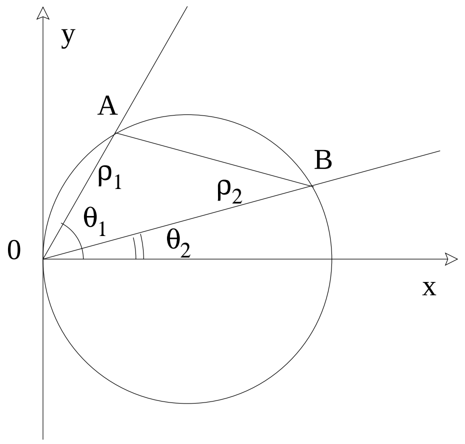
Figure de l'exercice 15. L'ellipse et les notations utilisées.

Exercice 16.

On veut, en utilisant la méthode des multiplicateurs de Lagrange, déterminer les triangles inscrits dans un cercle donné et d'aire maximum.

Pour cela, on se donne un cercle de rayon R. Il est alors toujours possible de choisir un point O quelconque sur ce cercle et de supposer qu'il est l'un des sommets d'un triangle.

C'est ce que l'on fera. Il restera à déterminer les deux autres sommets A et B sur le cercle compte tenu de la condition ci-dessus.



De plus, pour simplifier les expressions dans les calculs, on se placera en coordonnées polaires d'origine O, le cercle étant centré sur l'axe des x. Ce cercle aura donc pour équation : $\rho = 2R\cos\vartheta$

1. Les points A et B ayant pour coordonnées respectives (ρ_1, ϑ_1) et (ρ_2, ϑ_2), donner l'expression de l'aire du triangle OAB en fonction de ρ_1, ρ_2, ϑ_1 et ϑ_2. On supposera par la suite que $\vartheta_1 \geq \vartheta_2$.

2. En déduire, après avoir imposé les contraintes adéquates, les équations du problème.

3. Les résoudre et discuter du résultat obtenu.

Exercice 17.

x, y, z sont les coordonnées d'un point de $\mathbb{R}^3$ rapporté à sa base canonique. En utilisant la méthode des multiplicateurs de Lagrange, donner les équations auxquelles doivent obéir les extrema de $x^2 + y^2 + z^2$ sachant que x, y, z doivent satisfaire aux conditions :

$$\frac{x}{\sqrt{2}} + \frac{y}{\sqrt{3}} + \frac{z}{\sqrt{6}} = 0 \quad \text{et} \quad 3x^2 + 2y^2 + 6z^2 = 1$$

Les calculer et justifier l'interprétation géométrique de cet exercice.

Exercice 18.

L'intérêt de cet exercice est de montrer, dans le cas de la réfraction, qu'une particule relativiste n'obéit pas aux lois de Descartes, contrairement à ce qu'on peut lire parfois dans la littérature.

On considère un phénomène physique mesurable ponctuellement se déplaçant avec une vitesse constante dans un milieu homogène. L'espace physique est supposé rapporté aux trois axes de coordonnées d'un trièdre trirectangle.

On suppose de plus que pour aller d'un point à un autre de l'espace le phénomène en question parcourt la trajectoire qui lui permet de joindre les deux points en un temps minimum. Si le milieu est homogène, on a vu dans le cours que la trajectoire en question est le segment de droite joignant les deux points.

On suppose maintenant que l'espace est constitué de deux parties distinctes P_1 et P_2, chacune homogène, séparées par un plan Π d'équation :

$$ax + by + cz = d \tag{1}$$

et dans lesquels le phénomène physique considéré possède respectivement les vitesses v_1 et v_2. On suppose de plus que :

$$a^2 + b^2 + c^2 = 1 \tag{2}$$

et on rappelle, pour les étourdi(e)s, que les composantes (a, b, c) s'interprètent comme celles d'un vecteur $\overrightarrow{n}$ normal au plan.

On veut étudier maintenant la trajectoire qu'effectue le phénomène en question partant d'un point A_1 de coordonnées (x_1, y_1, z_1) de P_1 pour aboutir à un point A_2 de coordonnées (x_2, y_2, z_2) de P_2. Compte tenu de ce qui vient d'être dit précédemment, la trajectoire sera constituée de deux segments A_1M et MA_2 où M est un point de coordonnées (x, y, z) du plan Π. On notera pour simplifier :

$$d_i = \left((x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2\right)^{1/2} \quad \text{pour } i = 1, 2$$

Le but de cet exercice est de déterminer en utilisant les méthodes du calcul des variations la position de M et les lois qui régissent, dans ce cas, ce phénomène.

1. En tenant compte du fait que le temps mis pour parcourir toute la trajectoire est minimum et que M est astreint à se trouver sur le plan π, écrire toutes les équations déterminant ce phénomène.

2. Parmi les équations obtenues, trois se déduisent les unes des autres par permutation circulaire sur x, y, z et a, b, c. Montrer que ces dernières peuvent par combinaison se mettre sous la forme :

$$\frac{1}{v_1}\left(c\,\frac{y - y_1}{d_1} - b\,\frac{z - z_1}{d_1}\right) + \frac{1}{v_2}\left(c\,\frac{y - y_2}{d_2} - b\,\frac{z - z_2}{d_2}\right) = 0$$

les deux autres étant déduites par permutation circulaire sur x, y, z et a, b, c.

3. En posant :

$$\alpha_i = \frac{x - x_i}{d_i}\ ; \quad \beta_i = \frac{y - y_i}{d_i}\ ; \quad \gamma_i = \frac{z - z_i}{d_i} \qquad (i = 1, 2)$$

qui sont les composantes des vecteurs unitaires $\overrightarrow{u}_i$ $(i = 1, 2)$ portés par AM_i $(i = 1, 2)$, et en les reportant dans les équations ci-dessus, montrer que $\overrightarrow{n}$, $\overrightarrow{u}_1$ et $\overrightarrow{u}_2$ sont dans un même plan.

4. Interpréter géométriquement le résultat précédent pour en déduire les lois du phénomène et vérifier que ce ne sont pas les lois de Descartes.

Exercice 19.

Cet exercice est fortement inspiré d'un article de Pierre Curie, *Sur la formation des cristaux*, Bulletin de la Société minéralogique de France, **8**, 1885, p.145.

Lorsqu'un cristal se trouve dans des conditions adéquates (on dit qu'il est dans son eau mère), il se développe, c'est-à-dire qu'il grossit ; une des particularités étonnantes est que sa forme se conserve au cours du grossissement. Le but de cet exercice est de montrer sur un exemple aisément généralisable comment il est possible d'expliquer ce phénomène.

On définira un cristal comme un corps dont les faces sont planes et ont une forme géométrique régulière et l'on admettra que lors de sa formation dans son eau-mère que la seule énergie qui intervient est celle correspondant à la tension superficielle sur ses faces Cette énergie s'exprime pour une face par la quantité AS où S est la valeur de l'aire de la face et A est une constante caractéristique de la forme de la face.

On admettra que l'énergie du cristal est alors la somme des énergies correspondant aux tensions superficielles de chaque face, et que le système évolue pour un cristal donné de façon que cette énergie soit minimum.

On considère donc, et l'exercice commence ici, un cristal ayant la forme d'un prisme quadratique (c'est un parallélépipède rectangle dont la base est un carré de côté x et dont la hauteur a pour valeur y). L'énergie superficielle d'une face latérale sera xyA et celle sur une base x^2B. Pour la forme stable, la quantité

$$E(x, y) = 4xyA + 2x^2B$$

devra donc être minimum pour un volume $V = x^2y$ donné.

1. En traduisant la loi ci-dessus à l'aide de la méthode des multiplicateurs de Lagrange, montrer que le cristal conservera homothétiquement sa forme, c'est-à-dire que, quel que soit le volume V, le rapport x/y aura une valeur constante ne dépendant que des données géométriques du problème ; puis calculer explicitement x et y en fonction de V, de A et de B.

2. Reportant la relation entre x et y dans $E(x, y)$, vérifier que la solution obtenue correspond à un minimum de E.

3. On considère maintenant deux cristaux du type précédent de volumes distincts V_1 et V_2 . On pose $V = V_1 + V_2$. Montrer que les longueurs x et y définissant les dimensions d'un cristal du même type <u>et</u> de volume V seront supérieures à celles de chacun des deux cristaux.

En remarquant que l'énergie superficielle d'un cristal quelconque de volume V peut s'écrire

$$E = 6VB/y = 6VA/x$$

montrer que l'énergie de tension superficielle d'un cristal de volume V sera inférieure à la somme des énergies de deux cristaux de volumes V_1 et V_2 tels que $V = V_1 + V_2$.

4. En quelques lignes, montrer que les questions précédentes justifient la loi : *dans une eau mère saturée, le plus gros cristal mange les petits.*

CAHIERS D'ENSEIGNEMENT

Du même auteur et dans la même collection

FASCICULE I – Systèmes linéaires. Une introduction à l'algèbre linéaire.

FASCICULE II – (TAP 1) Présentation du cours.

FASCICULE III – (TAP 2) Analyse numérique et Ordinateurs.

FASCICULE IV – (TAP 3) Extrema et Différentielles.

FASCICULE V – (TAP 4) Calcul des variations.

FASCICULE VI – (TAP 5) Formes quadratiques et Opérateurs associés.

FASCICULE VII – (TAP 6) Principes variationnels pour des systèmes de dimension finie.

FASCICULE VIII – (TAP 7) Méthodes 1.

FASCICULE IX – (TAP 8) Méthodes 2.

FASCICULE X – Étude et tracé des courbes du type $f(x, y) = 0$.