

## Chapitre II – Dérivabilité, théorème des fonctions implicites et applications

*Sommaire.* Les dérivées d'une fonction en un point permettent de comprendre son comportement au voisinage de ce point (formule de Taylor); dans la pratique ce sont souvent les dérivées d'ordre 1 ou 2 seulement qui sont utilisées. Par exemple, il existe un critère bien connu pour qu'une fonction d'une variable  $f(x)$  admette un extremum (maximum ou un minimum) local en un point  $a$  : il faut que la dérivée de  $f$  en  $a$  soit nulle, ce qui s'interprète géométriquement par le fait que la tangente au graphe de  $f$  au point  $(a, f(a))$  doit être horizontale. C'est sur ce critère que se basent la méthode des multiplicateurs de Lagrange pour la recherche d'extrema liés, et les équations d'Euler-Lagrange pour la recherche d'extrema dans des espaces de dimension non finie (calcul des variations).

### 1. Dérivabilité, différentiabilité

#### 1.1 NORME D'UNE APPLICATION LINÉAIRE

Soit  $A : \mathbb{R}^n \rightarrow \mathbb{R}^p$  une application linéaire. Les coefficients de la matrice de  $A$  par rapport aux bases naturelles  $e_j$ ,  $j = 1, \dots, n$ , et  $e_i$ ,  $i = 1, \dots, p$ , de  $\mathbb{R}^n$  et  $\mathbb{R}^p$  respectivement sont définis par :

$$A(e_j) = \sum_{i=1}^p a_{i,j} e_i \quad , \quad j = 1, \dots, n$$

et alors, si  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  est un vecteur de  $\mathbb{R}^n$ , on a :

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1,j} x_j \\ \vdots \\ \sum_{j=1}^n a_{p,j} x_j \end{pmatrix}_{i=1, \dots, p} .$$

On déduit facilement de cette expression de  $A$  que c'est une application continue.

**1.1 Définition – norme d'une application linéaire.** La norme de l'application linéaire  $A : \mathbb{R}^n \rightarrow \mathbb{R}^p$  relative à des normes données sur  $\mathbb{R}^n$  et  $\mathbb{R}^p$  est définie par :

$$\|A\| = \sup \left\{ \|A(x)\|_{\mathbb{R}^p} \mid \|x\|_{\mathbb{R}^n} \leq 1 \right\} .$$

---

Puisque  $A$  est continue et que  $\{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$  est compact, l'ensemble  $\{A(x) \mid \|x\| \leq 1\}$  est borné et donc cette définition a un sens. La valeur de  $\|A\|$  dépend des normes choisies sur  $\mathbb{R}^n$  et  $\mathbb{R}^p$ , même si cela n'est pas noté explicitement.

#### 1.2 Proposition.

(1) Pour tout  $x \in \mathbb{R}^n$  on a :

$$\|A(x)\|_{\mathbb{R}^p} \leq \|A\| \|x\|_{\mathbb{R}^n} .$$

(2)  $\|A\|$  définit une norme sur l'espace vectoriel  $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$  des applications linéaires de  $\mathbb{R}^n$  dans  $\mathbb{R}^p$ .

(3)  $\|A\| = \inf \{K \mid \|A(x)\| \leq K \cdot \|x\|, \forall x \in \mathbb{R}^n\}$

*Preuve:* Si  $x = 0$ , l'inégalité affirmée dans (1) devient  $0 \leq 0$ . Sinon,  $\left\| \frac{x}{\|x\|} \right\| = 1$  et alors :

$$\|A(x)\| = \left\| A\left(\frac{x}{\|x\|}\right) \right\| \|x\| \leq \|A\| \|x\| \quad .$$

Vérifions que  $\|A\|$  est une norme sur  $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$  (cf. définition **1.2** du chap. I).

- (1) Si  $\|A\| = 0$ , alors il suit de la première partie de cette preuve que  $A = 0$ .
- (2) Si  $\lambda \in \mathbb{R}$ ,  $\|\lambda A(x)\| = |\lambda| \|A(x)\|$ ,  $\forall x \in \mathbb{R}^n$  et il en suit que  $\|\lambda A\| = |\lambda| \|A\|$ .
- (3) Si  $A, B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$  on a :

$$\|(A+B)(x)\| = \|A(x) + B(x)\| \leq \|A(x)\| + \|B(x)\| \leq \|A\| + \|B\| \quad \forall x \in \mathbb{R}^n, \|x\| \leq 1$$

et donc  $\|A+B\| = \sup \{\|(A+B)(x)\|, \|x\| \leq 1\} \leq \|A\| + \|B\|$ .

Soit  $X = \{K \mid \|A(x)\| \leq K \cdot \|x\|, \forall x \in \mathbb{R}^n\}$ . D'après (1),  $\|A\| \in X$ . D'autre part, si  $K \in X$ ,  $\|x\| \leq 1 \Rightarrow \|A(x)\| \leq K \|x\| \leq K$ , et donc  $\|A\| = \sup \{\|A(x)\| \mid \|x\| \leq 1\} \leq K$ . Donc  $\|A\| = \inf(X)$ .

*q.e.d.*

Par exemple, si l'on munit  $\mathbb{R}^n$  et  $\mathbb{R}^p$  de la norme  $\|\cdot\|_\infty$ , on peut estimer la norme de  $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$  à l'aide des coefficients de la matrice  $(a_{i,j})_{\substack{i=1,\dots,n \\ j=1,\dots,p}}$  de  $A$ :

$$\left| \sum_{j=1}^n a_{i,j} x_j \right| \leq \left( \sum_j |a_{i,j}| \right) \|x\|_\infty, \quad \forall i = 1, \dots, p \Rightarrow \|A(x)\|_\infty \leq \sup_{i=1,\dots,p} \left\{ \sum_j |a_{i,j}| \right\} \|x\|_\infty$$

et donc  $\|A\| \leq \sup_{i=1,\dots,p} \left\{ \sum_j |a_{i,j}| \right\} \leq \sup \{|a_{i,j}|, i = 1, \dots, p, j = 1, \dots, n\}$ .

## 1.2 L'INÉGALITÉ FONDAMENTALE DE L'INTÉGRALE

Rappelons que si  $\phi : [a, b] \rightarrow \mathbb{R}$  est une fonction continue, on définit son intégrale comme étant le nombre réel qui est limite de sommes sur les partages de  $[a, b]$ :

$$(1-1) \quad \int_a^b \phi(x) dx = \lim_{\delta(P) \rightarrow 0} \left( \sum_{i=1, \dots, \ell(P)} \phi(x_i) \Delta x_i \right)$$

où  $P = \{a = x_0 < x_1 < \dots < x_k = b\}$  est un partage de  $[a, b]$ ,  $\ell(P) = k$ ,  $\Delta x_i = x_i - x_{i-1}$ ,  $i = 1, \dots, \ell(P)$  et  $\delta(P) = \sup \{\Delta x_i \mid i = 1, \dots, \ell(P)\}$ .

Dans le cas d'une application continue  $f : [a, b] \rightarrow \mathbb{R}^p$ , l'intégrale de  $f$  sur  $[a, b]$  est le vecteur de  $\mathbb{R}^p$  obtenu en intégrant les composantes  $f_i$ ,  $i = 1, \dots, p$ , de  $f$ :

$$\int_a^b f(x) dx = \left( \int_a^b f_1(x) dx, \dots, \int_a^b f_p(x) dx \right)$$

et en appliquant (1-1) aux  $f_i$  on voit que:

$$\int_a^b f(x) dx = \lim_{\delta(P) \rightarrow 0} \left( \sum_{i=1, \dots, \ell(P)} f(x_i) \Delta x_i \right) \quad .$$

Puisque

$$\left\| \sum_{i=1, \dots, \ell(P)} f(x_i) \Delta x_i \right\| \leq \sum_{i=1, \dots, \ell(P)} \|f(x_i)\| \Delta x_i$$

en faisant tendre  $\delta(P)$  vers 0, on obtient l'inégalité suivant, appelée inégalité fondamentale de l'intégrale:

$$(1-2) \quad \left\| \int_a^b f(x) dx \right\| \leq \int_a^b \|f(x)\| dx$$

## 1.3 DÉRIVABILITÉ, DIFFÉRENTIABILITÉ

Soient  $\Omega \subset \mathbb{R}^n$  un ouvert et  $f : \Omega \rightarrow \mathbb{R}^p$  une application.

**1.3 Définition - dérivabilité.** Soit  $a \in \Omega$ . On dit que  $f$  est dérivable (ou différentiable) au point  $a$  s'il existe une application linéaire  $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$  telle que, pour  $\|h\|$  assez petit :

$$f(a+h) = f(a) + A(h) + r(h) \quad , \quad \text{avec} \quad \lim_{h \rightarrow 0} \frac{r(h)}{\|h\|} = 0 \quad .$$

Autrement dit :

$$\left\| \frac{f(a+h) - f(a) - A(h)}{\|h\|} \right\| \rightarrow 0 \quad \text{si} \quad h \rightarrow 0$$

ou encore :

$$\|h\| \leq \delta_\varepsilon \Rightarrow \|f(a+h) - f(a) - A(h)\| \leq \varepsilon \cdot \|h\| \quad .$$

Si l'on pose  $\phi(h) = \frac{f(a+h) - f(a) - A(h)}{\|h\|}$ , cela veut encore dire que :

$$f(a+h) = f(a) + A(h) + \phi(h) \cdot \|h\| \quad , \quad \text{avec} \quad \phi(h) \rightarrow 0 \quad \text{si} \quad h \rightarrow 0$$

## 1.4 Remarques.

(1) Si  $f$  est dérivable au point  $a$ , elle est continue en ce point, car :

$$\lim_{h \rightarrow 0} f(a+h) = \lim_{h \rightarrow 0} (f(a) + A(h) + \phi(h) \cdot \|h\|) = f(a)$$

(2) Si  $f$  est dérivable au point  $a \in \Omega$  et  $v \in \mathbb{R}^n$ ,  $v \neq 0$ , la limite suivante, dans laquelle  $t \in \mathbb{R}$ , existe :

$$\frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \frac{f(a+t \cdot v) - f(a)}{t}$$

et elle est égale à  $A(v)$ . En effet :

$$\frac{f(a+tv) - f(a)}{t} = \frac{A(t \cdot v) + \phi(t \cdot v) \cdot |t| \cdot \|v\|}{t} = \frac{A(t \cdot v)}{t} \pm \phi(t \cdot v) \cdot \|v\| \rightarrow A(v) \quad \text{si} \quad t \rightarrow 0 \quad .$$

On appelle  $\frac{\partial f}{\partial v}(a)$  la dérivée de  $f$  dans la direction du vecteur  $v$  au point  $a$ . Puisque  $\frac{\partial f}{\partial v}(a) = A(v)$ , cela montre que  $A$  est entièrement déterminée par  $f$ ; on l'appelle la dérivée de  $f$  au point  $a$ , et on la note  $df_a$  ou encore  $f'(a)$ . En particulier, si  $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$  est linéaire, en tout point  $a \in \mathbb{R}^n$  elle coïncide avec sa propre dérivée en ce point :  $df_a = f, \forall a \in \mathbb{R}^n$ .

Notons que si  $v = 0$ , alors  $\frac{\partial f}{\partial v} = 0$ .

(3) Si  $f = (f_1, \dots, f_p) : \Omega \rightarrow \mathbb{R}^p$  est dérivable au point  $a \in \Omega$ , sa dérivée  $df_a = f'(a) : \mathbb{R}^n \rightarrow \mathbb{R}^p$  peut être représentée par une matrice, qu'on appelle matrice jacobienne; elle s'écrit :

$$\left( \frac{\partial f_i}{\partial x_j}(a) \right) \quad , \quad \text{où} \quad \frac{\partial f_i}{\partial x_j}(a) = \lim_{t \rightarrow 0} \frac{f_i(a+t \cdot e_j) - f_i(a)}{t} = \frac{\partial f_i}{\partial e_j}(a)$$

où  $e_j = \underbrace{(0, \dots, 0, 1, 0, \dots, 0)}_j$  dénote le  $j$ -ème vecteur de la base naturelle de  $\mathbb{R}^n$ . Les  $\frac{\partial f_i}{\partial x_j}(a)$  sont appelées

dérivées partielles de  $f$  au point  $a$ . Nous verrons un peu plus loin (1.10) le théorème suivant, qui permet de passer de l'existence des dérivées partielles de  $f$  à la dérivée de  $f$  au sens de la définition 1.3 : si les  $\frac{\partial f_i}{\partial x_j}(x)$  existent pour  $x$  dans un voisinage de  $a$ , et sont continues sur ce voisinage, alors  $f$  est dérivable.

(4) Dérivabilité et différentiabilité.

Autrefois on appelait *différentielle* d'une fonction  $f(x)$  (ou encore *différentielle totale*) l'accroissement des valeurs de  $f$  lorsqu'on donnait un accroissement "infinitement petit" à la variable  $x$ ; pour une fonction de deux variables, on écrivait :

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 \quad ;$$

le reste a disparu, parce que c'est un "infinitement petit d'ordre supérieur" par rapport aux  $dx_i$  (voir par exemple Ed. Goursat, Cours d'analyse mathématique, Gauthier-Villars, Paris (1910), page 52.) Le terme de dérivée était réservé aux dérivées partielles; dans les ouvrages contemporains, les notions de dérivabilité et différentiabilité sont équivalentes, et correspondent à notre définition **1.3**.

Dans le cas de fonctions d'une variable, on écrivait  $df = f'(x)dx$ , d'où l'on tire :  $\frac{df}{dx}(x) = f'(x)$ . Ce n'est pas très rigoureux, mais la notation  $\frac{df}{dx}(x)$  pour  $f'(x)$  peut être utile, parce que l'on explicite le nom de la variable par rapport à laquelle on dérive. Cette notation sera utilisée dans la formule d'Euler-Lagrange au § 4.

**1.5 Proposition – Dérivation de fonctions composées.** Soient  $\Omega \subset \mathbb{R}^m$ ,  $\Omega' \subset \mathbb{R}^n$  des ouverts,  $f : \Omega \rightarrow \mathbb{R}^n$ ,  $g : \Omega' \rightarrow \mathbb{R}^p$ , des applications, et supposons que  $f(\Omega) \subset \Omega'$ . Supposons que  $f$  soit dérivable au point  $a \in \Omega$ , et que  $g$  soit dérivable au point  $b = f(a)$ . Alors la composée  $g \circ f$  est dérivable au point  $a$  et sa dérivée en ce point est la composée de la dérivée de  $f$  en  $a$  avec la dérivée de  $g$  en  $b = f(a)$  :

$$d(g \circ f)_a = dg_{f(a)} \circ df_a \quad .$$

*Preuve:*

$$\begin{aligned} g(f(a+h)) - g(f(a)) &= g'(f(a)) \left( f(a+h) - f(a) \right) + \phi_g(f(a+h) - f(a)) \cdot \|f(a+h) - f(a)\| = \\ &= g'(f(a)) \left( f'(a)(h) \right) + \underbrace{g'_{f(a)}(\phi_f(h)) \cdot \|h\| + \phi_g \left( f(a+h) - f(a) \right) \cdot \left\| f'_a(h) + \phi_f(h) \cdot \|h\| \right\|}_{=\rho(h)} \end{aligned}$$

et

$$\|\rho(h)\| \leq \|g'(b)\| \cdot \|\phi_f(\|h\|)\| \cdot \|h\| + \|\phi_g(f(a+h) - f(a))\| \cdot (\|f'(a)\| + \|\phi_f(h)\|) \cdot \|h\|$$

d'où on déduit que  $\frac{\rho(h)}{\|h\|} \rightarrow 0$  si  $h \rightarrow 0$ , donc  $\rho(h)$  vérifie bien les conditions de terme de reste de la définition **1.3**.

*q.e.d.*

On aimerait avoir une estimation explicite de l'accroissement  $f(a+h) - f(a)$  en termes de  $\|h\|$ ; pour des fonctions à une variable, à valeurs dans  $\mathbb{R}$ , on connaît le théorème des accroissements finis, qui dit que si  $f : [a, b] \rightarrow \mathbb{R}$  est dérivable, alors il existe  $\xi \in [a, b]$  tel que

$$\spadesuit \quad f(b) - f(a) = f'(\xi)(b - a) \quad .$$

En général, on ne peut pas s'attendre à avoir des formules analogues; par exemple, si l'on prend l'application  $\varphi(t) = (t^2, t^3)$ ,  $\varphi'(t) = (2t, 3t)$  :

$$\varphi(1) - \varphi(0) = \varphi'(\xi)(1 - 0) = \varphi'(\xi) \Rightarrow (1, 1) = (2\xi, 3\xi^2) \Rightarrow \xi = 1/2 \text{ et } \xi = 1/3$$

il n'y a donc pas de solutions. Par contre, on déduit de  $\spadesuit$  que, si  $|f'(\xi)| \leq M$ ,  $\xi \in [a, b]$ , alors  $|f(b) - f(a)| \leq M(b - a)$  et cette inégalité se généralise, comme nous allons voir maintenant (théorème **1.8**.)

**1.6 Définition - application de classe  $\mathcal{C}^1$ .** Soit  $f : \Omega \rightarrow \mathbb{R}^p$  dérivable en tout point  $a \in \Omega$ ; en associant à tout point  $a \in \Omega$  la dérivée en ce point  $df_a \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$  on définit une application  $df : \Omega \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ . On dit que  $f$  est de classe  $\mathcal{C}^1$ , ou 1 fois continûment dérivable, si l'application  $df : \Omega \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$  est continue, c'est-à-dire si toutes les dérivées partielles  $\frac{\partial f_i}{\partial x_j}(x)$ ,  $i = 1, \dots, p$ ,  $j = 1, \dots, n$  sont continues sur  $\Omega$ .

**1.7 Proposition.** Soit  $f : \Omega \rightarrow \mathbb{R}^p$ ,  $\Omega \subset \mathbb{R}^n$ , de classe  $\mathcal{C}^1$ ; soit  $v \in \mathbb{R}^n$  et supposons que le segment  $\{a + tv \mid 0 \leq t \leq 1\}$  soit contenu dans  $\Omega$ . Alors

$$f(a + v) - f(a) = \int_0^1 \frac{\partial f}{\partial v}(a + tv) dt$$

*Preuve:* On pose  $\varphi(t) = f(a + tv)$ , pour  $0 \leq t \leq 1$ , de sorte que  $\varphi(1) - \varphi(0) = f(a + v) - f(a)$ . Il suit de 1.4(2) que

$$\lim_{s \rightarrow 0} \left( \frac{\varphi(t + s) - \varphi(t)}{s} \right) = \lim_{s \rightarrow 0} \left( \frac{f(a + (t + s)v) - f(a + tv)}{s} \right) = \frac{\partial f}{\partial v}(a + tv)$$

et en particulier  $\varphi$  est aussi de classe  $\mathcal{C}^1$ . Donc

$$f(a + v) - f(a) = \varphi(1) - \varphi(0) = \int_0^1 \varphi'(t) dt = \int_0^1 \frac{\partial f}{\partial v}(a + tv) dt \quad .$$

*q.e.d.*

**1.8 Théorème des accroissements finis.** Soit  $f : \Omega \rightarrow \mathbb{R}^p$ ,  $\Omega \subset \mathbb{R}^p$ ,  $\Omega$  ouvert; supposons que  $f$  soit de classe  $\mathcal{C}^1$ . Soient  $a, b \in \Omega$ ; supposons que le segment  $[a, b] = \{a + t(b - a) \mid 0 \leq t \leq 1\}$  soit contenu dans  $\Omega$ . Alors :

$$\|f(b) - f(a)\| \leq \sup \{ \|df_{a+t(b-a)}\| \mid 0 \leq t \leq 1 \} \cdot \|b - a\|$$

*Preuve:* Tout d'abord, du fait que  $f$  est  $\mathcal{C}^1$ , il suit que l'application  $t \mapsto df_{a+t(b-a)}$  est continue; comme  $[0, 1]$  est compact, le sup  $\{ \|df_{a+t(b-a)}\| \mid 0 \leq t \leq 1 \}$  a un sens.

Posons  $v = b - a$ . Il résulte de 1.7 et de l'inégalité fondamentale de l'intégrale que

$$\begin{aligned} \|f(b) - f(a)\| &= \left\| \int_0^1 \frac{\partial f}{\partial v}(a + tv) dt \right\| \leq \int_0^1 \left\| \frac{\partial f}{\partial v}(a + tv) \right\| dt \\ &\leq \sup \{ \|df_{a+tv}(v)\| \mid 0 \leq t \leq 1 \} \leq \sup \{ \|df_{a+tv}\| \mid 0 \leq t \leq 1 \} \cdot \|v\| \end{aligned}$$

*q.e.d.*

Une conséquence immédiate :

**1.9 Corollaire.** Soient  $f : \Omega \rightarrow F$ ,  $a$  et  $b$  comme dans le théorème précédent; supposons que  $\overline{B}(a, r) \subset \Omega$  et que  $\|df_x\| \leq M$ ,  $\forall x \in \overline{B}(a, r)$ . Alors si  $b \in \overline{B}(a, r)$  on a :

$$\|f(b) - f(a)\| \leq M \|b - a\| \quad .$$

■

Soient  $\Omega \subset \mathbb{R}^m \times \mathbb{R}^n$ ,  $f : \Omega \rightarrow \mathbb{R}^p$ ,  $\Omega$  un ouvert; notons par  $(x, y)$  les éléments de  $\mathbb{R}^m \times \mathbb{R}^n$ , avec  $x \in \mathbb{R}^m$ ,  $y \in \mathbb{R}^n$ . Si  $(a, b) \in \Omega$ , on note par  $\frac{\partial f}{\partial x}(a, b)$  la dérivée au point  $a$  de l'application  $\xi \mapsto f(a + \xi, b)$ , et par  $\frac{\partial f}{\partial y}(a, b)$  la dérivée au point  $b$  de l'application  $\xi \mapsto f(a, b + \xi)$ . La proposition suivante, qui est utile pour le calcul de dérivées, généralise le théorème mentionné sous 1.4(3).

**1.10 Proposition.** *Supposons que les dérivées  $\frac{\partial f}{\partial y}(a,b)$  et  $\frac{\partial f}{\partial x}(x,y)$ ,  $(x,y)$  dans un voisinage de  $(a,b)$ , existent, et que l'application  $(x,y) \mapsto \frac{\partial f}{\partial x}(x,y)$  soit continue dans un voisinage du point  $(a,b)$ . Alors  $f$  est dérivable au point  $(a,b)$  et :*

$$df_{(a,b)}(v,w) = \frac{\partial f}{\partial x}(a,b)(v) + \frac{\partial f}{\partial y}(a,b)(w) \quad .$$

*Preuve:*

$$\begin{aligned} & f((a,b) + (v,w)) - f(a,b) - \frac{\partial f}{\partial x}(a,b)(v) - \frac{\partial f}{\partial y}(a,b)(w) = \\ & \underbrace{f((a+v, b+w)) - f((a, b+w)) - \frac{\partial f}{\partial x}(a,b)(v) + f(a, b+w)}_{\text{I}} - \underbrace{f(a, b+w) - f(a,b) - \frac{\partial f}{\partial y}(a,b)(w)}_{\text{II}} \quad . \end{aligned}$$

Soit  $\varepsilon > 0$ ; il s'agit de montrer que  $\|\text{I} + \text{II}\| \leq \varepsilon \|(v,w)\|$  si  $\|(v,w)\|$  est assez petit.

Posons  $\varphi(v) = f(a+v, b+w) - \frac{\partial f}{\partial x}(a,b)(v)$ ; alors  $\varphi(v) - \varphi(0) = \text{I}$  et  $d\varphi_v = \frac{\partial f}{\partial x}(a+v, b+w) - \frac{\partial f}{\partial x}(a,b)$ . On a que  $d\varphi_0 = 0$ , et alors si  $\|(v,w)\|$  est assez petite,  $\|d\varphi_v\| < \varepsilon$ . Il suit de **1.9** que  $\|\text{I}\| \leq \varepsilon \|v\|$  pour  $\|(v,w)\|$  assez petite.

L'existence de  $\frac{\partial f}{\partial y}(a,b)$  implique directement que  $\|\text{II}\| \leq \varepsilon \|w\|$  si  $\|w\|$  est assez petite.

*q.e.d.*

Cette proposition se généralise sans autre aux cas de plus de 2 facteurs : si  $f : \Omega \rightarrow \mathbb{R}^p$ ,  $\Omega$  ouvert de  $\mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_k}$ , espace vectoriel normé,  $x = (x_1, \dots, x_k) \in \Omega$ ,  $x_i \in \mathbb{R}^{n_i}$ , on pose

$$f_j^x(\xi_j) = f(x_1, \dots, x_{j-1}, \xi_j, x_{j+1}, \dots, x_k) \quad ,$$

qui est définie pour  $\xi_j \in (\{x_1\} \times \dots \times \{x_{j-1}\} \times \mathbb{R}^{n_j} \times \{x_{j+1}\} \times \dots \times \{x_k\}) \cap \Omega$  et si  $f_j^x$  est dérivable en  $\xi_j = x_j$ , on dénote sa dérivée par  $\partial_j f_x$ . Il suit de 1.10, par induction sur  $k$ , que si  $\partial_1 f_a$  existe et que les  $\partial_j f_x$ ,  $j = 2, \dots, k$ , sont continues au voisinage de  $a$ , alors  $f$  est dérivable et

$$d_a f(h_1, \dots, h_k) = \partial_1 f_a(h_1) + \dots + \partial_k f_a(h_k) \quad .$$

Dans le cas d'une application  $f : \Omega \rightarrow \mathbb{R}$ ,  $\Omega$  ouvert de  $\mathbb{R}^n$ , en utilisant la décomposition  $\mathbb{R}^n = \underbrace{\mathbb{R} \times \dots \times \mathbb{R}}_n$ ,

on a que de  $\partial_i f(x) = \frac{\partial f}{\partial x_i}(x)$ .

**1.11 Corollaire.** *Soit  $f : \Omega \rightarrow \mathbb{R}^p$ ,  $f = (f_1, \dots, f_p)$ ,  $a \in \Omega$ , et supposons que toutes les dérivées partielles  $\frac{\partial f}{\partial x_i}(x)$  existent et soient continues pour  $x \in \Omega$ . Alors  $f$  est dérivable en tout point  $a \in \Omega$  et la matrice de  $df_a$  est égale à*

$$\left( \frac{\partial f_i}{\partial x_j}(a) \right)_{i=1, \dots, p, j=1, \dots, n}$$

■

En particulier, si  $f : \Omega \rightarrow \mathbb{R}^p$  et les dérivées  $\frac{\partial f_i}{\partial x_j}(x)$  existent et sont continues sur  $\Omega$ ,  $f$  est  $\mathcal{C}^1$  : la condition que  $f$  soit dérivable demandée dans la définition **1.6** est automatiquement satisfaite.

La proposition ci-dessous établit une propriété élémentaire, mais efficace, à laquelle nous aurons recours pour des applications au § suivant.

Soit  $f : \Omega \rightarrow \mathbb{R}$  une application,  $\Omega \subset \mathbb{R}^n$  ouvert et  $a \in \Omega$ . On dit que  $a$  est un minimum (respectivement maximum) local de  $f$  s'il existe un voisinage  $V$  de  $a$  tel que  $f(x) \geq f(a)$  si  $x \in V$  (respectivement  $f(x) \leq f(a)$ ). On dit que  $a$  est un extremum local si c'est un minimum ou un maximum local.

**1.12 Proposition.** Soit  $f$  comme ci-dessus. Supposons que  $f$  soit dérivable au point  $a$ . Alors, si  $a$  est un extremum local de  $f$ , on a  $df_a = 0$ .

*Preuve:* Supposons qu'il s'agisse d'un maximum local; soit  $v \in E$  et  $t > 0$  assez petit. Alors on a :

$$\frac{f(a+tv) - f(a)}{t} \leq 0 \leq \frac{f(a-tv) - f(a)}{-t}$$

et comme les 2 fractions ont pour limite  $\frac{\partial f}{\partial v}(a)$ , on doit avoir  $\frac{\partial f}{\partial v}(a) = df_a(v) = 0$ , ceci pour tout  $v \in E$ . Le cas d'un minimum local est semblable.

*q.e.d.*

#### 1.4 DÉRIVÉES D'ORDRE SUPÉRIEUR ET FORMULE DE TAYLOR

Nous allons définir la notion d'application  $f : \Omega \rightarrow \mathbb{R}^p$  de classe  $\mathcal{C}^k$ ,  $k \in \mathbb{N}$ , et ses dérivées d'ordre  $\ell \leq k$  par induction sur  $k$  :

**1.13 Définition.** On dit que  $f$  est de classe  $\mathcal{C}^0$  si  $f$  est continue, et dans ce cas il n'y a point de dérivées à définir.

Au §1.2 on a introduit la notion d'application de classe  $\mathcal{C}^1$  (définition 1.6). Cela revient à supposer que les dérivées partielles  $\frac{\partial f}{\partial x_i}(a)$  existent pour tout  $a \in \Omega$  et sont continues, ce qui entraîne par le corollaire 1.11 que  $f$  est dérivable en tout point  $a \in \Omega$ .

Supposons d'avoir défini la notion d'application de classe  $\mathcal{C}^\ell$  et les applications

$$\frac{\partial^\ell f}{\partial x_{i_1} \dots \partial x_{i_\ell}} : \Omega \rightarrow \mathbb{R}^p \quad , \quad 1 \leq i_h \leq n, h = 1, \dots, \ell$$

pour tout  $\ell \leq k-1$ . On dira que alors que  $f$  est de classe  $\mathcal{C}^k$  si les fonctions ci-dessus sont de classe  $\mathcal{C}^1$  et on pose:

$$\forall a \in \Omega \quad , \quad \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}(a) = \frac{\partial}{\partial x_{i_1}} \left( \frac{\partial^{k-1} f}{\partial x_{i_2} \dots \partial x_{i_k}} \right)(a)$$

On dit encore que  $f : \Omega \rightarrow \mathbb{R}^p$  est  $\mathcal{C}^\infty$  si  $f$  est de classe  $\mathcal{C}^k$  pour tout  $k \geq 0$

Dans le cas d'une application  $g : ]a, b[ \rightarrow \mathbb{R}^p$  de classe  $\mathcal{C}^k$ , il n'y a qu'une dérivée d'ordre  $\ell$ ,  $1 \leq \ell \leq k$ , celle qui correspond à la suite  $i_1 = \dots = i_k = 1$ . On la note  $g^{(\ell)}(a)$ , ou encore  $\frac{d^\ell g}{dt^\ell}(a)$ , si  $t$  denote la variable du domaine de définition de  $g$ . On note aussi  $g^{(0)}(a) = g(a)$ .

On peut généraliser aussi la notion de dérivée directionnelle; si  $v_1, \dots, v_k \in \mathbb{R}^n$ , on définit par induction sur  $k$ :

$$\frac{\partial^k f}{\partial v_1 \dots \partial v_k}(a) = \frac{\partial}{\partial v_1} \left( \frac{\partial^{k-1} f}{\partial v_2 \dots \partial v_k} \right)(a) \quad .$$

Il suit du fait que  $\frac{\partial f}{\partial v}(a)$  est linéaire en  $v$  que pour tout  $i = 1, \dots, k$ ,  $\frac{\partial^k f}{\partial v_1 \dots \partial v_i \dots \partial v_k}(a)$  est linéaire par rapport à  $v_i$ .

La définition suivante sera utile pour mieux comprendre les dérivées d'ordre supérieur.

**1.14 Définition.** Soit  $f : \Omega \rightarrow \mathbb{R}^p$  une application (quelconque). Pour tout  $a \in \Omega$  et  $v \in \mathbb{R}^n$  tels que  $[a, a+v] \subset \Omega$  on pose:

$$\Delta_v f(a) = f(a+v) - f(a)$$

On peut regarder  $\Delta_v f$  comme un opérateur dépendant du paramètre  $v$ : à l'application  $f$  il associe une nouvelle application, celle qui à  $a$  fait correspondre l'accroissement de  $f$  en  $a$  relatif à l'accroissement  $v$  de la variable. La nouvelle application  $\Delta_v f$  ne sera pas toujours définie, mais puisque  $\Omega$  est ouvert, pour tout

$a \in \Omega$  il existe un  $r > 0$  tel que  $a + v \in \Omega$  si  $\|v\| < r$ , et donc  $\Delta_v f$  sera bien définie pour tout  $v$  assez petit. Dans ce qui suit on supposera, sans le dire explicitement chaque fois, que les accroissements  $v_i$  sont assez petits pour que les opérateurs que l'on écrira soient bien définis.

Considérons l'expression

$$\begin{aligned} \Delta_{v_2} \Delta_{v_1} f(a) &= \Delta_{v_2} (f(a + v_1) - f(a)) = f(a + v_1 + v_2) - f(a + v_2) - (f(a + v_1) - f(a)) \\ &= f(a + v_1 + v_2) - f(a + v_1) - f(a + v_2) + f(a) \quad . \end{aligned}$$

Elle représente l'accroissement (relatif à  $v_2$ ) de l'accroissement (relatif à  $v_1$ ) de  $f$  et nous sera utile pour comprendre la 2-ième dérivée comme "taux d'accroissement du taux d'accroissement"; de même, à l'aide de l'opérateur  $\Delta$  itéré  $k$  fois on obtiendra une expression de la  $k$ -ième dérivée (voir corollaire 1.16).

Notons que l'expression ci-dessus est symétrique en  $(v_1, v_2)$  :

$$\Delta_{v_2} \Delta_{v_1} f(a) = \Delta_{v_1} \Delta_{v_2} f(a) \quad .$$

**1.15 Proposition.** Soit  $f : \Omega \rightarrow \mathbb{R}^p$  de classe  $\mathcal{C}^k$ ; on a:

$$\Delta_{v_k} \Delta_{v_{k-1}} \dots \Delta_{v_1} f(a) = \int_0^1 \dots \left( \int_0^1 \frac{\partial^k f}{\partial v_k \dots \partial v_1} (a + t_k v_k + \dots + t_1 v_1) dt_k \right) \dots dt_1 .$$

*Preuve:* On procède par induction sur  $k$ . Si  $k = 1$ , cela résulte de la proposition 1.7.

Supposons que la formule soit vraie pour  $k - 1$  et posons  $w_{k-1} = t_{k-1} v_{k-1} + \dots + t_1 v_1$ . On a:

$$\begin{aligned} \Delta_{v_k} \dots \Delta_{v_1} f(a) &= \Delta_{v_{k-1}} \dots \Delta_{v_1} f(a + v_k) - \Delta_{v_{k-1}} \dots \Delta_{v_1} f(a) \stackrel{\text{hyp. induction}}{=} \\ &= \int_0^1 \dots \int_0^1 \frac{\partial^{k-1} f}{\partial v_{k-1} \dots \partial v_1} (a + v_k + w_{k-1}) dt_{k-1} \dots dt_1 - \int_0^1 \dots \int_0^1 \frac{\partial^{k-1} f}{\partial v_{k-1} \dots \partial v_1} (a + w_{k-1}) dt_{k-1} \dots dt_1 \\ &= \int_0^1 \dots \int_0^1 \left( \frac{\partial^{k-1} f}{\partial v_{k-1} \dots \partial v_1} (a + v_k + w_{k-1}) - \frac{\partial^{k-1} f}{\partial v_{k-1} \dots \partial v_1} (a + w_{k-1}) \right) dt_{k-1} \dots dt_1 \\ &= \int_0^1 \dots \int_0^1 \frac{\partial^k f}{\partial v_k \dots \partial v_1} (a + t_k v_k + \dots + t_1 v_1) dt_k \dots dt_1 \end{aligned}$$

où la dernière égalité résulte du cas  $k = 1$  appliqué à  $\frac{\partial^k f}{\partial v_{k-1} \dots \partial v_1} (a + w_{k-1})$ .

*q.e.d.*

Le corollaire suivant exprime une dérivée d'ordre supérieur de  $f$  directement à partir de  $f$ , plutôt que de passer par des dérivations successives comme il est fait dans la définition 1.13:

**1.16 Corollaire.** Si  $f$  est de classe  $\mathcal{C}^k$ ,

$$\frac{\partial^k f}{\partial v_k \dots \partial v_1} (a) = \lim_{s_1, \dots, s_k \rightarrow 0} \left( \frac{\Delta_{s_k v_k} \dots \Delta_{s_1 v_1} f(a)}{s_k \dots s_1} \right)$$

*Preuve:* D'après 1.15:

$$\begin{aligned} \frac{\Delta_{s_k v_k} \dots \Delta_{s_1 v_1} f(a)}{s_k \dots s_1} &= \frac{1}{s_k \dots s_1} \int_0^1 \dots \int_0^1 \frac{\partial^k f}{\partial (s_k v_k) \dots \partial (s_1 v_1)} (a + s_k t_k v_k + \dots + s_1 t_1 v_1) dt_k \dots dt_1 \\ &= \int_0^1 \dots \int_0^1 \frac{\partial^k f}{\partial v_k \dots \partial v_1} (a + s_k t_k v_k + \dots + s_1 t_1 v_1) dt_k \dots dt_1 \end{aligned}$$



où la dernière égalité utilise le fait que  $\frac{\partial^k f}{\partial(s_k v_k) \dots \partial(s_1 v_1)}(x) = s_1 \dots s_k \frac{\partial^k f}{\partial v_k \dots \partial v_1}(x)$  (linéarité par rapport à  $v_i$ ,  $i = 1, \dots, k$ ). Puisque  $\frac{\partial^k f}{\partial v_k \dots \partial v_1}(x)$  est continue en  $x$ , il suffit de poser  $s_i = 0$  dans la dernière expression pour calculer la limite cherchée.

*q.e.d.*

**1.17 Corollaire.** *Si  $f$  est de classe  $\mathcal{C}^k$ ,  $a \in \Omega$ , alors  $\frac{\partial^k f}{\partial v_k \dots \partial v_1}(a)$  est symétrique en  $v_1, \dots, v_k$ , ce qui veut dire que si  $\sigma : \{1, \dots, k\} \rightarrow \{1, \dots, k\}$  est une bijection, alors*

$$\frac{\partial^k f}{\partial v_k \dots \partial v_1}(a) = \frac{\partial^k f}{\partial v_{\sigma(k)} \dots \partial v_{\sigma(1)}}(a) \quad .$$

*Preuve:* On a déjà remarqué que  $\Delta_{v_2} \Delta_{v_1} f(a)$  était symétrique en  $v_1, v_2$ . Il en suit que  $\Delta_{v_k} \dots \Delta_{v_1} f(a)$  est symétrique en  $v_1, \dots, v_k$ , et le corollaire suit alors de **1.16**.

*q.e.d.*

La formule suivante, due à Leibniz, exprime la  $\ell$ ème dérivée de 2 fonctions d'une variable en termes des dérivées de chacune des fonctions :

**1.18 Proposition.** *Soient  $\alpha, \beta : ]t_0 - r, t_0 + r[ \rightarrow \mathbb{R}$  deux fonctions de classe  $\mathcal{C}^\ell$ . Alors la  $\ell$ -ième dérivée de  $\alpha(t) \cdot \beta(t)$  a pour expression :*

$$(\alpha(t) \cdot \beta(t))^{(\ell)}(t) = \sum_{h=0}^{\ell} \binom{\ell}{h} \alpha^{(h)}(t) \cdot \beta^{(\ell-h)}(t)$$

*Preuve:* Par induction sur  $\ell$ . Pour  $\ell = 1$ , la formule se réduit à l'expression bien connue :

$$(\alpha(t) \cdot \beta(t))'(t) = \alpha'(t) \cdot \beta(t) + \alpha(t) \cdot \beta'(t) \quad .$$

Supposons d'avoir montré que

$$(\alpha(t) \cdot \beta(t))^{(\ell-1)}(t) = \sum_{h=0}^{\ell-1} \binom{\ell-1}{h} \alpha^{(h)}(t) \cdot \beta^{(\ell-1-h)}(t) \quad .$$

On dérive en appliquant la formule pour  $\ell = 1$  à chaque terme de la somme :

$$\begin{aligned} (\alpha(t) \cdot \beta(t))'(t) &= \sum_{h=0}^{\ell-1} \binom{\ell-1}{h} \left( \alpha^{(h+1)}(t) \cdot \beta^{(\ell-1-h)}(t) + \alpha^{(h)}(t) \cdot \beta^{(\ell-h)}(t) \right) \\ &= \sum_{h=0}^{\ell} \alpha^{(h)}(t) \cdot \beta^{(\ell-h)}(t) \left( \binom{\ell-1}{h-1} + \binom{\ell-1}{h} \right) \end{aligned}$$

et on conclut en remarquant que :

$$\begin{aligned} \binom{\ell-1}{h-1} + \binom{\ell-1}{h} &= \frac{(\ell-1)!}{(h-1)!(\ell-1-(h-1))!} + \frac{(\ell-1)!}{h!(\ell-1-h)!} \\ &= \frac{(\ell-1)!h}{h!(\ell-h)!} + \frac{(\ell-1)!(\ell-h)}{h!(\ell-h)!} = \frac{(\ell-1)!h + (\ell-1)!(\ell-h)}{h!(\ell-h)!} = \binom{\ell}{h} \quad . \end{aligned}$$

*q.e.d.*

Suivent des résultats préliminaires à la formule de Taylor.

**1.19 Proposition.** Soit  $g : ]a, b[ \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^{k+1}$ ,  $t > 0$ ,  $[0, t] \subset ]a, b[$ . Supposons que

$$g(0) = g'(0) = \dots = g^{(k)}(0) = 0 \quad .$$

Alors

$$(1-3) \quad g(t) = t^{k+1} \cdot \gamma(t)$$

où

$$(1-4) \quad \gamma(t) = \frac{1}{k!} \int_0^1 g^{(k+1)}(t \cdot u)(1-u)^k du$$

et en particulier  $\gamma(0) = \frac{1}{(k+1)!} g^{(k+1)}(0)$ .

*Preuve:* Montrons l'affirmation par induction sur  $k$ . Pour  $k=0$ , on a que  $g(0) = 0$  et donc :

$$g(t) = \int_0^t g'(s) ds \quad .$$

On fait le changement de variable  $s = u \cdot t$  dans l'intégrale :

$$g(t) = \int_0^1 g'(t \cdot u) \cdot t du = t \int_0^1 g'(t \cdot u) du \quad .$$

Supposons que  $g(0) = g'(0) = \dots = g^{(k)}(0) = 0$  et que (1-3) et (1-4) soient vraie pour  $k-1$ :

$$g(t) = \frac{t^k}{(k-1)!} \int_0^1 g^{(k)}(t \cdot u)(1-u)^{k-1} du$$

On intègre par partie en remarquant que

$$(1-u)^{k-1} = \left( -\frac{1-u}{k} \right)'$$

et compte tenu du fait que  $g^{(k)}(0) = 0$  on obtient:

$$\begin{aligned} g(t) &= \frac{t^k}{(k-1)!} \left( \int_0^1 g^{(k)}(t \cdot u) \left( -\frac{1-u}{k} \right)' du \right) = \\ &= \frac{t^k}{(k-1)!} \left( \underbrace{g^{(k)}(t \cdot u) \cdot \left( -\frac{1-u}{k} \right)}_{=0} \Big|_0^t - \int_0^1 g^{(k+1)}(t \cdot u) \cdot t \cdot \left( -\frac{1-u}{k} \right) du \right) = \frac{t^{k+1}}{k!} \int_0^1 g^{(k+1)}(t \cdot u)(1-u)^k du . \end{aligned}$$

Puisque  $\int_0^1 (1-u)^k du = \int_0^1 u^k du = \frac{1}{k+1}$ , on a bien que  $\gamma(0) = \frac{1}{(k+1)!} g^{(k+1)}(0)$ .

*q.e.d.*

**1.20 Corollaire – formule de Taylor pour les fonctions d'une variable.** Soit  $\varphi : ]a-r, a+r[ \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^{k+1}$ . Alors, pour  $|t| < r$  :

$$\varphi(a+t) = \varphi(a) + \sum_{\ell=1}^k \frac{\varphi^{(\ell)}(a)}{\ell!} t^\ell + \frac{t^{k+1}}{k!} \int_0^1 \varphi^{(k+1)}(a+t \cdot u)(1-u)^k du \quad .$$

*Preuve:* On pose

$$g(t) = \varphi(a+t) - \varphi(a) - \left( \sum_{\ell=0}^k \frac{\varphi^{(\ell)}(0)}{\ell!} t^\ell \right)$$

de sorte que  $g^{(\ell)}(0) = 0, \forall \ell \leq k$  et  $g^{(k+1)}(t) = \phi^{k+1}(t)$ , puis on applique **1.19** à  $g$ .

*q.e.d.*

Le cas de fonctions à plusieurs variables pourra se ramener au cas d'une variable grâce au lemme suivant. Tout d'abord on introduit une notation permettant de regrouper les dérivées supérieures qui ne diffèrent que par l'ordre des dérivations. Soit  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$  un vecteur à coefficients entiers. On pose:

$$|\alpha| = \alpha_1 + \dots + \alpha_n \quad , \quad \alpha! = (\alpha_1!) \cdot \dots \cdot (\alpha_n!)$$

et si  $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ , on pose  $h^\alpha = h_1^{\alpha_1} \cdot \dots \cdot h_n^{\alpha_n} \in \mathbb{R}$ .

Soit  $\alpha \in \mathbb{N}^n, |\alpha| = k$ ; on peut lui associer la suite de  $k$  entiers  $i(\alpha) = (i(\alpha)_1, \dots, i(\alpha)_k)$  définie ainsi :

$$i(\alpha) = \underbrace{(1, \dots, 1)}_{\alpha_1 \text{ fois}}, \underbrace{(2, \dots, 2)}_{\alpha_2 \text{ fois}}, \dots, \underbrace{(n, \dots, n)}_{\alpha_n \text{ fois}} \quad .$$

Si  $f : \Omega \rightarrow \mathbb{R}^p$  est de classe  $\mathcal{C}^k$ , on pose :

$$\frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a) = \frac{\partial^{|\alpha|} f}{\partial x_{i(\alpha)_k} \dots \partial x_{i(\alpha)_1}}(a)$$

ce qui revient à dériver  $f$   $\alpha_1$ -fois par rapport à  $x_1$ ,  $\alpha_2$ -fois par rapport à  $x_2$ , et ainsi de suite, au point  $a$ .

**1.21 Lemme.** Soit  $f : \Omega \rightarrow \mathbb{R}^p$  de classe  $\mathcal{C}^k$ ,  $h \in \mathbb{R}^n$  tel que  $\{a + th \mid 0 \leq t \leq 1\} \subset \Omega$ . Posons  $\varphi(t) = f(a + th), 0 \leq t \leq 1$ . Alors,  $\forall \ell \leq k$  on a:

$$\varphi^{(\ell)}(t) = \sum_{1 \leq i_1, \dots, i_\ell \leq n} \frac{\partial^\ell f}{\partial x_{i_\ell} \dots \partial x_{i_1}}(a+th) h_{i_1} \dots h_{i_\ell} = \sum_{\alpha \in \mathbb{N}^n, |\alpha|=\ell} \frac{\ell!}{\alpha!} \frac{\partial^\ell f}{\partial x^\alpha}(a+th) h^\alpha$$

*Preuve:* On montre d'abord la première égalité par induction sur  $\ell$ .

Pour  $\ell = 1$ , puisque  $h = \sum_{i=1}^n h_i e_i$ , où les  $e_i$  dénotent les vecteurs de la base standard de  $\mathbb{R}^n$ , il suit de la définition de  $\frac{\partial f}{\partial h}$  et du fait qu'elle est linéaire en  $h$  que:

$$\varphi'(t) = \frac{\partial f}{\partial h}(a+th) = \sum_{1 \leq i \leq n} \frac{\partial f}{\partial x_i}(a+th) h_i \quad .$$

Supposons d'avoir démontré que

$$\varphi^{(\ell-1)}(t) = \sum_{1 \leq i_1, \dots, i_{\ell-1} \leq n} \frac{\partial^{\ell-1} f}{\partial x_{i_{\ell-1}} \dots \partial x_{i_1}}(a+th) h_{i_1} \dots h_{i_{\ell-1}} \quad .$$

Alors

$$\begin{aligned} \varphi^{(\ell)}(t) &= \sum_{i=1}^n \frac{\partial}{\partial x_i} \left( \sum_{1 \leq i_1, \dots, i_{\ell-1} \leq n} \left( \frac{\partial^{\ell-1} f}{\partial x_{i_{\ell-1}} \dots \partial x_{i_1}}(a+th) h_{i_1} \dots h_{i_{\ell-1}} \right) \right) h_i = \\ &\quad \sum_{1 \leq i_1, \dots, i_\ell \leq n} \frac{\partial^\ell f}{\partial x_{i_\ell} \dots \partial x_{i_1}}(a+th) h_{i_1} \dots h_{i_\ell} \quad . \end{aligned}$$

Dans la somme ci-dessus, la suite d'indices

$$(i_1, \dots, i_\ell) = \underbrace{(1, \dots, 1)}_{\alpha_1\text{-fois}} \underbrace{(2, \dots, 2)}_{\alpha_2\text{-fois}} \dots \underbrace{(n, \dots, n)}_{\alpha_n\text{-fois}}$$

avec  $\alpha_1 + \dots + \alpha_n = \ell$  apparaît  $\frac{\ell!}{\alpha!}$  fois à permutation près; en effet, il faut diviser le nombre  $\ell!$  de toutes les permutations de la suite  $\underbrace{(1, \dots, 1)}_{\alpha_1\text{-fois}} \underbrace{(2, \dots, 2)}_{\alpha_2\text{-fois}} \dots \underbrace{(n, \dots, n)}_{\alpha_n\text{-fois}}$  par le nombre de celles qui la laissent invariante, qui est  $(\alpha_1!) \dots (\alpha_n!) = \alpha!$ . La dernière égalité énoncée suit alors du corollaire **1.16** (symétrie des dérivées d'ordre supérieure) .

*q.e.d.*

**1.22 Formule de Taylor.** Soit  $f = (f_1, \dots, f_p) : \Omega \rightarrow \mathbb{R}^p$  de classe  $\mathcal{C}^{k+1}$ ,  $a \in \Omega$  et  $h \in \mathbb{R}^n$  tel que  $\{a + th \mid 0 \leq t \leq 1\} \subset \Omega$ . Alors:

$$f(a + h) = f(a) + \sum_{\ell=1}^k \left( \sum_{|\alpha|=\ell} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a) \cdot h^\alpha \right) + r_k(a, h)$$

où le terme  $r_k(a, h)$ , appelé terme de reste, s'écrit :

$$(1-5) \quad r_k(a, h) = (k+1) \sum_{|\alpha|=k+1} \frac{1}{\alpha!} h^\alpha \left( \int_0^1 \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a+uh) (1-u)^k du \right) .$$

Si  $R > 0$  est tel que  $\overline{B(a, R)} \subset \Omega$  et  $\|h\| \leq R$ , on a:

$$(1-6) \quad \|r_k(a, h)\| \leq C_k \|h\|^{k+1}$$

où  $C_k$  est une constante, et en particulier  $\lim_{h \rightarrow 0} \frac{r_k(a, h)}{\|h\|^k} = 0$ .

*Preuve:* On pose  $\varphi(t) = (\varphi_1(t), \dots, \varphi_p(t)) = f(a + th)$ ,  $0 \leq t \leq 1$ . La formule (1-5) se déduit alors du corollaire **1.20** appliqué aux composantes de  $\varphi(t)$  et du lemme **1.21** appliqués aux composantes de  $f$ .

D'autre part, en reprenant (1-5), on a que

$$\|r_k(a, h)\| = \left\| \frac{1}{k!} \sum_{1 \leq i_1, \dots, i_{k+1} \leq n} h^\alpha \left( \int_0^1 \frac{\partial^{k+1} f}{\partial x_{i_{k+1}} \dots \partial x_{i_1}}(a+uh) (1-u)^k du \right) \right\| .$$

Il y a  $n^{k+1}$  terme dans la somme sous l'intégrale, et  $\int_0^1 (1-u)^k ds = \frac{1}{k+1}$ . On sait que la norme donnée sur  $\mathbb{R}^n$  est équivalente à la norme  $\|h\|_\infty = \sup \{|h_i| \mid i = 1, \dots, n\}$ , donc il existe  $K$  tel que  $\|h\|_\infty \leq K \cdot \|h\|$ . Or, si  $|\alpha| = k+1$  :

$$|h^\alpha| = |h_1^{\alpha_1} \dots h_n^{\alpha_n}| \leq \|h\|_\infty^{k+1} \leq K^{k+1} \cdot \|h\|^{k+1}$$

Si on pose :

$$C_k = \sup \left\{ \left\| \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a + uh) \right\| \mid |\alpha| = k+1, 0 \leq u \leq 1 \right\} \cdot \frac{n^{k+1}}{(k+1)!} K^{k+1}$$

l'inégalité (1-6) suit.

*q.e.d.*

**1.23 Remarques.**

(1) Considérons la fonction:

$$f(x) = \begin{cases} e^{-1/t^2} & \text{si } t > 0 \\ 0 & \text{si } t \leq 0 \end{cases} .$$

On montre que  $f$  est  $\mathcal{C}^\infty$  en 0, avec toutes ses dérivées nulles en ce point. Il en résulte que tous les termes de la formule de Taylor sont nuls, sauf le terme de reste. On en déduit que  $|f(t)| \leq |t|^k$ , ceci pour  $|t|$  assez petit et pour tout  $k$ .

(2) Un théorème dû à E. Borel affirme que pour toute suite  $\{a_k\}_{k \in \mathbb{N}} \subset \mathbb{R}$  de nombres réels, il existe une fonction  $f : [-1, 1] \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^\infty$  telle que  $f^{(k)}(0) = a_k$ . Cela exprime l'extrême flexibilité des fonctions différentiables.

On a un énoncé analogue pour les fonctions de plusieurs variables: si l'on se donne une famille  $\{a_\alpha\}_{\alpha \in \mathbb{N}^n} \subset \mathbb{R}$ , il existe  $f : U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$  ouvert,  $U \ni 0$  telle que  $\frac{\partial^{|\alpha|} f}{\partial x^\alpha}(0) = a_\alpha$ .

(3) Si  $f$  est  $\mathcal{C}^\infty$  et ses dérivées sont bornées sur une boule fermée de rayon fixe, on déduit de (1-5) que la série

$$f(a) + \sum_{k=1}^{\infty} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a) h^\alpha$$

converge uniformément sur cette même boule vers  $f(a+h)$ . Cela peut s'appliquer à des fonctions telles que  $\sin(x)$ ,  $\cos(x)$ ,  $e^{x+y}$ .

### Etude des extrema locaux de fonctions

**1.24 Définition.** Soit  $f : \Omega \rightarrow \mathbb{R}$ ,  $\Omega \subset \mathbb{R}^n$  ouvert,  $a \in \Omega$ . On dit que  $a$  est un minimum local strict (respectivement un maximum local strict) s'il existe un ouvert  $V \subset \Omega$ ,  $V \ni a$ , tel que :

$$x \in V, x \neq a \Rightarrow f(x) > f(a) \quad (\text{respectivement } f(x) < f(a)) \quad .$$

Si l'on remplace les inégalités strictes ci-dessus par des inégalités non strictes on obtient les notions de minimum local, resp. maximum local, non strict.

La formule de Taylor fournit un critère permettant parfois de décider si un point est un extremum local ou non d'une fonction.

**1.25 Proposition.** Soit  $f : \Omega \rightarrow \mathbb{R}$ ,  $\Omega \subset \mathbb{R}^n$  ouvert, de classe  $\mathcal{C}^{k+1}$ ,  $k \geq 2$ ,  $a \in U$ . Supposons que toutes les dérivées de  $f$  d'ordre au plus  $k-1$  s'annulent en  $a$  :

$$\frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a) = 0 \quad \forall \alpha \text{ tel que } 1 \leq |\alpha| \leq k-1 \quad .$$

Posons

$$\varphi_k(h) = \sum_{|\alpha|=k} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(a) h^\alpha \quad .$$

On a :

- (1) Si  $\varphi_k(h) > 0 \forall h \neq 0$ , alors  $a$  est un minimum local.
- (2) Si  $\varphi_k(h) < 0 \forall h \neq 0$ , alors  $a$  est un maximum local.
- (3) S'il existe  $h_1, h_2 \in \mathbb{R}^n$  tels que  $\varphi_k(h_1) > 0$ ,  $\varphi_k(h_2) < 0$  alors  $a$  n'est ni un minimum ni un maximum local, même non strict.

*Preuve:* Remarquons que  $\varphi_k(h)$  est homogène de degré  $k$ , ce qui veut dire que  $\forall \lambda \in \mathbb{R}$ ,  $\varphi_k(\lambda h) = \lambda^k \varphi_k(h)$ . Soit  $h_0 \in \mathbb{R}^n$  tel que  $\varphi_k(h_0) \neq 0$ , avec  $\|h_0\| = 1$ . Alors il suit de la formule de Taylor que

$$f(a + th_0) - f(a) = \varphi_k(th_0) + r_k(a, th_0) = t^k \left( \varphi_k(h_0) + \frac{r_k(a, th_0)}{t^k} \right) \quad .$$

Posons  $\varepsilon = |\varphi_k(h_0)|$ ; il suit de l'inégalité (1-6) de **1.22** qu'il existe  $\delta_\varepsilon$  tel que si  $|t| < \delta_\varepsilon$ ,

$$\frac{|r_k(a, th_0)|}{\underbrace{t^k \|h_0\|^k}_{=1}} < \varepsilon = |\varphi_k(h_0)|$$

et donc  $f(a + th_0) - f(a)$  a le même signe que  $\varphi_k(h_0)$ . Cela entraîne immédiatement l'affirmation (3). Pour (1) et (2), il suffit de poser

$$\varepsilon = \inf \{ |\varphi_k(h_0)| \mid \|h_0\| = 1 \} \quad ;$$

puisque  $|\varphi_k(h)|$  est continue, non nulle sur l'ensemble compact  $\{h \mid \|h\| = 1\}$ , il suit que  $\varepsilon > 0$ . La condition d'extremum local strict sera satisfaite avec  $V = \{a + th_0 \mid |t| < \delta_\varepsilon\}$ ,

*q.e.d.*

### 1.26 Remarques.

(1) Du fait que  $\varphi_k(-h) = (-1)^k \varphi_k(h)$  on déduit que :

- si la condition (1) ou la condition (2) de **1.25** sont satisfaites, alors  $k$  est pair.
- si  $a$  est un extremum local et  $k$  est l'ordre de la première dérivée non nulle en  $a$ , alors  $k$  doit être pair, sinon on pourrait appliquer (3) de **1.25** avec  $h_1$  un vecteur tel que  $\varphi_k(h_1) \neq 0$  et  $h_2 = -h_1$ .

(2) Il se peut qu'une fonction  $f$  a un minimum local en restriction à chaque droite passant par le point  $a$ , sans que  $a$  soit un minimum local pour  $f$  (même discours pour les maxima).

C'est le cas pour la fonction  $f(x, y) = (y - x^2)(y - 2x^2)$ . On voit sur la figure 9 que  $f$  admet un zéro isolé à l'origine de toute droite  $\ell$  passant par  $(0, 0)$ , et qu'elle est strictement positive sur un intervalle de  $\ell$  contenant  $(0, 0)$ . Cependant, on ne peut pas trouver une longueur d'intervalle uniforme, qui fonctionne sur toute droite. Et en fait il y a des points arbitrairement près de l'origine en lesquels  $f$  est positive, d'autres en lesquels  $f$  est négative.

(3) Si  $k = 2$  et  $\det\left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a)\right) \neq 0$ , alors on peut décider si  $a$  est un minimum local strict, un maximum, ou bien ni l'un ni l'autre en examinant la forme quadratique

$$q(h) = \phi_2(h) = \sum_{|\alpha|=2} \frac{1}{\alpha!} \frac{\partial^2 f}{\partial x^\alpha}(a) h^\alpha \quad .$$

En effet,  $q$  est soit définie positive, auquel cas  $a$  est un minimum local strict, soit définie négative, et alors  $a$  est un maximum local strict, soit elle est indéfinie, et alors  $a$  n'est ni minimum, ni maximum local, même non strict. On peut décider en mettant  $q$  sous forme diagonale.

Dans le cas de fonctions à 2 variables, si  $\det\left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a)\right) > 0$  alors  $q$  est définie, si  $\det\left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a)\right) < 0$  elle est indéfinie; si elle est définie, elle est définie positive si  $\frac{\partial^2 f}{\partial x_1^2}(a) > 0$  (ou  $\frac{\partial^2 f}{\partial x_2^2}(a) > 0$ ), elle est définie négative sinon.

(4) La fonction  $x^2 + y^4$  possède visiblement un minimum absolu en  $(0, 0)$ , mais les critères de **1.25** ne s'appliquent pas.

## 2. Le théorème des fonctions implicites

Soit  $f(x, y)$ ,  $(x, y) \in U \subset \mathbb{R}^2$ , une fonction de 2 variables. On peut lui associer une correspondance entre  $x$  et  $y$  de la manière suivante :

à tout  $x$  fixé on fait correspondre les valeurs de  $y$  qui sont solution de  $f(x, y) = 0$ .

On dira que cette correspondance entre  $x$  et  $y$  peut être rendue explicite s'il existe une fonction  $y(x)$  telle que :

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

En général, il n'est pas possible de passer de la relation implicite  $f(x, y) = 0$  à la relation explicite  $y = y(x)$ . Souvent c'est possible, mais seulement localement, au voisinage d'une solution particulière de  $f(x, y) = 0$ .

### 2.1 Exemples.

(1) Soit  $f(x, y) = ax + by + c = 0$  l'équation d'une droite. Si  $b \neq 0$ , on peut en tirer l'équation explicite :  $y = -\frac{a}{b}x - \frac{c}{b}$ , et si  $a \neq 0$ , on en tire que  $x = -\frac{b}{a} - \frac{c}{a}$ . Le premier cas revient à supposer que  $\frac{\partial f}{\partial y} \neq 0$ , ou encore que la droite n'est pas verticale; dans le deuxième,  $\frac{\partial f}{\partial x} \neq 0$ , ou encore que la droite n'est pas horizontale.

(2) De l'équation implicite  $x^2 + y^2 - 1 = 0$  on peut tirer 4 relations explicites :

$$\begin{aligned} y &= +\sqrt{1-x^2} \quad , |x| < 1, y > 0 \quad ; \quad y = -\sqrt{1-x^2} \quad , |x| < 1, y < 0 \\ x &= +\sqrt{1-y^2} \quad , |y| < 1, x > 0 \quad ; \quad x = -\sqrt{1-y^2} \quad , |y| < 1, x < 0 \end{aligned}$$

(3) De la relation implicite  $x^2 - y^2 = 0$  on peut tirer les 2 relations explicites :  $y = x$  et  $y = -x$ . Ensemble, elles décrivent toutes les solutions de  $x^2 - y^2 = 0$ . Mais au voisinage de la solution  $(0, 0)$ , aucune des 2 ne donne une solution explicite complète. En fait il n'y a pas de solution explicite au voisinage de  $(0, 0)$  (voir figure II.1).

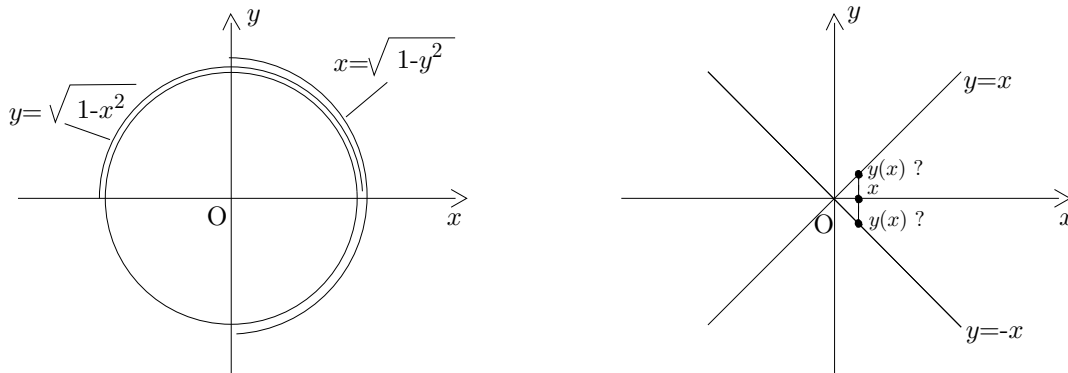


Figure II.1 – Passage d'une relation implicite à une relation explicite

Le théorème suivant nous donne une condition suffisante pour passer localement d'une relation implicite à une relation explicite. Considérons la décomposition en produit  $\mathbb{R}^n = \mathbb{R}^{n-p} \times \mathbb{R}^p$  et notons  $(x, y) \in \mathbb{R}^{n-p} \times \mathbb{R}^p$ ,  $x = (x_1, \dots, x_{n-p})$ ,  $y = (y_1, \dots, y_p)$ .

**2.2 Théorème des fonctions implicites.** Soit  $U \subset \mathbb{R}^n$  un ouvert et  $f : U \rightarrow \mathbb{R}^p$  continue, admettant une dérivée partielle  $\frac{\partial f}{\partial y}(x, y) \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^p)$  dépendant continûment de  $(x, y) \in U$ . Supposons que  $(x_0, y_0) \in U$  soit tel que :

i)  $f(x_0, y_0) = 0$

ii)  $\frac{\partial f}{\partial y}(x_0, y_0) : \mathbb{R}^p \rightarrow \mathbb{R}^p$  est bijective.

Alors il existe  $r_0, R_0 > 0$  et une application continue  $g : \overline{B(x_0, r_0)} \rightarrow \overline{B(y_0, R_0)}$  tels que:

- i)  $\overline{B(x_0, r_0)} \times \overline{B(y_0, R_0)} \subset U$   
 ii)  $f(x, g(x)) = 0 \quad \forall x \in \overline{B(x_0, r_0)}$  et pour tout  $(x, y) \in \overline{B(x_0, r_0)} \times \overline{B(y_0, R_0)}$  on a :

$$f(x, y) = 0 \iff y = g(x).$$

*Preuve:* Soient  $r, R > 0$  tels que  $\overline{B(x_0, r)} \times \overline{B(y_0, R)} \subset U$  et posons:

$$X_{r,R} = \left\{ \phi : \overline{B(x_0, r)} \rightarrow \overline{B(y_0, R)} \mid \phi \text{ continue} \right\} \quad ;$$

c'est un sous-ensemble fermé de l'espace vectoriel complet  $\left( \mathcal{C} \left( \overline{B(x_0, r)}, \mathbb{R}^p \right), \| \cdot \|_\infty \right)$  et donc  $X_{r,R}$  est complet pour la métrique induite par  $\| \cdot \|_\infty$  (voir **I.3.10**). Pour  $\phi \in X_{r,R}$  posons:

$$T(\phi)(x) = \phi(x) - \left( \frac{\partial f}{\partial y}(x_0, y_0) \right)^{-1} (f(x, \phi(x))) \quad ;$$

$T(\phi)$  est une application continue de  $\overline{B(x_0, r)}$  dans  $\mathbb{R}^p$ . Remarquons que

$$T(\phi) = \phi \iff f(x, \phi(x)) = 0 \quad \forall x \in \overline{B(x_0, r)}.$$

On a donc ramené la recherche d'une solution explicite à un problème de point fixe. En fait, la définition de  $T$  s'inspire de la 1ère variante de la méthode de Newton (voir **I.4.2**), que l'on retrouve lorsque  $p = 1$  et si l'on considère  $x$  comme un paramètre. Le reste de la preuve consiste pour l'essentiel à montrer que si  $r, R$  sont assez petits, alors  $T$  est une transformation contractante.

Définissons l'application auxiliaire

$$\Psi : \overline{B(x_0, r)} \times \overline{B(y_0, R)} \rightarrow \mathbb{R}^p \quad , \quad \Psi(x, y) = y - \left( \frac{\partial f}{\partial y}(x_0, y_0) \right)^{-1} (f(x, y))$$

de sorte que  $T(\phi)(x) = \Psi(x, \phi(x))$  et  $\Psi(x_0, y_0) = y_0$ . On a que

$$\frac{\partial \Psi}{\partial y}(x, y) = \mathbb{I}_p - \left( \frac{\partial f}{\partial y}(x_0, y_0) \right)^{-1} \left( \frac{\partial f}{\partial y}(x, y) \right)$$

où  $\mathbb{I}_p : \mathbb{R}^p \rightarrow \mathbb{R}^p$  dénote l'application identité.

Choisissons  $q$  tel que  $0 < q < 1$ . Il suit de l'hypothèse que  $\frac{\partial f}{\partial y}(x, y)$  est continue que  $\frac{\partial \Psi}{\partial y}(x, y)$  l'est aussi, et puisque  $\frac{\partial \Psi}{\partial y}(x_0, y_0) = \mathbb{I}_p - \left( \frac{\partial f}{\partial y}(x_0, y_0) \right)^{-1} \left( \frac{\partial f}{\partial y}(x_0, y_0) \right) = 0$ , il existe  $r_1 > 0$  et  $R_0 > 0$  tel que:

$$\|x - x_0\| \leq r_1, \|y - y_0\| \leq R_0 \Rightarrow \left\| \frac{\partial \Psi}{\partial y}(x, y) \right\| \leq q$$

ce qui entraîne encore, par le théorème des accroissements finis **1.9** que

$$\|x - x_0\| \leq r_1, \|y_1 - y_0\|, \|y_2 - y_0\| \leq R_0 \Rightarrow \|\Psi(x, y_1) - \Psi(x, y_2)\| \leq q \|y_1 - y_2\|.$$

Soit  $\delta > 0$  tel que

$$\|x - x_0\| \leq \delta \Rightarrow \|\Psi(x, y_0) - \Psi(x_0, y_0)\| \leq R_0(1 - q)$$

et posons  $r_0 = \inf \{r_1, \delta\}$ . Alors

$$\begin{aligned} \|x - x_0\| \leq r_0, \|y - y_0\| \leq R_0 &\Rightarrow \\ \|\Psi(x, y) - y_0\| &\leq \|\Psi(x, y) - \Psi(x, y_0)\| + \|\Psi(x, y_0) - \Psi(x_0, y_0)\| \\ &\leq q \|y - y_0\| + R_0(1 - q) \leq qR_0 + R_0(1 - q) = R_0. \end{aligned}$$



Ainsi, si  $\phi \in X_{r_0, R_0}$ ,

$$\|T(\phi) - y_0\|_\infty = \sup \{ \|\Psi(x, \phi(x))\|, \|x - x_0\| \leq r_0 \} \leq R_0$$

et si  $\phi_1, \phi_2 \in X_{r_0, R_0}$ ,

$$\begin{aligned} \|T(\phi_1) - T(\phi_2)\| &= \sup \{ \|\Psi(x, \phi_1(x)) - \Psi(x, \phi_2(x))\|, \|x - x_0\| \leq r_0 \} \\ &\leq q \cdot \sup \{ \|\phi_1(x) - \phi_2(x)\|, \|x - x_0\| \leq r_0 \} = q \|\phi_1 - \phi_2\|_\infty. \end{aligned}$$

Donc  $T$  est une transformation contractante de  $X_{r_0, R_0}$  et on peut lui appliquer le théorème du point fixe (chap. I, th. 5.1, page 16). L'unique point fixe de  $T$  sera une application continue  $g : \overline{B(x_0, r_0)} \rightarrow \overline{B(y_0, R_0)}$  telle que

$$f(x, g(x)) = 0.$$

L'unicité du point fixe de  $T$  nous dit que  $g$  est l'unique application avec la propriété ci-dessus, cependant cela n'empêche pas a priori que pour un  $x \in \overline{B(x_0, r_0)}$  il existe  $y \in \overline{B(y_0, R_0)}$ ,  $y \neq g(x)$  avec  $f(x, y) = 0$ . Mais si  $x \in \overline{B(x_0, r_0)}$  est fixé, on peut considérer la transformation

$$t_x : \overline{B(y_0, R_0)} \rightarrow \overline{B(y_0, R_0)}, t_x(y) = \Psi(x, y).$$

On déduit immédiatement des estimations qu'on a vues sur  $\Psi$  que  $t_x$  est une transformation contractante, donc elle admet un unique point fixe, qui ne peut être que  $g(x)$ .

*q.e.d.*

### 2.3 Remarques.

(1) Si l'on préfère travailler avec des boules ouvertes, on peut choisir  $r'_0 \leq r_0$  tel que  $g(B(x_0, r'_0)) \subset B(y_0, R_0)$ , et on aura encore:

$$\text{pour } (x, y) \in B(x_0, r'_0) \times B(y_0, R_0) \quad , \quad f(x, y) = 0 \iff y = g(x) \quad .$$

(2) Soit  $f : U \rightarrow \mathbb{R}^p$ ,  $U \subset \mathbb{R}^n$ ,  $x_0 \in U$ ,  $f(x_0) = 0$  et supposons que  $df_{x_0}$  soit surjective. Alors, quitte à renuméroter les vecteurs de la base naturelle de  $\mathbb{R}^n$ , on peut supposer que  $\det \left( \frac{\partial f_i}{\partial x_j}(x_0) \right)_{i=1, \dots, p, j=n-p+1, \dots, n} \neq 0$ . Si l'on prend la décomposition  $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{n-p}$  et l'on note  $(x', x'') \in \mathbb{R}^{n-p} \times \mathbb{R}^p$ , on aura que  $\frac{\partial f}{\partial x''}(x_0)$  est bijective et on pourra appliquer le théorème des fonctions explicites pour exprimer  $x''$  en fonction de  $x'$  près de  $x_0 = (x'_0, x''_0)$ .

(3) Désignons par  $Z(f) = \{(x, y) \in \mathbb{R}^n \mid (x, y) = 0\}$  l'ensemble des zéros de  $f$ . L'application  $h : B(x'_0, r) \rightarrow B(x''_0, R)$ ,  $x' \mapsto (x', g(x'))$ , a pour image  $Z(f) \cap (B(x'_0, r) \times B(x''_0, R))$ ; c'est une paramétrisation locale de  $X$ .

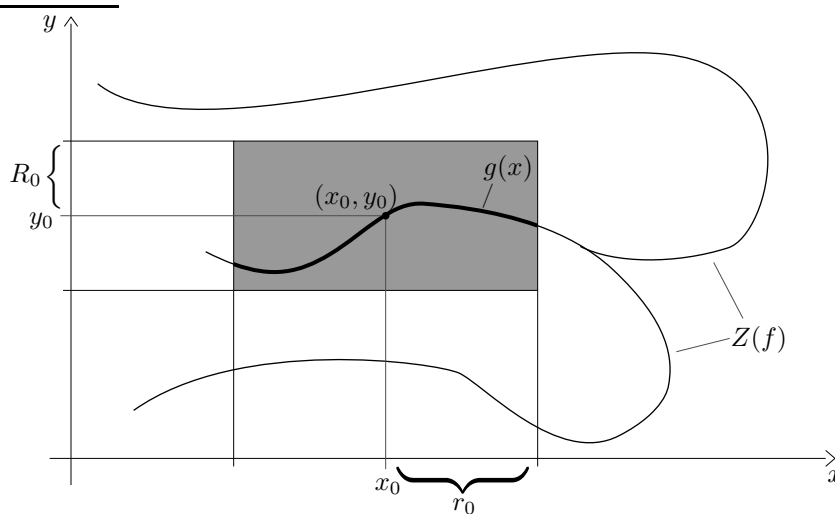


Figure II.2 – Le théorème des fonctions implicites

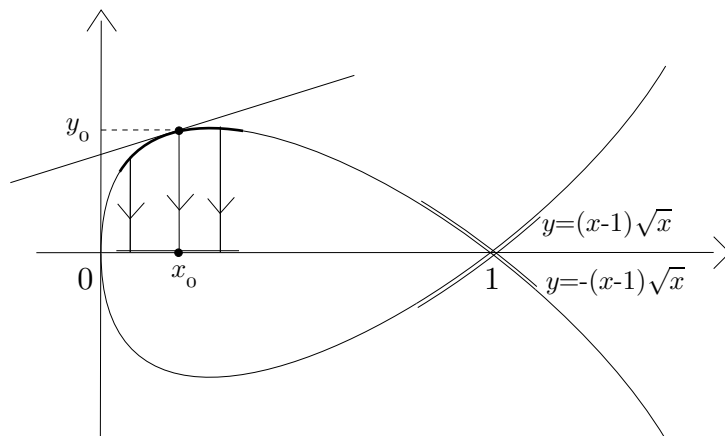


Figure II.3 – La courbe d'équation  $y^2 - x(x-1)^2 = 0$

## 2.4 Exemples.

(1) *Courbes planes.* Généralement, si  $x \in \mathbb{R}$  et  $y \in \mathbb{R}$ , l'équation  $f(x, y) = 0$  définit une courbe plane, notée  $Z(f)$ . Si  $(x_0, y_0) \in Z(f)$  et  $(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0)) \neq (0, 0)$ , on dit que  $(x_0, y_0)$  est un point régulier, sinon c'est un point singulier. On appelle droite tangente à  $Z(f)$  en un point régulier  $(x_0, y_0)$  la droite d'équation :

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

elle est proche de  $Z(f)$  près de  $(x_0, y_0)$ , parce que  $f(x, y) - f(x_0, y_0) = f(x, y)$  est approché par  $\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$ . Si  $(x_0, y_0)$  est un point régulier, le théorème des fonctions implicites nous dit qu'on peut paramétrer  $Z(f) \cap B(x_0) \times B(y_0, R_0)$  par une application de la forme  $x \mapsto (x, g(x))$ ,  $x \in B(x_0, r_0)$ , si  $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ , ou  $y \mapsto (y, g(y))$  si  $\frac{\partial f}{\partial x}(x_0, y_0) \neq 0$ . La condition  $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$  signifie que la droite tangente à  $Z(f)$  en  $(x_0, y_0)$  n'est pas verticale, et elle se projette donc bijectivement sur l'axe  $OX$ ; le théorème des fonctions implicites nous dit qu'alors la courbe elle-même, au voisinage de  $(x_0, y_0)$ , se projette bijectivement sur  $OX$  (voir figure II.3).

Par exemple, considérons la fonction  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f(x, y) = y^2 - x(x-y)^2$ . On a :

$$\frac{\partial f}{\partial y}(x, y) = 2y \quad , \quad \frac{\partial f}{\partial x}(x, y) = -3x^2 + 4x - 1$$

et

$$f(x, y) = 0, \quad \frac{\partial f}{\partial y}(x, y) = 0 \Rightarrow (x, y) = (0, 0) \text{ ou } (1, 0).$$

On pourra donc résoudre explicitement  $y$  en fonction de  $x$ , sauf aux 2 points  $(0, 0)$  et  $(1, 0)$ . Puisque  $\frac{\partial f}{\partial x}(0, 0) = -1$ , au voisinage de ce point on pourra résoudre explicitement  $x$  en fonction de  $y$ . Par contre  $(1, 0)$  est un point singulier, et on se rend compte sur la figure II.3 qu'il n'est pas possible de résoudre explicitement, ni  $y$  en fonction de  $x$ , ni  $x$  en fonction de  $y$ .

(2) *Surfaces de l'espace.* Une équation de la forme  $F(x, y, z) = 0$ ,  $x, y, z \in \mathbb{R}$ , définit généralement une surface de l'espace. Si  $P = (x_0, y_0, z_0) \in Z(F)$  et  $dF_P = (\frac{\partial F}{\partial x}(P), \frac{\partial F}{\partial y}(P), \frac{\partial F}{\partial z}(P)) \neq (0, 0, 0)$ , on dit que  $P$  est un point régulier, sinon c'est un point singulier. On appelle plan tangent à  $Z(F)$  en un point régulier  $P = (x_0, y_0, z_0)$  la plan d'équation :

$$\frac{\partial F}{\partial x}(P)(x - x_0) + \frac{\partial F}{\partial y}(P)(y - y_0) + \frac{\partial F}{\partial z}(P)(z - z_0) = 0 \quad ;$$

il est proche de  $Z(F)$  près de  $(x_0, y_0, z_0)$ , parce que  $F(x, y, z) - F(x_0, y_0, z_0) = F(x, y, z)$  est approché par  $\frac{\partial F}{\partial x}(x_0, y_0, z_0)(x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0, z_0)(y - y_0) + \frac{\partial F}{\partial z}(x_0, y_0, z_0)(z - z_0)$ .

Au voisinage d'un point régulier  $P$ , si par exemple  $\frac{\partial F}{\partial z}(P) \neq 0$ , on peut paramétrer  $Z(F)$  par  $(x, y, z) \mapsto (x, y, g(x, y))$ , pour  $(x, y)$  proche de  $(x_0, y_0)$ , où  $g(x, y)$  nous est fournie par le théorème des fonctions implicites.

Par exemple, considérons  $F(x, y, z) = x^2 + y^2 - z^2$ ,  $dF_{(x,y,z)} = (2x, 2y, -2z)$ . Donc  $(0, 0, 0) \in Z(F)$  est l'unique point singulier. Au voisinage de  $(0, 1, 1)$  on peut paramétrer  $Z(F)$  par  $(x, y) \mapsto (x, y, \sqrt{x^2 + y^2})$ ,  $(x, y)$  proche de  $(0, 1)$ .

**(3) Courbes dans l'espace.** Si  $\varphi = (\varphi^1, \varphi^2) : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ , généralement  $Z(\varphi)$  est une courbe dans l'espace. Un point  $P = (x_0, y_0, z_0) \in Z(\varphi)$  est dit régulier si  $d\varphi_P$  est de rang 2, sinon on dit qu'il est singulier. Si  $P$  est régulier, cela signifie que  $d\varphi_P^1$  et  $d\varphi_P^2$  sont linéairement indépendants; en particulier ils sont non nuls, et donc  $P$  est un point régulier sur chacune des deux surfaces  $Z(\varphi^1)$  et  $Z(\varphi^2)$ . Dire que  $d\varphi_P^1$  et  $d\varphi_P^2$  sont linéairement indépendants signifie exactement que ces plans tangents sont distincts; leur intersection est la droite tangente à  $Z(\varphi)$  en  $P$ . Au voisinage d'un point régulier, on peut paramétrer  $Z(\varphi)$  par projection sur  $OX$ ,  $OY$  ou  $OZ$ .

Par exemple prenons  $\varphi^1(x, y, z) = x^2 + y^2 - z^2$ ,  $\varphi^2(x, y, z) = y^2 + z^2 - 1$ ;  $Z(\varphi_1)$  est un cône circulaire d'axe  $OZ$ ,  $Z(\varphi_2)$  est un cylindre circulaire de rayon 1, d'axe  $OX$ . En posant  $\varphi = (\varphi^1, \varphi^2)$ , on a :

$$d\varphi = \begin{pmatrix} 2x & 2y & -2z \\ 0 & 2y & 2z \end{pmatrix} .$$

Le  $2 \times 2$  mineurs de cete matrice valent  $8yz$ ,  $4xy$  et  $4xz$ . Pour que le rang de  $d\varphi$  soit inférieur à 2 il faut donc que 2 des 3 coordonnées  $x, y, z$  s'annulent, et on vérifie que cela ne se produit en aucun point de  $Z(\varphi)$ ; donc tous les points de  $Z(\varphi)$  sont réguliers.

Au point  $(1, 0, 1)$ , le mineur de  $d\varphi$  correspondant à la première et dernière colonne est non nul; on peut paramétrer  $Z(\varphi)$  au voisinage de ce point par  $y \mapsto (\sqrt{1 - 2y^2}, \sqrt{1 - y^2})$ .

Dans les exemples précédents, on a utilisé souvent l'adverbe "généralement", parce qu'il est des équations particulières surprenantes; par exemple, si on prend  $F(x, y, z) = x^2 + y^2 + z^2 = 0$ , alors  $Z(F) = \{(0, 0, 0)\}$ .

## 2.1 DÉPENDANCE DES RACINES SIMPLES D'UNE FAMILLE DE POLYNÔMES PAR RAPPORT À DES PARAMÈTRES

D'abord un lemme qui nous donne une estimation des racines d'un polynôme à une variable en fonction de la taille de ses coefficients. On va supposer que le polynôme est distingué, ce qui veut dire que le coefficient du terme du plus haut degré vaut 1.

**2.5 Lemme.** Soit  $f(x) = x^d + \sum_{i=1}^d a_i x^{d-i}$ ,  $a_i \in \mathbb{R}$  et supposons que  $|a_i| \leq M$ ,  $i = 1, \dots, d$ . Alors, si  $\alpha \in \mathbb{R}$  et  $f(\alpha) = 0$  on a :

$$|\alpha| < 1 + M$$

*Preuve:* Si  $|\alpha| \leq 1$  il n'y a rien à démontrer. Sinon :

$$\sum_{i=1}^d \frac{1}{|\alpha|^i} < \sum_{i=1}^{\infty} \frac{1}{|\alpha|^i} = \frac{1}{|\alpha|} \frac{1}{1 - 1/|\alpha|}$$

et alors

$$\begin{aligned} 0 = |\alpha^d| \left| 1 + \sum_{i=0}^{d-1} a_i \alpha^i \right| &\geq |\alpha|^d \left( 1 - M \sum_{i=1}^d \frac{1}{|\alpha|^i} \right) > \\ &|\alpha|^d \left( 1 - \frac{M}{|\alpha|} \left( \frac{1}{1 - 1/|\alpha|} \right) \right) = |\alpha|^d \left( 1 - M \frac{1}{|\alpha| - 1} \right) \\ &\implies 1 - M \frac{1}{|\alpha| - 1} < 0 \implies |\alpha| < 1 + M \end{aligned}$$

et donc  $|\alpha| < 1 + M$  dans tous les cas.

*q.e.d.*

**2.6 Théorème.** Soit  $F(x, \lambda) = x^d + \sum_{i=0}^{d-1} a_i(\lambda) x^i$  une famille de polynômes de degré  $d$ , où les coefficients sont des fonctions continues  $a_i(\lambda) : U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$ . Supposons que pour un  $\lambda_0 \in U$  l'équation

$$F(x, \lambda_0) = 0$$

possède exactement  $\nu$  racines, toutes simples,  $\alpha_1, \dots, \alpha_\nu \in \mathbb{R}$  :

$$F(x, \lambda_0) = 0 \implies \exists i \text{ t.q. } x = \alpha_i \quad , \quad \text{et} \quad \frac{\partial F}{\partial x}(\alpha_i, \lambda_0) \neq 0 \quad , \quad i = 1, \dots, \nu$$

Alors il existe  $r > 0$  et  $\nu$  fonctions continues  $\gamma_i : B(\lambda_0, r) \rightarrow \mathbb{R}$  telles que :

- i)  $\gamma_i(\lambda_0) = \alpha_i$
- ii)  $F(x, \lambda) = 0, \lambda \in B(\lambda_0, r) \iff \exists i \text{ t.q. } x = \gamma_i(\lambda)$

*Preuve:* D'après le théorème des fonctions implicites 2.2 il existe  $r > 0$ ,  $R_i > 0$ ,  $i = 1, \dots, \nu$  et des fonctions continues  $\gamma_i : B(\lambda_0, r) \rightarrow ]\alpha_i - R_i, \alpha_i + R_i[$ ,  $i = 1, \dots, \nu$ , telles que

$$F(x, \lambda) = 0, x \in B(\alpha_i, R_i), \lambda \in B(\lambda_0, r) \iff \exists i \text{ t.q. } x = \gamma_i(\lambda) \quad .$$

Il reste à voir que si  $r$  est choisi suffisamment petit il n'y a pas d'autres racines.

Soit  $r' < r$ ; alors, puisque  $\{\lambda \mid \|\lambda - \lambda_0\| \leq r'\}$  est compact et les  $a(\lambda)$  continues, il existe  $M$  tel que  $|a(\lambda)| \leq M$  pour  $\|\lambda - \lambda_0\| \leq r'$ . Alors il suit du lemme qui précède que  $F(x, \lambda) \neq 0$  si  $\lambda \in B(\lambda_0, r')$  et  $|x| \geq 1 + M$ . Posons

$$\mu = \inf \left\{ |F(x, \lambda_0)| \mid x \in [-1 - M, 1 + M] \setminus \cup_{i=1, \dots, \nu} B(\alpha_i, R_i) \right\} \quad ;$$

alors  $\mu > 0$ . Il existe  $r'' \leq r'$  tel que  $|F(x, \lambda)| \geq \mu/2$  si  $\|\lambda - \lambda_0\| \leq r''$ ,  $x \in [-1 - M, 1 + M] \setminus \cup_{i=1, \dots, \nu} B(\alpha_i, R_i)$ ; les racines de  $F(x, \lambda)$  ne peuvent donc être que dans  $\cup_{i=1, \dots, \nu} B(\alpha_i, R_i)$ .

*q.e.d.*

Remarquons que si l'on admet des racines non simples, le théorème précédent est faux :  $x^2 + \lambda$  admet une racine pour  $\lambda = 0$ , qui disparaît dès que  $\lambda > 0$ . Aussi, le théorème est en défaut pour la famille de polynômes  $\lambda x^2 - x$  : pour  $\lambda = 0$ ,  $x = 0$  est l'unique racine, et elle est simple, mais dès que  $\lambda \neq 0$  il y a 2 racines simples :  $x = 0$  et  $x = \frac{1}{\lambda}$ .

On peut néanmoins étendre un peu la validité de **2.6** :

**2.7 Corollaire.** Soit  $F(x, \lambda) = \sum_{i=0}^d a_i(\lambda)x^i$  une famille de polynômes de degré  $d$ , où les coefficients sont des fonctions continues  $a_i(\lambda) : U \rightarrow \mathbb{R}$ ,  $U \subset \mathbb{R}^n$ . Soit  $\lambda_0 \in U$  tel que l'on ait :

- i)  $a_d(\lambda_0) \neq 0$
- ii) l'équation

$$F(x, \lambda_0) = 0$$

possède exactement  $\nu$  racines, toutes simples.

Alors la conclusion du théorème **2.6** reste valable.

*Preuve:* En posant  $F'(x, \lambda) = 1/a_d(\lambda)$ , pour  $\lambda$  dans un voisinage de  $\lambda_0$ , on est ramené à **2.6**.

*q.e.d.*

On en déduit une méthode pour dessiner des courbes planes décrites par une équation polynomiale en deux variable  $x$  et  $y$ , que nous allons esquisser.

Soit  $f(x, y) = \sum_{i,j=0}^d a_{i,j}x^i y^j$  un polynôme en  $x$  et  $y$  et soit

$$X = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = 0\}$$

la courbe plane qu'il définit. Posons  $f_x(y) = f(x, y) = a_0(x) + a_1(x)y + \dots + a_d(x)y^d$ , que l'on considère comme famille de polynômes en  $y$  dépendants du paramètre  $x$ . L'idée de cette méthode est de fixer  $x \in \mathbb{R}$ , puis de chercher l'ensemble des  $y$  solutions de l'équation  $f_x(y) = 0$ ; ce sera un ensemble  $S_x \subset \mathbb{R}$  qui sera soit fini, soit égal à  $\mathbb{R}$ , si la courbe  $X$  contient la droite verticale  $\{x\} \times \mathbb{R}$ . Puis on essaie de comprendre comment  $S_x$  se comporte lorsque  $x$  varie. Des problèmes vont surgir lorsque les racines de  $f_x$  ne dépendent pas continûment de  $x$ . D'après **2.7** cela ne peut arriver que dans les cas suivants :

- lorsque  $x$  annule le coefficient de degré maximal en  $y$  :  $a_d(x) = 0$
- lorsque  $\frac{\partial f}{\partial y}(x, y) = 0$  et  $f(x, y) = 0$

En ces points, qui sont généralement en nombre fini, il faut examiner de plus près l'équation. On va noter  $\Sigma \subset \mathbb{R}$  l'ensemble de ces points. Son complémentaire  $\mathbb{R} \setminus \Sigma$  est constitué d'un nombre fini d'intervalles  $I_1, \dots, I_N$ . Il suit de **2.7** que  $X \cap I_h$  se compose d'une réunion finie d'arcs disjoints, chacun se projetant bijectivement sur  $I_h$ , pour  $h = 1, \dots, N$ .

### 2.8 Exemple.

Considérons la courbe d'équation :

$$y^2(1-x) - x(2x-1)^2 = 0 \quad .$$

On calcule que

$$\frac{\partial f}{\partial x} = -y^2 - 6x^2 + 8x - 1 \quad , \quad \frac{\partial f}{\partial y} = 2y - 2xy \quad .$$

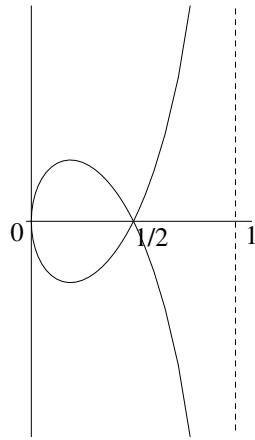
Les solutions de  $f(x, y) = 0$ ,  $\frac{\partial f}{\partial y} = 0$  sont  $(0, 0)$  et  $(0, 1/2)$ . Le coefficient du terme de degré maximum en  $y$  est  $1-x$ , nul en  $x = 1$ . Donc

$$\Sigma = \{0, 1/2, 1\} \quad .$$

On vérifie que l'équation  $f_x(y) = 0$  possède 2 solutions symétrique par rapport à l'axe  $Ox$  lorsque  $0 < x < 1/2$  et  $1/2 < x < 1$ . Le point  $(0, 0)$  n'est pas singulier, alors que  $(1/2, 0)$  l'est. Pour comprendre ce qui se passe en  $(1/2, 0)$  on regarde le développement de Taylor de  $f$  en ce point :

$$f(x, y) = 1/2y^2 - 2(x - 1/2) + r_2$$

et il est donc probable que  $X$  ressemble à la paire de droites  $1/2y^2 - 2(x - 1/2)^2 = 0$  au voisinage de ce point. Enfin, si  $x \rightarrow 1$ , les deux solutions de  $f_x(y) = 0$  ne peuvent que tendre vers l'infini, car  $f(1, y) = -1 \neq 0$  (figure **II.4**).



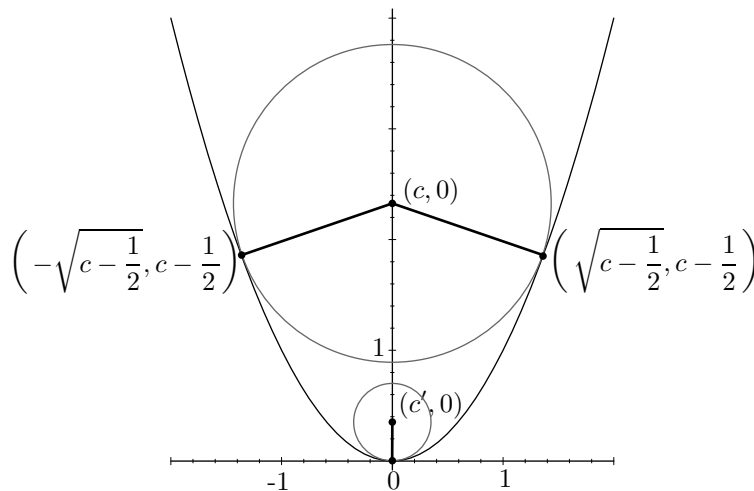
**Figure II.4** – La courbe d'équation  $y^2(1-x) - x(2x-1)^2 = 0$ .

## 2.2 MÉTHODE DES MULTIPLICATEURS DE LAGRANGE POUR LA RECHERCHE D'EXTREMA LIÉS

Soit  $\Omega \subset \mathbb{R}^n$  un ouvert et  $F : \Omega \rightarrow \mathbb{R}^k$  une application. Posons  $X = \{x \in \mathbb{R}^n \mid F(x) = 0\} = Z(F)$ ; si  $f : X \rightarrow \mathbb{R}$  est une fonction, on aimerait chercher les extrema locaux de la restriction de  $f$  à  $X$  : on dit que  $a \in X$  est un maximum (resp. minimum) local de  $f|_X$  si  $\exists r > 0$  tel que  $x \in X$ ,  $d(x, a) < r \Rightarrow f(x) \leq f(a)$  (resp.  $f(x) \geq f(a)$ ). On dit alors que  $a$  est un extremum local de  $f$ , lié aux conditions  $F_i(x) = 0$ ,  $i = 1, \dots, k$ .

Par exemple, cherchons les points de la parabole  $y = x^2$  qui sont à distance minimale du point  $P = (0, c)$ . Dans ce cas,  $F(x, y) = y - x^2$ , et on peut prendre pour  $f$  la fonction  $f(x, y) = x^2 + (y - c)^2$ , c'est-à-dire le carré de la distance de  $(x, y)$  à  $P$ ; minimiser  $f(x, y)$  ou la distance proprement dite  $\sqrt{f(x, y)}$  revient au même, mais  $f(x, y)$  est plus simple à écrire. On peut résoudre directement ce problème à l'aide de la paramétrisation globale de  $Z(F)$  :  $\alpha(x) = (x, x^2)$ ; alors  $f(\alpha(x)) = x^2 + (x^2 - c)^2 = \phi(x)$ ,  $\phi'(x) = 2x + 2(x^2 - c)2x = 2x(2x^2 - 2c + 1)$ .

On doit chercher la valeur de  $x$  pour laquelle  $\phi(x)$  est minimale, et alors  $\phi'(x) = 0$ , d'où  $x = 0$  ou  $x = \sqrt{c - 1/2}$  si  $c > 1/2$ . Or  $\phi(0) = c^2$ ,  $\phi(\sqrt{c - 1/2}) = c - 1/4$  ( $c \geq 1/2$ ). On vérifie que  $c^2 \geq c - 1/4$  si  $c \geq 1/2$ ; donc que la distance minimale est atteinte par le point  $(0, 0)$  si  $c < 1/2$ , sinon par les deux points  $(\pm\sqrt{c - 1/2}, c - 1/2)$  (figure II.5).



**Figure II.5** – Exemple d'extremum lié : distance minimale d'une courbe à un point donné

Dans cet exemple, on s'est ramené, à l'aide d'une paramétrisation de la courbe  $g(x, y) = 0$ , à la recherche de points critiques d'une fonction d'une variable. En général, il n'est pas possible d'explicitement une paramétrisation d'une courbe décrite par une équation : le théorème des fonctions implicites est un pur théorème d'existence. Néanmoins, un complément à ce théorème va nous donner l'expression de la première dérivée d'une paramétrisation locale de la courbe  $g = 0$ ; c'est ce que va exploiter la méthode des multiplicateurs de Lagrange, qui permet sous certaines conditions de donner des équations des extremas d'une fonction liés à une contrainte.

**2.9 Complément au théorème des fonctions implicites.** *Sous les hypothèses du théorème 2.2 (théorème des fonctions implicites), si  $f$  est dérivable en  $(x_0, y_0)$ , alors  $g$  est dérivable en  $x_0$  et*

$$(3) \quad dg_{x_0} = - \left( \frac{\partial f}{\partial y}(x_0, y_0) \right)^{-1} \left( \frac{\partial f}{\partial x}(x_0, y_0) \right).$$

De plus, si  $f$  est de classe  $C^r$ ,  $g$  l'est aussi.

*Preuve:* En dérivant les 2 membres de l'équation:

$$f(x, g(x)) = 0$$

on obtient que

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

et de là on tire que

$$dg_{x_0} = - \left( \frac{\partial f}{\partial y}(x_0, y_0) \right)^{-1} \left( \frac{\partial f}{\partial x}(x_0, y_0) \right)$$

ceci à condition que l'on sache déjà que  $g$  est dérivable en  $x_0$ . C'est précisément ce que l'on doit commencer par démontrer, mais le calcul précédent explique d'où vient la formule énoncée.

Pour montrer que  $g$  est dérivable en  $x_0$ , donnons un accroissement  $h$  à la variable  $x$  en  $x_0$ , un accroissement  $\Delta g = g(x_0 + h) - g(x_0)$  à  $y$  en  $y_0$  et utilisons la dérivabilité de  $f$  en  $(x_0, y_0)$  et le fait que  $f(x, g(x)) = 0$ :

$$(4) \quad 0 = f(x_0 + h, g(x_0 + h)) - f(x_0, g(x_0)) = \frac{\partial f}{\partial x}(x_0, y_0)(h) + \frac{\partial f}{\partial y}(x_0, y_0)(\Delta g) + r(h, \Delta g)$$

avec  $r(h, \Delta g) \leq \varepsilon (\|h\| + \|\Delta g\|)$  si  $\|h\|, \|\Delta g\| \leq \delta_\varepsilon$ , où l'on prend partout la norme  $\| \cdot \|_1$ . Puisque  $g$  est continue

$$\|h\| \leq \delta_\varepsilon^1 \Rightarrow \|\Delta g\| \leq \delta_\varepsilon$$

donc si  $\delta_\varepsilon^2 = \inf \{ \delta_\varepsilon, \delta_\varepsilon^1 \}$ ,

$$(5) \quad \|h\| \leq \delta_\varepsilon^2 \Rightarrow r(h, \Delta g) \leq \varepsilon (\|h\| + \|\Delta g\|)$$

et on déduit de (4) que

$$(6) \quad \Delta g = - \frac{\partial f}{\partial y}^{-1} \left( \frac{\partial f}{\partial x}(h) + r(h, \Delta g) \right)$$

et donc, en supposant que  $\varepsilon < 1$

$$\|h\| \leq \delta_\varepsilon^2 \Rightarrow \|\Delta g\| \leq \left\| \frac{\partial f}{\partial y}^{-1} \right\| \left( \left( \left\| \frac{\partial f}{\partial x} \right\| + 1 \right) \|h\| + \varepsilon \|\Delta g\| \right)$$

où cela va sans dire que les dérivées sont prises au point  $(x_0, y_0)$ .

Posons  $C = \left\| \frac{\partial f^{-1}}{\partial y} \right\| \left( \left\| \frac{\partial f}{\partial x} \right\| + 1 \right)$ ; en prenant  $\varepsilon = \frac{1}{2 \left\| \frac{\partial f^{-1}}{\partial y} \right\|}$  et en appelant  $\delta^3$  la valeur correspondante de  $\delta_\varepsilon^2$ , on obtient que

$$\|h\| \leq \delta^3 \Rightarrow \|\Delta g\| \leq C \|h\| + \frac{1}{2 \left\| \frac{\partial f^{-1}}{\partial y} \right\|} \left\| \frac{\partial f^{-1}}{\partial y} \right\| \|\Delta g\| = C \|h\| + \frac{1}{2} \|\Delta g\|$$

$$(7) \Rightarrow \frac{1}{2} \|\Delta g\| \leq C \|h\| \Rightarrow \|\Delta g\| \leq 2C \|h\|$$

Si on reprend (6) on voit que

$$(8) \quad \Delta g = -\frac{\partial f^{-1}}{\partial y} \left( \frac{\partial f}{\partial x}(h) \right) + \rho(h) \text{ où } \rho(h) = -\frac{\partial f^{-1}}{\partial y} (r(h, \Delta g))$$

et il suit de (5), (7) et (8) que

$$\|h\| \leq \inf \{ \delta_\varepsilon, \delta^3 \} \Rightarrow \|\rho(h)\| \leq \varepsilon \left\| \frac{\partial f^{-1}}{\partial y} \right\| (1 + 2C) \|h\| \quad ;$$

cela montre que  $g$  est dérivable et que sa dérivée est donnée par (3).

Si l'on suppose  $f$  de classe  $\mathcal{C}^r$ ,  $r \geq 1$ ,  $\frac{\partial f}{\partial x}$  et  $\frac{\partial f}{\partial y}$  sont de classe  $\mathcal{C}^{r-1}$ . L'expression (3) est valable sur un ouvert  $U' \ni x_0$  et montre que  $dg$  est de classe  $\mathcal{C}^{r-1}$  et donc que  $g$  est de classe  $\mathcal{C}^r$ .

*q. e. d.*

**2.10 Théorème – Methode des multiplicateurs de Lagrange.** Soient  $F : \Omega \rightarrow \mathbb{R}^k$ , et  $f : \Omega \rightarrow \mathbb{R}$ ,  $\Omega \subset \mathbb{R}^n$  un ouvert,  $F, f$  de classe  $\mathcal{C}^1$ . Posons  $X = F^{-1}(0)$  et soit  $a \in X$ ; supposons que  $dF_a$  soit surjective. Alors, une condition nécessaire pour que  $f|_X$  ait un extremum local en  $a$  est qu'il existe  $k$  nombres réels  $\lambda_1, \dots, \lambda_k \in \mathbb{R}$  tel que :

$$df_a = \sum_{i=1}^k \lambda_i \cdot F'_i(a)$$

où  $F'_i(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$  dénote la dérivée de  $F_i$  au point  $a$ . En coordonnées, cela s'écrit :

$$\frac{\partial f}{\partial x_j}(a) = \sum_{i=1}^k \lambda_i \cdot \frac{\partial F_i}{\partial x_j}(a) \quad , \quad j = 1, \dots, n$$

Notons que pour la recherche des extremas liés nous aurons :

$$n + k \text{ inconnues} : a_1, \dots, a_n, \lambda_1, \dots, \lambda_k$$

$$n + k \text{ équations} : F_1(a) = \dots = F_k(a) = 0, \quad \frac{\partial f}{\partial x_j}(a) = \sum_{i=1}^k \lambda_i \cdot \frac{\partial F_i}{\partial x_j}(a), \quad j = 1, \dots, n$$

et et il est donc raisonnable d'espérer de résoudre ce système d'équations.

Les points  $a \in \Omega$  qui satisfont les  $n + k$  équations ci-dessus, sans préjuger de leur qualité de maximum ou minimum local lié, sont appelés **points stationnaires**.

Lorsque  $n = 2$  et  $k = 1$ , les niveaux de  $f$  et  $F$  sont généralement des courbes. La condition du théorème devient  $\text{grad}(f)_a = \lambda \cdot \text{grad}(F)_a$ , donc les 2 courbes  $F(x_1, x_2) = 0$  et  $f(x_1, x_2) - f(a) = 0$  ont même tangente en  $a$ ; on le constate par exemple sur les figures **II.5** et **II.6** (voir aussi remarque **2.11**).

*Preuve du théorème 2.10 :* Quitte à renuméroter les vecteurs de base, on peut supposer que

$$\det \left( \frac{\partial F_i}{\partial x_j}(a) \right)_{i=1, \dots, k, j=n-k+1, \dots, n} \neq 0 \quad ;$$



on prend la décomposition  $\mathbb{R}^n = \mathbb{R}^{n-k} \times \mathbb{R}^k$  et on note  $x = (x', x'') \in \mathbb{R}^{n-k} \times \mathbb{R}^k$ . Il suit du théorème des fonctions implicites **2.2** qu'il existe  $R, r > 0$  et une application  $g : B(a', r) \rightarrow B(a'', R)$  de classe  $\mathcal{C}^1$  telle que  $X \cap (B(a', r) \times B(a'', R)) = \{x', g(x') \mid x' \in B(a', r)\}$ . On a donc que la fonction  $x' \mapsto f(x', g(x'))$  a un extremum local en  $a'$ , et alors d'après la proposition **1.12** la dérivée de  $x' \mapsto f(x', g(x'))$  est nulle en  $a'$ , c'est-à-dire :

$$\frac{\partial f}{\partial x'}(a) + \frac{\partial f}{\partial x''}(a) dg_{a'} = 0 \quad .$$

D'autre part, par **2.9** :

$$dg_{a'} = - \left( \frac{\partial F}{\partial x''}(a) \right)^{-1} \frac{\partial F}{\partial x'}(a)$$

et donc

$$\heartsuit \quad \frac{\partial f}{\partial x'}(a) = \frac{\partial f}{\partial x''}(a) \left( \frac{\partial F}{\partial x''}(a) \right)^{-1} \frac{\partial F}{\partial x'}(a) \quad .$$

Posons

$$\Lambda = \frac{\partial f}{\partial x''}(a) \left( \frac{\partial F}{\partial x''}(a) \right)^{-1} \quad ,$$

alors  $\heartsuit$  s'écrit :

$$\frac{\partial f}{\partial x'}(a) = \Lambda \circ \left( \frac{\partial F}{\partial x'}(a) \right)$$

et on a aussi trivialement :

$$\frac{\partial f}{\partial x''}(a) = \frac{\partial f}{\partial x''}(a) \left( \frac{\partial F}{\partial x''}(a) \right)^{-1} \frac{\partial F}{\partial x''}(a) = \Lambda \circ \left( \frac{\partial F}{\partial x''}(a) \right)$$

donc finalement :

$$df_a = \Lambda \circ dF_a \quad .$$

Si  $(\lambda_1, \dots, \lambda_k)$  est la matrice de  $\Lambda$ , le dernière égalité s'écrit :

$$df_a = \sum_{i_1}^k \lambda_{i_1} \cdot F'_{i_1}(a) \quad .$$

*q.e.d.*

**2.11 Remarque.** Quelques définitions vont nous permettre de mieux comprendre le théorème précédent. Si  $x_0 \in Z(f)$ , où  $F : \Omega \rightarrow \mathbb{R}^k$ ,  $\Omega \subset \mathbb{R}^n$ , on dit que  $a$  est un point régulier de  $Z(F)$  si  $dF_a$  est surjective, et alors, quitte à renuméroter les coordonnées de  $\mathbb{R}^n$  comme dans la preuve de **2.10**, on peut supposer qu'on a une paramétrisation locale de  $Z(F)$  de la forme  $x' \mapsto (x', g(x'))$ . On appelle espace tangent à  $Z(F)$  au point régulier  $a$ , noté  $T(Z(F))_a$ , le noyau de la dérivée de  $F$  en  $a$  :

$$T(Z(F))_a = \{v \in \mathbb{R}^n \mid dF_a(v) = 0\} \quad .$$

Notons que la droite tangente ou le plan tangent définis dans **2.4** sont en fait les translatés au point  $a$  de l'espace tangent tel qu'on vient de le définir.

La dérivée de la paramétrisation locale  $x' \mapsto (x', g(x'))$  de  $Z(F)$  au voisinage de  $a$  s'écrit  $v' \mapsto (v', dg_a(v'))$ , et on voit qu'elle est injective. Puisque  $F(x, g(x)) = 0 \forall x' \in B(a', r) \Rightarrow dF_a(v', dg_a(v')) = 0 \forall v' \in \mathbb{R}^{n-k}$ , on a que :

$$\{(v', dg_a(v')), v' \in \mathbb{R}^{n-k}\} \subset \ker(dF_a)$$

et comme ces deux espaces ont même dimension  $n - k$ , ils coïncident :

$$T(Z(F))_a = \{(v', dg_a(v')), v' \in \mathbb{R}^{n-k}\} = \ker(dF_a) \quad .$$

Dans la preuve de **2.10**, on a utilisé que  $\frac{\partial f}{\partial x'}(a) + \frac{\partial f}{\partial x''}(a) dg_{a'} = 0$ , ce qui signifie que :

$$\frac{\partial f}{\partial x'}(a)(v') + \frac{\partial f}{\partial x''}(a)(dg_{a'}(v')) = 0 \quad \forall v' \in \mathbb{R}^{n-k}$$

ou encore :

**l'application linéaire  $df_a : \mathbb{R}^n \rightarrow \mathbb{R}$  doit s'annuler sur le sous-espace  $T(Z(F))_a$**

De là on a déduit ensuite que  $df_a = \sum_{i=1}^k \lambda_i \cdot F'_i(a)$ . Cela résulte aussi de l'affirmation suivante, qui est laissée au lecteur : Soit  $A : \mathbb{R}^n \rightarrow \mathbb{R}^k$  une application linéaire surjective et  $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$  une application linéaire. Alors :

$$\exists \Lambda : \mathbb{R}^k \rightarrow \mathbb{R} \text{ linéaire t.q. } \alpha = \Lambda \circ A \Leftrightarrow \ker(\alpha) \supset \ker(A) \quad .$$

## 2.12 Exemples.

(1) Soit  $f(x, y) = xy$  et cherchons ses valeurs maximales et minimales sur le cercle  $F(x, y) = x^2 + y^2 - 1 = 0$ ; on sait qu'elles existent parce que le cercle est compact. Cherchons d'abord ses points stationnaires :

$$\left. \begin{array}{l} x^2 + y^2 - 1 = 0 \\ y = 2\lambda x \\ x = 2\lambda y \end{array} \right\} \Rightarrow x^2 = y^2 = 1/2 \Rightarrow a = (\pm\sqrt{1/2}, \pm\sqrt{1/2})$$

Il y a donc quatre points stationnaires.  $f(\pm(\sqrt{1/2}, \sqrt{1/2})) = 1/2$ ,  $f(\pm(\sqrt{1/2}, -\sqrt{1/2})) = -1/2$ . Donc la valeur maximale est  $1/2$ , la valeur minimale est  $-1/2$  (figure **II.6**).

(2) Cherchons les valeurs maximales et minimales de la fonction  $f(x, y, z) = x + y + z$  sur l'ellipsoïde  $F(x, y, z) = \frac{x^2}{2} + \frac{y^2}{4} + \frac{z^2}{6} - 1 = 0$ . De nouveau, on sait a priori que  $f$  admet une valeur minimale et maximale, parce que l'ellipsoïde est compact. On a  $\text{grad}(f) = (1, 1, 1)$ ,  $\text{grad}(F) = (x, \frac{y}{2}, \frac{z}{3})$  et on doit résoudre le système d'équations :

$$\left. \begin{array}{l} 1 = \lambda x \\ 1 = \frac{1}{2}\lambda y \\ 1 = \frac{1}{3}\lambda z \end{array} \right\} \Rightarrow \lambda = \frac{1}{x}, y = 2x, z = 3x \Rightarrow \frac{x^2}{2} + \frac{4x^2}{4} + \frac{9x^2}{6} = 1 \Rightarrow x = \pm\frac{\sqrt{3}}{3} \quad .$$

Les points stationnaires sont donc  $P = \sqrt{3}(\frac{1}{3}, \frac{2}{3}, 1)$  et  $Q = -\sqrt{3}(\frac{1}{3}, \frac{2}{3}, 1)$ . Donc  $f(P) = 2\sqrt{3}$  est la valeur maximale,  $f(Q) = -2\sqrt{3}$  est la valeur minimale de  $f$  sur l'ellipsoïde.

(3) Si au lieu de l'ellipsoïde on prend la quadrique  $h(x, y, z) = -\frac{x^2}{2} + \frac{y^2}{4} + \frac{z^2}{6} - 1 = 0$ , des calculs semblables à ceux ci-dessus montrent qu'il y a 2 points stationnaires  $P' = \sqrt{2}(\frac{1}{2}, -1, -\frac{3}{2})$  et  $Q' = \sqrt{2}(\frac{1}{2}, -1, -\frac{3}{2})$ , et  $f(P') = -2\sqrt{2}$ ,  $f(Q') = 2\sqrt{2}$ . Mais  $f$  n'a ni valeur maximale, ni minimale sur cette surface! En effet, prenons  $z_0 = \sqrt{6}$ ; alors, les points  $(x, \sqrt{2}x, z_0) \in Z(h)$ , et  $f(x, \sqrt{2}x, z_0) = x(1 + \sqrt{2}) + \sqrt{6}$  prend des valeurs aussi grandes que l'on veut.

(4) La méthode ne permet pas toujours de trouver des solutions, même si elles existent. Prenons la parabole semi-cubique (figure **II.7**), d'équation  $g(x, y) = y^2 - x^3 = 0$  et cherchons la distance minimale du point  $P = (-1, 0)$  à cette courbe, en appliquant la méthode de Lagrange à  $f(x, y) = (x + 1)^2 + y^2$ , qui est la distance au carré de  $(x, y)$  à  $(-1, 0)$ , soumise à la contrainte  $F(x, y) = 0$ . On voit tout de suite que la distance minimale est atteinte au point  $(0, 0)$ , et elle vaut 1. Mais  $\text{grad}(F)_{(0,0)} = (0, 0)$ ,  $\text{grad}(f)_{(0,0)} = (1, 0)$ ; on ne pourra pas trouver de  $\lambda$  tel que  $\text{grad}(f)_{(0,0)} = \lambda \cdot \text{grad}(F)_{(0,0)}$ .

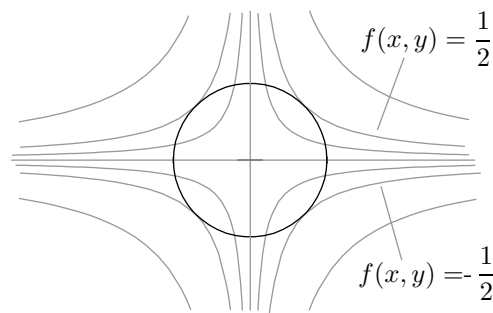


Figure II.6 – Valeurs extrémales de  $x \cdot y$  sur le cercle  $x^2 + y^2 - 1 = 0$

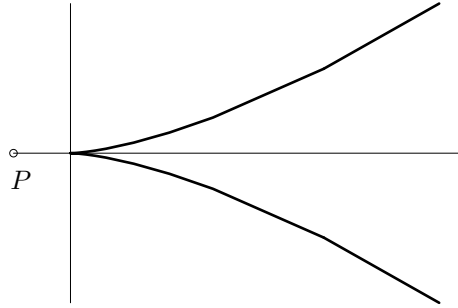


Figure II.7 – La parabole semi-cubique, d'équation  $y^2 - x^3 = 0$

### 3. Éléments de calcul des variations

On aimerait aborder des problèmes du type suivant : parmi les fonctions  $\varphi : [a, b] \rightarrow \mathbb{R}$ , de classe  $\mathcal{C}^1$ , qui vérifie la condition au bord  $\varphi(a) = A$ ,  $\varphi(b) = B$ , trouver celles pour lesquelles l'intégrale suivante

$$F(\varphi) = \int_a^b f(x, \varphi(x), \varphi'(x)) dx$$

est minimale, respectivement maximale, où  $f : \Omega \rightarrow \mathbb{R}$  est une fonction continue, et  $\Omega \ni (x, \varphi(x), \varphi'(x))$ ,  $\forall x \in [a, b]$ . Traditionnellement, on appelle  $F(\varphi)$  une *fonctionnelle* et le  $\varphi$  cherché une *extremale* de cette fonctionnelle.

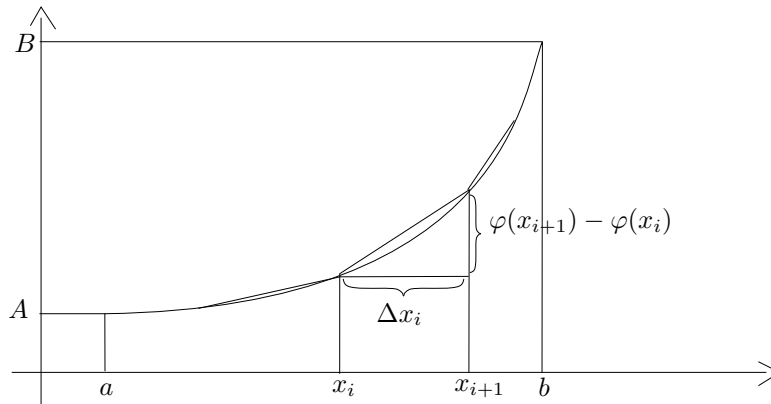


Figure II.8 – Calcul de la longueur du graphe de  $\varphi$

#### 3.1 Exemples.

(1) Soient  $P = (a, A), Q = (b, B) \in \mathbb{R}^2$ . Trouver la courbe réalisant la plus courte distance entre  $P$  et  $Q$ . On peut supposer que  $a < b$  et que la solution est donnée par le graphe d'une fonction  $\varphi : [a, b] \rightarrow \mathbb{R}$ , avec

$\varphi(a) = A, \varphi(b) = B$ . La longueur du graphe de  $\varphi$  est donnée par l'intégrale :

$$\ell(\varphi) = \int_a^b \sqrt{1 + (\varphi'(x))^2} dx \quad .$$

En effet, si  $a = x_0 < x_1 \dots < x_n = b$  est un partage de  $[a, b]$ , la longueur du segment joignant  $(x_i, \varphi(x_i))$  à  $(x_{i+1}, \varphi(x_{i+1}))$ ,  $i = 0, \dots, n-1$ , vaut, d'après le théorème de Pythagore :

$$\sqrt{\Delta x_i^2 + \Delta \varphi_i^2} \quad , \quad \text{où } \Delta x_i = x_{i+1} - x_i \quad , \quad \Delta \varphi_i = \varphi(x_{i+1}) - \varphi(x_i)$$

et la longueur de la courbe polygonale réunion de ces segments vaut

$$\sum_{i=0}^{n-1} \sqrt{\Delta x_i^2 + \Delta \varphi_i^2} = \sum_{i=0}^{n-1} \sqrt{1 + \left(\frac{\Delta \varphi_i}{\Delta x_i}\right)^2} \Delta x_i$$

ce qui tend vers  $\ell(\varphi) = \int_a^b \sqrt{1 + (\varphi'(x))^2} dx$  lorsque la maille du partage  $\sup\{\Delta x_i \mid i = 0, \dots, n-1\}$  tend vers 0. On a donc ramené le problème à la recherche de fonctions  $\varphi$  qui minimisent la fonctionnelle  $F(\varphi)$ .

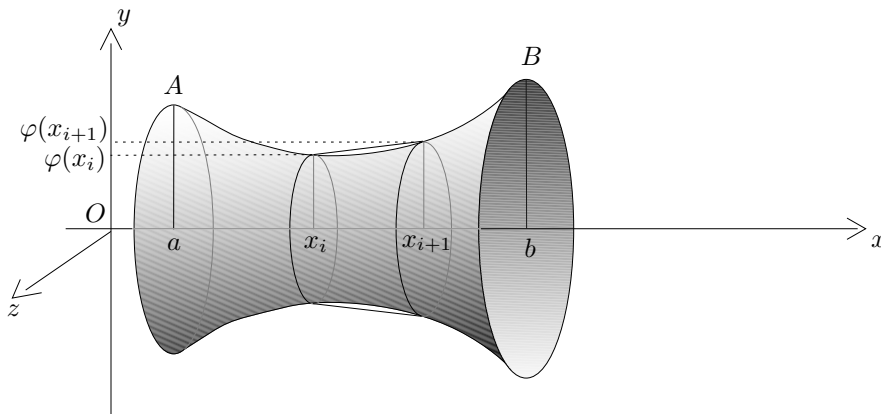
(2) On place un système de coordonnées rectangulaires  $(x, y, z)$  dans l'espace. Etant donné deux points  $P = (a, B)$  et  $Q = (b, B)$  dans le plan  $(x, y)$ , trouver  $\varphi : [a, b] \rightarrow \mathbb{R}$ ,  $\varphi(a) = A$ ,  $\varphi(b) = B$ , de sorte que l'aire  $A(\varphi)$  de la surface obtenue par rotation du graphe de  $\varphi$  autour de l'axe  $Ox$  soit minimale. Cette aire peut s'exprimer ainsi:

$$A(\varphi) = 2\pi \int_a^b \varphi(x) \sqrt{1 + \varphi'(x)^2} dx \quad .$$

En effet, rappelons que l'aire latérale d'un tronc de cône est égale à  $2\pi \frac{(R_1+R_2)}{2} r$ , où  $R_1$  et  $R_2$  sont les rayons de la base inférieure et supérieure respectivement et  $r$  est la largeur du côté. Si  $a = x_0 < \dots < x_i < \dots < x_N = b$  est un partage de l'intervalle  $[a, b]$ , l'aire de la surface engendrée par rotation de la ligne polygonale constituée par les segments d'extrémités  $(x_i, f(x_i))$ ,  $(x_{i+1}, f(x_{i+1}))$  est égale à

$$\begin{aligned} & \sum_{i=0}^{N-1} 2\pi \sqrt{\Delta x_i^2 + \Delta \varphi_i^2} \frac{(\varphi(x_i) + \varphi(x_{i+1}))}{2} = \\ & 2\pi \sum_{i=0}^{N-1} \frac{1}{2} (\varphi(x_i) + \varphi(x_{i+1})) \Delta x_i \sqrt{1 + \left(\frac{\Delta \varphi_i}{\Delta x_i}\right)^2} \end{aligned}$$

ce qui tend vers l'expression voulue lorsque les mailles du partage tendent vers 0.



**Figure II.9** – Calcul de l'aire de la surface engendrée par rotation du graphe de  $\varphi$

Supposons que l'on veuille trouver un  $\varphi_0 : [a, b] \rightarrow \mathbb{R}$  qui soit une extrémale de la fonctionnelle  $F(\varphi) = \int_a^b f(x, \varphi(x), \varphi'(x))dx$ , et vérifiant  $\varphi_0(a) = A, \varphi_0(b) = B$ . Posons :

$$E = \{ \theta : [a, b] \rightarrow \mathbb{R} \mid \theta \text{ de classe } \mathcal{C}^1 \text{ et } \theta(a) = \theta(b) = 0 \} \quad .$$

C'est un espace vectoriel, que l'on peut considérer comme l'espace des déformations de  $\varphi_0$  : pour voir que  $\varphi_0$  est une extrémale de  $F$ , il suffit d'étudier les valeurs de  $F(\varphi_0 + t \cdot \theta)$ , pour  $\theta \in E$  et  $t \in \mathbb{R}$  proche de 0.

**3.2 Proposition.** Soit  $\theta \in E$  et, pour  $t \in \mathbb{R}$  proche de 0 posons

$$h_\theta(t) = F(\varphi_0 + t \cdot \theta) = \int_a^b f(x, \varphi_0(x) + t \cdot \theta(x), \varphi_0'(x) + t \cdot \theta'(x))dx \quad .$$

Alors

$$h'_\theta(0) = \int_a^b \left\{ \frac{\partial f}{\partial y}(x, \varphi_0(x), \varphi_0'(x)) \cdot \theta(x) + \frac{\partial f}{\partial z}(x, \varphi_0(x), \varphi_0'(x)) \cdot \theta'(x) \right\} dx \quad .$$

*Preuve:* On dérive l'expression définissant  $h_\theta(t)$  sous le signe intégrale.

*q.e.d.*

**3.3 Corollaire.** Si  $\varphi_0$  est une extrémale de  $F$ , alors

$$\heartsuit \quad \int_a^b \left\{ \frac{\partial f}{\partial y}(x, \varphi_0(x), \varphi_0'(x)) \cdot \theta(x) + \frac{\partial f}{\partial z}(x, \varphi_0(x), \varphi_0'(x)) \cdot \theta'(x) \right\} dx = 0 \quad \forall \theta \in E$$

et si  $f$  est de classe  $\mathcal{C}^2$ , alors :

$$\int_a^b \left\{ \frac{\partial f}{\partial y}(x, \varphi_0(x), \varphi_0'(x)) - \frac{d}{dx} \left( \frac{\partial f}{\partial z}(x, \varphi_0(x), \varphi_0'(x)) \right) \right\} \cdot \theta(x) dx = 0 \quad \forall \theta \in E$$

*Preuve:* Si  $\varphi_0$  est une extrémale de  $F$ , alors pour tout  $\theta \in E$   $0 \in \mathbb{R}$  est un extremum pour  $h_\theta(t)$ , et donc, d'après **1.12**,  $h'_\theta(0) = 0$ , ce qui implique l'équation  $\heartsuit$  par la proposition **3.2**.

Si  $f$  est de classe  $\mathcal{C}^2$ , on peut intégrer par parties :

$$\int_a^b \frac{\partial f}{\partial z}(x, \varphi_0(x), \varphi_0'(x)) \cdot \theta'(x) dx = \underbrace{\frac{\partial f}{\partial z}(t, \varphi_0(t), \varphi_0'(t)) \cdot \theta(t)}_{=0} \Big|_a^b - \int_a^b \frac{d}{dx} \left( \frac{\partial f}{\partial z}(x, \varphi_0(x), \varphi_0'(x)) \right) \cdot \theta(x) dx$$

et la deuxième égalité en suit.

*q.e.d.*

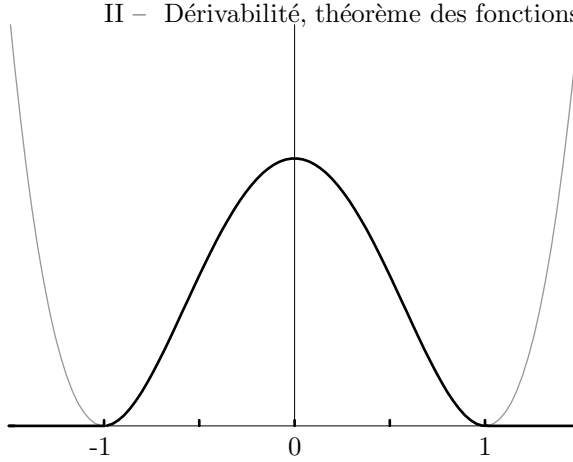
**3.4 Lemme.** Soit  $g : [a, b] \rightarrow \mathbb{R}$  continue et supposons que

$$\int_a^b g(x)\theta(x)dx = 0 \quad \forall \theta \in E \quad .$$

Alors  $g$  est identiquement nulle.

*Preuve:* Si  $g$  est identiquement nulle sur  $]a, b[$  alors  $g$  est identiquement nulle sur  $[a, b]$ . Donc si la conclusion du lemme est fautive, il existe  $x_0 \in ]a, b[$  avec  $g(x_0) \neq 0$ . On peut supposer, quitte à remplacer  $g$  par  $-g$ , que  $g(x_0) > 0$ . Puisque  $g$  est continue, il existe  $\delta > 0$  tel que  $g(x) \geq \frac{g(x_0)}{2}$  si  $|x - x_0| \leq \delta$ , et on peut supposer que  $[x_0 - \delta, x_0 + \delta] \subset [a, b]$ . On vérifie facilement que la fonction

$$f(x) = \begin{cases} (x - 1)^2(x + 1)^2 & \text{si } |x| \leq 1 \\ 0 & \text{sinon} \end{cases}$$



**Figure II.10** – Graphe de la fonction ”en cloche”  $(x-1)^2(x+1)^2$ , prolongée par 0 en dehors de  $[-1, 1]$

est de classe  $\mathcal{C}^1$  (voir figure **II.10**); posons  $\theta(x) = f\left(\frac{x-x_0}{\delta}\right)$ . Alors  $\theta$  est aussi  $\mathcal{C}^1$ ,  $\theta(a) = \theta(b) = 0$  et  $\theta(x) \geq 0$ . Donc  $g(x) \cdot \theta(x) \geq 0$ , mais  $g \cdot \theta$  est non identiquement nulle, et elle est continue. Alors, d’après le lemme **I.1.3**

$$\int_a^b g(x)\theta(x)dx > 0$$

ce qui contredit l’hypothèse.

*q.e.d.*

**3.5 Remarque.** On peut même construire des fonctions  $\mathcal{C}^\infty$  du même type que la fonction  $\theta(x)$  de la preuve du lemme précédent en posant :

$$\theta(x) = \begin{cases} e^{\frac{-1}{\delta^2 - (x-x_0)^2}} & \text{si } |x - x_0| < \delta \\ 0 & \text{sinon} \end{cases}$$

et on démontre qu’elle est  $\mathcal{C}^\infty$  de manière analogue à l’exemple que l’on trouve dans **H-W, chap. III(7.12)** où l’on traite le cas de  $f(x) = e^{-1/x^2}$  si  $x > 0$ ,  $f(x) = 0$  sinon.

**3.6 Théorème.** *Supposons que  $f$  soit de classe  $\mathcal{C}^2$ ; une condition nécessaire pour que  $\varphi_0$  soit une extrémale de la fonctionnelle :*

$$\int_a^b f(x, \varphi(x), \varphi'(x))dx$$

*est que l’équation différentielle suivante, appelé équation d’Euler-Lagrange, soit satisfaite :*

$$(I) \quad \frac{\partial f}{\partial y}(x, \varphi_0(x), \varphi_0'(x)) - \frac{d}{dx} \left( \frac{\partial f}{\partial z}(x, \varphi_0(x), \varphi_0'(x)) \right) = 0$$

*et si  $f$  ne dépend pas de  $x$ , alors l’équation suivante est satisfaite :*

$$(II) \quad \varphi_0'(x) \frac{\partial f}{\partial z}(\varphi_0(x), \varphi_0'(x)) - f(\varphi_0(x), \varphi_0'(x)) = \text{Constante}$$

*Preuve:* (I) suit de **3.3** et **3.4**. Si  $f$  ne dépend que de  $y$  et  $z$ , alors  $\frac{\partial f}{\partial x} \equiv 0$  et il donc :

$$\frac{d}{dx} \left( \varphi_0' \frac{\partial f}{\partial z} - f \right) = \varphi_0'' \frac{\partial f}{\partial z} + \varphi_0' \frac{d}{dx} \left( \frac{\partial f}{\partial z} \right) - \frac{\partial f}{\partial y} \varphi_0' - \frac{\partial f}{\partial z} \varphi_0'' = \varphi_0' \left( \frac{d}{dx} \left( \frac{\partial f}{\partial z} \right) - \frac{\partial f}{\partial y} \right) = 0$$

où les dérivées sont toutes prises au point  $(x, \varphi_0(x), \varphi'_0(x))$ ; d'où l'on tire (II).

*q.e.d.*

### 3.7 Exemples.

On reprend les exemples du début de ce paragraphe; nous nous contenterons d'utiliser les équations d'Euler-Lagrange pour déterminer les solutions possibles, sans montrer qu'il s'agit effectivement de minima (voir G.A. Bliss, *Calculus of variations*, the Open Court Publishing Company, Chicago (1925) pour une analyse plus complète de ces problèmes.)

(1) Pour trouver la plus courte distance entre  $P = (a, A)$  et  $Q = (b, B)$ , on doit minimiser la fonctionnelle :

$$\ell(\varphi) = \int_a^b \sqrt{1 + \varphi'(x)^2} dx \quad .$$

Ici on a :

$$f(x, y, z) = f(z) = \sqrt{1 + z^2} \quad , \quad \frac{\partial f}{\partial z} = \frac{z}{\sqrt{1 + z^2}}$$

et on déduit de **3.6(II)** que si  $\varphi_0$  est un minimum de  $\ell(\varphi)$ , alors :

$$\varphi'_0 \frac{\varphi'_0}{\sqrt{1 + \varphi_0'^2}} - \sqrt{1 + \varphi_0'^2} = C$$

d'où l'on tire que

$$\varphi_0'^2 - 1 - \varphi_0'^2 = C \sqrt{1 + \varphi_0'^2} \Rightarrow \varphi_0'^2 = \text{Const.} \Rightarrow \varphi_0 = \text{const.}$$

et donc  $\varphi_0$  est de la forme  $\varphi_0(x) = \alpha x + \beta$ , comme on pouvait s'y attendre.

(2) Pour trouver la fonction dont le graphe engendre par rotation autour de  $Ox$  la surface d'aire minimale, on doit minimiser la fonctionnelle:

$$A(\varphi) = 2\pi \int_a^b \varphi(x) \sqrt{1 + \varphi'(x)^2} dx \quad .$$

Le coefficient  $2\pi$  ne joue pas de rôle; on prend donc :

$$f(y, z) = y \sqrt{1 + z^2} \quad , \quad \frac{\partial f}{\partial z} = \frac{yz}{\sqrt{1 + z^2}}$$

d'où l'on tire, en appliquant **3.6(II)**, l'équation:

$$\varphi'_0 \frac{\varphi_0 \varphi'_0}{\sqrt{1 + \varphi_0'^2}} - \varphi_0 \sqrt{1 + \varphi_0'^2} = C \quad .$$

Il en suit que

$$\frac{\varphi'_0}{\sqrt{(\varphi_0/C)^2 - 1}} = 1$$

Or

$$\left( \operatorname{arccosh}(\varphi_0/C) \right)' = \frac{\varphi'_0}{\sqrt{\varphi_0^2 - C^2}} = 1/C$$

où  $\operatorname{arccosh}(y)$  est la fonction inverse de  $\cosh(x) = \frac{e^x + e^{-x}}{2}$ , d'où l'on tire que

♠

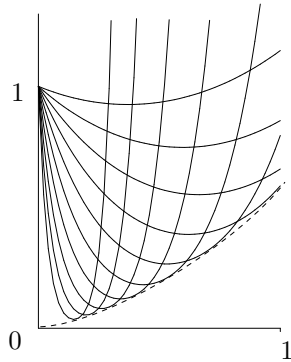
$$\varphi_0(x) = C \cdot \cosh \left( \frac{x}{C} - \alpha \right)$$

où  $\alpha$  et  $C$  sont des constantes.

Il reste à savoir si pour  $P$  et  $Q$  donnés on peut déterminer les constantes  $\alpha$  et  $C$  telles que le graphe du  $\varphi_0$  correspondant passe par  $P$  et  $Q$ . Or la discussion de ce problème est assez compliquée, et nous nous contenterons de l'esquisser. On peut supposer que  $P = (0, 1)$  sans perte de généralité; Il se trouve que si  $Q$  est à gauche de la courbe en pointillé dans la figure **II.11**, il existe une solution qui passe par  $P$  et  $Q$ . Si  $Q$  se trouve à droite de la courbe en pointillé, il n'existe pas de solution qui passe par  $P$  et  $Q$ . Dans ce cas il y a une solution au problème, à condition de l'énoncer un peu différemment. Par rotation autour de l'axe  $Ox$ , les points  $P$  et  $Q$  décrivent deux cercles  $C_P$  et  $C_Q$ ; le problème est de trouver la surface d'aire minimale ayant les cercles  $C_P$  et  $C_Q$  comme bord. Lorsque  $Q$  est au-delà de la courbe en pointillé, on a comme solution tout simplement les disques dont  $C_P$  et  $C_Q$  sont les bords.

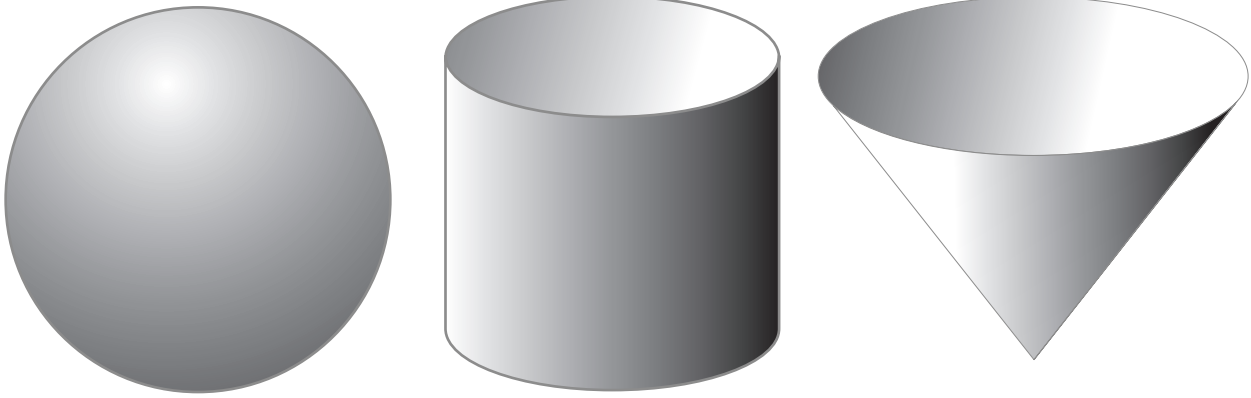
On ne peut pas donner une formule explicite pour la courbe en pointillé, qui est en fait l'enveloppe de la sous-famille des courbes de la famille  $\spadesuit$  qui passent par  $P$ . (voir C. Carathéodory, *Variationsrechnung*, Teubner Verlag (1935), Leipzig und Berlin, page 301).

Physiquement, on peut obtenir ces surfaces en réalisant  $C_P$  et  $C_Q$  en fil de fer et la surface cherchée par une pellicule d'eau savonneuse. On voit bien que si on part de deux cercles proches, bords d'une pellicule d'eau savonneuse d'un seul tenant, et qu'on éloigne de plus en plus les deux cercles, la pellicule finira par se séparer pour former (avec un peu de chance) deux disques, de bord  $C_P$ , respectivement  $C_Q$ .



**Figure II.11** – Graphes engendrant des surfaces minimales par rotation autour de l'axe  $Ox$





**Figure II.12** – La sphère, le cylindre et le cône

### 3.1 GÉODÉSQUES SUR LES SURFACES

Nous considérerons des surfaces de  $\mathbb{R}^3$  données par une équation de la forme  $G(x, y, z) = 0$ , où  $\Omega \subset \mathbb{R}^3$  est ouvert et  $G : \Omega \rightarrow \mathbb{R}$  est de classe au moins  $\mathcal{C}^2$ . Voici 3 exemples concrets de surface, où  $\Omega = \mathbb{R}^3$  :

- la sphère, d'équation  $x^2 + y^2 + z^2 - 1 = 0$
- le cylindre, d'équation  $x^2 + y^2 - 1 = 0$
- le cône, d'équation  $x^2 + y^2 - z^2 = 0$ .

Une courbe paramétrique régulière de l'espace est donnée par une application  $\gamma : [a, b] \rightarrow \mathbb{R}^3$ , que nous supposons de classe  $\mathcal{C}^2$ , vérifiant  $\gamma'(t) \neq 0 \forall t \in [a, b]$ . Nous noterons  $\gamma(t) = (x(t), y(t), z(t))$  et  $\gamma'(t) = (\dot{x}(t), \dot{y}(t), \dot{z}(t))$  sa dérivée. Nous noterons par  $G_x, G_y$  et  $G_z$  les dérivées partielles respectives de  $G$ .

Etant donnée une surface  $Z(G)$  d'équation  $G(x, y, z) = 0$  et deux points  $P, Q \in Z(G)$ , on se pose le problème de trouver la courbe régulière  $\gamma : [a, b] \rightarrow \mathbb{R}^3$  contenue dans la surface  $Z(G)$ , qui joint  $P$  à  $Q$ , et dont la longueur est minimale. Une telle courbe, qu'on appelle géodésique de la surface  $Z(G)$ , doit satisfaire :

- 1)  $G(x(t), y(t), z(t)) = 0 \forall t \in [a, b]$
- 2)  $\gamma(a) = P, \gamma(b) = Q$
- 3) La fonctionnelle

$$\int_0^1 \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2} \quad ,$$

qui, d'après un raisonnement semblable à celui de l'exemple **3.1(1)** représente la longueur de  $\gamma$ , doit être minimale.

Supposons plus généralement que nous cherchions les extremales d'une fonctionnelle de la forme

$$(3-1) \quad \int_a^b f(x, y, z, \dot{x}, \dot{y}, \dot{z}) dt$$

où  $x, y, z, \dot{x}, \dot{y}, \dot{z}$  représentent les coordonnées d'une courbe sur la surface  $Z(G)$  et de sa dérivée; une solution sera une courbe  $(x(t), y(t), z(t))$  qui rend la fonctionnelle ci-dessus extrémale, et qui de plus satisfait  $G(x(t), y(t), z(t)) = 0$ . Au voisinage d'un point régulier de  $Z(G)$ , quitte à échanger les rôles de  $x, y$  et  $z$ , on peut supposer que  $G_z \neq 0$ , et il suit alors du théorème des fonctions implicites que l'on peut paramétrer localement  $Z(G)$  par une application de la forme  $(x, y) \mapsto (x, y, g(x, y))$ . On aura :

$$(3-2) \quad g_x = -\frac{G_x}{G_z} \quad , \quad g_y = -\frac{G_y}{G_z}$$

et si  $\gamma(t) = (x(t), y(t), z(t))$  est une courbe sur  $Z(G)$ , on aura que  $z(t) = g(x(t), y(t))$ , et donc  $\dot{z}(t) = g_x(x(t), y(t)) \cdot \dot{x}(t) + g_y(x(t), y(t)) \cdot \dot{y}(t)$ . Posons :

$$(3-3) \quad h(x, y) = f(x, y, g(x, y), \dot{x}, \dot{y}, \dot{x}, \dot{y}, g_x(x, y) \cdot \dot{x} + g_y(x, y) \cdot \dot{y}) \quad .$$

Alors le problème revient à trouver les extremales de la fonctionnelle

$$(3-4) \quad \int_a^b h(x, y, \dot{x}, \dot{y}) dt \quad .$$

Ce que nous avons gagné par rapport à (3-1) c'est que les variables  $x, y, \dot{x}, \dot{y}$  sont libres, alors que dans (3-1) on avait la contrainte supplémentaire  $G(x, y, z) = 0$ .

La fonctionnelle de (3-3) est plus générale que celle du type considéré pour obtenir les équations d'Euler-Lagrange. Mais on a le théorème suivant, qui se démontre de manière analogue au théorème 3.6 :

**3.8 Théorème.** *Si  $(x(t), y(t))$ ,  $t \in [a, b]$ , est une extremale de (3-4), alors elle satisfait les équations différentielles suivantes :*

$$\begin{cases} h_x(x(t), y(t), \dot{x}(t), \dot{y}(t)) - \frac{d}{dt}(h_{\dot{x}}(x(t), y(t), \dot{x}(t), \dot{y}(t))) = 0 \\ h_y(x(t), y(t), \dot{x}(t), \dot{y}(t)) - \frac{d}{dt}(h_{\dot{y}}(x(t), y(t), \dot{x}(t), \dot{y}(t))) = 0 \end{cases}$$

■

Ces équations s'appellent encore équations d'Euler-Lagrange.

Dans notre cas, compte tenu de (3-3), elles donnent :

$$\begin{aligned} h_x - \frac{d}{dt}f_x + f_z \cdot g_x + f_{\dot{z}} \cdot (g_{xx} \cdot \dot{x} + g_{xy} \cdot \dot{y}) - \frac{d}{dt}(f_{\dot{x}}) - \underbrace{\frac{d}{dt}(f_{\dot{z}} \cdot g_x)} &= 0 \\ &= f_{\dot{z}} \cdot (g_{x,x} \cdot \dot{x} + g_{x,y} \cdot \dot{y}) + g_x \cdot \frac{d}{dt}(f_{\dot{z}}) \end{aligned}$$

d'où :

$$(3-5) \quad f_x - \frac{d}{dt}(f_{\dot{x}}) + g_x \cdot (f_z - \frac{d}{dt}(f_{\dot{z}})) = 0$$

et, en échangeant  $x$  et  $y$  :

$$(3-6) \quad f_y - \frac{d}{dt}(f_{\dot{y}}) + g_y \cdot (f_z - \frac{d}{dt}(f_{\dot{z}})) = 0$$

Définissons la fonction  $\lambda(t)$  par l'équation :

$$(3-7) \quad \frac{d}{dt}(f_{\dot{z}}) - f_z = \lambda \cdot G_z$$

on tire alors de (3-5) et (3-6), compte tenu de (3-2) :

$$\begin{aligned} \frac{d}{dt}f_{\dot{x}} - f_x &= g_x(f_z - \frac{d}{dt}(f_{\dot{z}})) = -\frac{G_x}{G_z}(f_z - \frac{d}{dt}(f_{\dot{z}})) = \lambda \cdot G_x \\ \frac{d}{dt}f_{\dot{y}} - f_y &= \lambda \cdot G_y \end{aligned}$$

donc finalement on obtient les 3 équations :

$$(3-8) \quad \begin{cases} \frac{d}{dt}f_{\dot{x}} - f_x = \lambda \cdot G_x \\ \frac{d}{dt}f_{\dot{y}} - f_y = \lambda \cdot G_y \\ \frac{d}{dt}f_{\dot{z}} - f_z = \lambda \cdot G_z \end{cases}$$

qui seront valables en tous les points réguliers de  $Z(G)$ . Dans le cas particulier où  $f(x, y, z, \dot{x}, \dot{y}, \dot{z}) = f(\dot{x}, \dot{y}, \dot{z}) = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}$ , on obtient le système d'équations suivant, que devront satisfaire les géodésiques :

$$(3-9) \quad \begin{cases} \frac{\ddot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} = \lambda \cdot G_x \\ \frac{\ddot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} = \lambda \cdot G_y \\ \frac{\ddot{z}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} = \lambda \cdot G_z \end{cases} .$$

Pour le résoudre, il est plus commode de supposer que  $(x(t), y(t), z(t))$  est une paramétrisation par la longueur d'arc, que nous allons maintenant rappeler.

Si  $\gamma : [a, b] \rightarrow \mathbb{R}^3$  une courbe régulière, posons

$$s(t) = \int_a^t \|\gamma'(\tau)\| d\tau$$

où  $\|\cdot\|$  est la norme euclidienne;  $s(t)$  représente la longueur du morceau de courbe  $\gamma([a, t])$ , pour  $a \leq t \leq b$ , et prend ses valeurs dans l'intervalle  $[0, L]$ , où  $L = s(b)$  est la longueur de la courbe. Puisque  $s'(t) = \|\gamma'(t)\| \neq 0$ ,  $s$  est monotone croissante, en fait une bijection entre  $[a, b]$  et  $[0, L]$ . On peut donc reparamétriser la courbe en posant  $\hat{\gamma}(s(t)) = \gamma(t)$ , ce qui définit une nouvelle paramétrisation  $\hat{\gamma} : [0, L] \rightarrow \mathbb{R}^3$  de la même courbe; on dit que  $\hat{\gamma}$  est une paramétrisation par la longueur d'arc. On a :

$$\hat{\gamma}(s(t))' = \gamma'(s(t)) \cdot s'(t) = \gamma'(t) \Rightarrow \|\hat{\gamma}'(s(t))\| \cdot \underbrace{\|s'(t)\|}_{=\|\gamma'(t)\|} = \|\gamma'(t)\| \Rightarrow \|\hat{\gamma}'(s)\| = 1 \quad .$$

On note maintenant par  $\gamma(t) = (x(t), y(t), z(t)) : [a, b] \rightarrow \mathbb{R}^3$  une courbe paramétrée par la longueur d'arc; puisque  $\|\gamma'(t)\| = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} = 1$ , le système d'équations (3-9) s'écrit :

$$(3-10) \quad \begin{cases} \ddot{x} = \lambda \cdot G_x \\ \ddot{y} = \lambda \cdot G_y \\ \ddot{z} = \lambda \cdot G_z \end{cases}$$

ou encore, sous forme vectorielle :

$$(3-11) \quad \ddot{\gamma} = \lambda dG \quad .$$

Autrement dit, l'accélération  $\ddot{\gamma}$  de la courbe géodésique doit être perpendiculaire au plan tangent à la surface (cf. **2.4(2)**).

### 3.9 Exemples.

(1) Prenons le plan  $G(x, y, z) = z = 0$ . Alors  $\gamma(t) = (x(t), y(t), 0)$  et l'équation (3-11) donne :

$$(\ddot{x}, \ddot{y}, 0) = \lambda(0, 0, 1) \Rightarrow \lambda = 0 \Rightarrow \ddot{x} = \ddot{y} = 0$$

ce montre que les géodésiques sont des droites.

(2) Pour la sphère  $G(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$ , on a que  $dG = 2(x, y, z)$ , et de (3-10) on tire alors que  $\ddot{\gamma} = 2\lambda \cdot \gamma$ . Soit  $N(t) = \gamma(t) \times \dot{\gamma}(t)$ , où  $\times$  dénote le produit vectoriel, le vecteur perpendiculaire au plan engendré par  $\gamma(t)$  et  $\dot{\gamma}(t)$ . Alors :

$$\dot{N} = \underbrace{\dot{\gamma} \times \dot{\gamma}}_{=0} + \gamma \times \ddot{\gamma} = 2\lambda \gamma \times \gamma = 0$$

et il en suit que  $N$  est constant; donc la géodésique se trouve dans le plan perpendiculaire au vecteur constant  $N$  et c'est donc une portion de grand cercle de la sphère.

#### 4. Théorèmes de l'application inverse et du rang

Ces deux théorèmes sont des conséquences du théorème des fonctions implicites. Le théorème de l'application inverse permet de construire des changements de coordonnées locaux; le théorème du rang permet d'écrire une application ayant une dérivée de rang maximum en un point comme une application linéaire, après changement de coordonnées local.

**4.1 Théorème de l'application inverse.** Soit  $f : U \rightarrow \mathbb{R}^n$ ,  $U \subset \mathbb{R}^n$  ouvert, de classe  $\mathcal{C}^1$ , et supposons qu'en un point  $x_0 \in U$  la dérivée  $df_{x_0} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$  soit inversible; posons  $y_0 = f(x_0)$ . Il existe un ouvert  $V \subset \mathbb{R}^n$ ,  $V \ni x_0$ ,  $r > 0$  et une application  $g : B(y_0, r) \rightarrow V$  de classe  $\mathcal{C}^1$  telle que

$$f \circ g = \mathbb{I}_{B(y_0, r)} \quad , \quad g \circ f|_V = \mathbb{I}_V$$

où  $\mathbb{I}_-$  dénote l'application identité. De plus,  $dg_{y_0} = df_{x_0}^{-1}$ .

On dira que  $g$  est un inverse local de  $f$  au voisinage de  $x_0$ .

*Preuve:* Posons

$$F(x, y) = y - f(x) \quad y \in \mathbb{R}^n, x \in U \quad .$$

Alors  $F(x_0, y_0) = 0$  et chercher l'inverse local de  $f$  revient à résoudre explicitement  $x$  par rapport à  $y$  dans l'équation  $F(x, y) = 0$ . Puisque  $\frac{\partial F}{\partial x}(x_0, y_0) = -df_{x_0}$  est inversible, on peut appliquer le théorème des fonctions implicites **2.2**, avec la remarque **2.3(1)** (notez que les rôles de  $x$  et  $y$  sont inversés) :

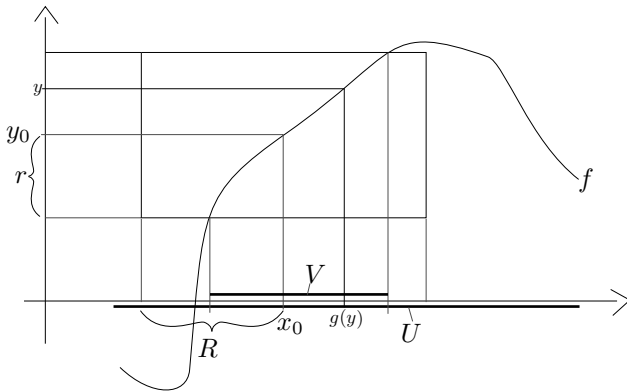
$\exists r, R > 0$  et  $g : B(y_0, r) \rightarrow B(x_0, R)$  tels que pour  $(x, y) \in B(y_0, r) \times B(x_0, R)$  on a :  $x = f(y) \Leftrightarrow y = g(x)$ .

et d'après **2.9**  $g$  est  $\mathcal{C}^1$ . Posons  $V = f^{-1}(B(y_0, r)) \cap B(x_0, R)$ . Montrons que  $f|_V$  et  $g$  sont inverse l'une de l'autre:

- si  $y \in B(y_0, r)$ ,  $g(y) = x \in B(x_0, R)$ , et donc  $F(x, y) = f(x) - y = 0 \Rightarrow y = f(x)$ ; cela entraîne aussi que  $g(x) \in V$ .
- si  $x \in V$ ,  $x \in B(x_0, R)$  et  $y = f(x) \in B(y_0, r)$ , donc  $g(y) = x$ .

Enfin, en dérivant les 2 membres de l'équation  $f(g(y)) = y$ , pour  $y \in B(y_0, r)$ , on obtient l'expression de  $dg_{y_0}$ .

*q.e.d.*



**Figure II.13** – Le théorème d'inversion locale

**4.2 Définition.** Soit  $h : U \rightarrow V$ ,  $U, V \subset \mathbb{R}^n$  des ouverts. On dit que  $h$  est un difféomorphisme si  $h$  est  $\mathcal{C}^1$ , bijective, et d'inverse aussi  $\mathcal{C}^1$ . On dit que  $h$  est un difféomorphisme local au voisinage de  $x_0 \in U$  s'il existe des ouverts  $U', V'$ ,  $x_0 \in U' \subset U$ ,  $h(x_0) \in V' \subset V$ , tel que  $h|_{U'}$  soit un difféomorphisme entre  $U'$  et  $V'$ .

On peut donc résumer l'énoncé du théorème de l'application inverse en disant que si la dérivée de  $f$  en un point  $x_0$  est inversible, alors  $f$  est un difféomorphisme local au voisinage de  $x_0$ .

### 4.3 Exemples.

(1) Coordonnées polaires. Il s'agit de l'application  $f(\rho, \theta) = (\rho \cos(\theta), \rho \sin(\theta))$ ,  $\rho, \theta \in \mathbb{R}^2$ ,  $\rho \geq 0$ . On a :

$$f'(\rho, \theta) = \begin{pmatrix} \cos(\theta) & -\rho \sin(\theta) \\ \sin(\theta) & \rho \cos(\theta) \end{pmatrix} \text{ et } \det(f'(\rho, \theta)) = \rho$$

donc  $f$  est un difféomorphisme local au voisinage de tout point  $(\rho, \theta)$  avec  $\rho > 0$ . Si on pose

$$U = \{(\rho, \theta) \mid \rho > 0, -\pi < \theta < \pi\} \quad , \quad V = \mathbb{R}^2 \setminus \{(x, 0) \mid x \leq 0\}$$

alors  $f$  est un difféomorphisme entre  $U$  et  $V$ . Son inverse peut s'écrire ainsi :

$$f^{-1}(x, y) = \left( \sqrt{x^2 + y^2}, \arcsin \left( \frac{y}{\sqrt{x^2 + y^2}} \right) \right)$$

où l'on prend arcsin comme étant l'inverse de  $\sin(\theta)$  restreint à l'intervalle  $]-\pi, \pi[$ . Cette application a la vertu de transformer des cercles centrés à l'origine en des segments de droite, et les disque centrés à l'origine en des rectangles.

(2) L'application  $f(x_1, x_2) = (x_1, x_1 \cdot x_2)$  a pour dérivée

$$f'(x_1, x_2) = \begin{pmatrix} 1 & 0 \\ x_2 & x_1 \end{pmatrix} \text{ et } \det(f'(x_1, x_2)) = x_1$$

c'est donc un difféomorphisme local au voisinage de tout point  $(x_1, x_2)$ , avec  $x_1 \neq 0$ . Si on pose

$$U = \{(x_1, x_2) \mid 0 < x_i < 1, i = 1, 2\} \text{ et } V = \{(x_1, x_2) \mid 0 < x_1 < 1, 0 < x_2 < x_1\}$$

on vérifie que  $f$  est un difféomorphisme entre  $U$  et  $V$ . Cette application a donc la vertu de transformer le carré  $U$  en le triangle  $V$ .

**4.4 Théorème du rang.** Soient  $U \subset \mathbb{R}^n$  un ouvert,  $f : U \rightarrow \mathbb{R}^p$  de classe  $\mathcal{C}^1$  et  $x_0 \in U$ . Supposons que  $df_{x_0} : \mathbb{R}^n \rightarrow \mathbb{R}^p$  soit une application linéaire de rang maximum, c'est à dire de rang  $n$  si  $n \leq p$ , de rang  $p$  si  $n \geq p$ . Alors:

(1) Si  $n \leq p$ , il existe un ouvert  $U' \subset U$ ,  $U' \ni x_0$  et un difféomorphisme  $H : V' \rightarrow V$ ,  $V$  et  $V'$  ouverts de  $\mathbb{R}^p$ , tels que  $f(U') \subset V$  et

$$H^{-1}(f(x_1, \dots, x_n)) = (x_1, \dots, x_n, \underbrace{0, \dots, 0}_{p-n}) \quad \forall (x_1, \dots, x_n) \in U' \quad .$$

(2) Si  $n \geq p$ , il existe un ouvert  $U' \ni x_0$  de  $\mathbb{R}^n$ ,  $U' \subset U$ , et un difféomorphisme  $h : U' \rightarrow V$ ,  $V$  ouvert de  $\mathbb{R}^n$ , tel que

$$f(h^{-1}(x_1, \dots, x_n)) = (x_1, \dots, x_p) \quad \forall (x_1, \dots, x_n) \in V \quad .$$

*Preuve:* (1) Quitte à renuméroter les coordonnées au but, on peut supposer que  $\det \left( \frac{\partial f_i}{\partial x_j}(x_0) \right)_{i,j=1,\dots,n} \neq 0$ . Posons

$$H(x_1, \dots, x_n, x_{n+1}, \dots, x_p) = f(x_1, \dots, x_n) + \underbrace{(0, \dots, 0)}_n, x_{n+1}, \dots, x_p \quad .$$

Alors

$$dH_{(x_0,0)} = \left( \begin{array}{c|cccc} \left( \frac{\partial f_i}{\partial x_j}(x_0) \right)_{i,j=1,\dots,p} & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \\ \hline & 1 & 0 & \cdots & 0 \\ & 0 & 1 & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & 0 & \cdots & 0 & 1 \end{array} \right)$$

est inversible et  $H(x_0, 0) = f(x_0)$ , donc par 4.1  $H$  restreint à un ouvert  $V' \ni (x_0, 0)$  est un difféomorphisme sur un ouvert  $V \ni f(x_0)$ .

Or

$$H(x_1, \dots, x_n, 0, \dots, 0) = f(x_1, \dots, x_n) \implies H^{-1}(f(x_1, \dots, x_n)) = (x_1, \dots, x_n, 0, \dots, 0) \quad .$$

(2) Si  $n \geq p$ , on peut supposer que  $\det \left( \frac{\partial f_i}{\partial x_j}(x_0) \right)_{i,j=1,\dots,p} \neq 0$ , quitte à renuméroter les coordonnées à la source. Si l'on pose

$$h(x_1, \dots, x_n) = (f(x_1, \dots, x_n), x_{p+1}, \dots, x_n)$$

alors:

$$dh_{x_0} = \left( \begin{array}{cccccc} \frac{\partial f_i}{\partial x_j}(x_0) & & & & & \\ \hline 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & & & 0 & 1 & \ddots & \vdots \\ \vdots & & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{array} \right)$$

est inversible, donc d'après 2.1  $h$  est un difféomorphisme d'un ouvert  $U' \ni x_0$  sur un ouvert  $V$  de  $\mathbb{R}^n$ . Or

$$\begin{aligned} (x_1, \dots, x_n) &= h(h^{-1}(x_1, \dots, x_n)) = (f_1(h^{-1}(x_1, \dots, x_n)), \dots, f_p(h^{-1}(x_1, \dots, x_n)), *, \dots, *) \\ \implies f(h^{-1}(x_1, \dots, x_n)) &= (x_1, \dots, x_p) \end{aligned}$$

*q.e.d.*

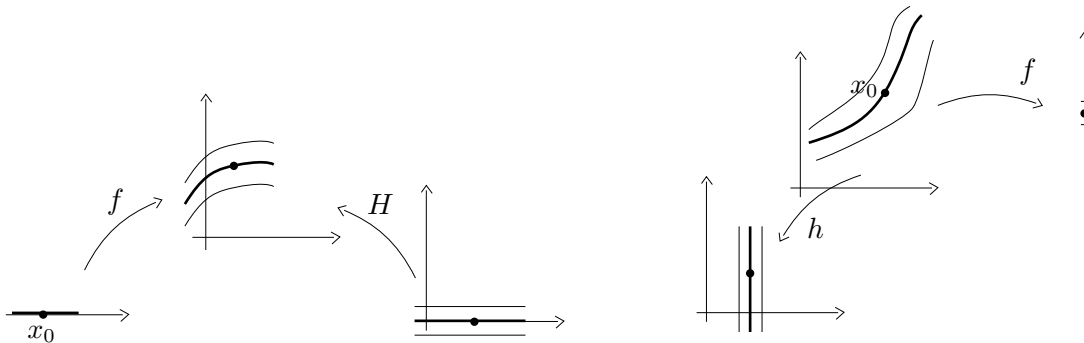


Figure II.14 – Le théorème du rang lorsque  $n = 1, p = 2$  et  $n = 2, p = 1$

4.1 SOUS-VARIÉTÉS DE  $\mathbb{R}^n$ 

**4.5 Définition.** On dit que  $X \subset \mathbb{R}^n$  est une sous-variété de dimension  $k$  si  $\forall x_0 \in X, \exists U_{x_0} \subset \mathbb{R}^n$  ouvert et  $f = (f_1, \dots, f_{n-k}) : U_{x_0} \rightarrow \mathbb{R}^{n-k}$  différentiable telle que:

- (1)  $X \cap U_{x_0} = f^{-1}(0) = \{x \in U_{x_0} \mid f_i(x) = 0, i = 1, \dots, n - k\}$
- (2)  $\forall x \in U_{x_0}$  la dérivée  $df_x : \mathbb{R}^n \rightarrow \mathbb{R}^{n-k}$  est surjective.

On appelle  $n - k$  la codimension de  $X$ . L'application  $f$  est appelée équation locale de  $X$  au voisinage de  $x_0$

En d'autres termes, une sous-variété est un sous-ensemble de  $\mathbb{R}^n$  qui admet en tout point un système d'équations qui satisfait les hypothèses du théorème **2.2** (théorème des fonctions implicites – voir remarque **2.3(2)**).

On parle de courbes lisses ou de surfaces lisses dans le cas de sous-variétés de dimension 1, respectivement 2.

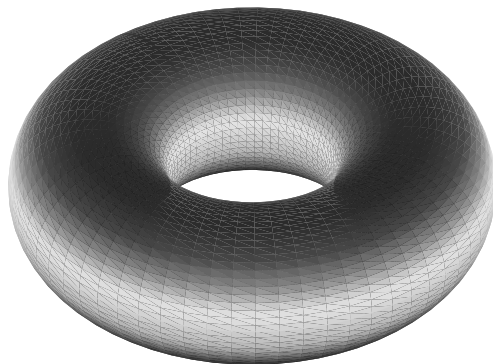
Plus généralement, on dira que  $X \subset \mathbb{R}^n$  est une sous-variété de dimension  $k$  au voisinage d'un point  $x_0 \in X$  s'il existe un ouvert  $U \subset \mathbb{R}^n, U \ni x_0$ , tel que  $U \cap X$  est une sous-variété de  $\mathbb{R}^n$ .

**4.6 Remarques.**

(1) Souvent une seule application  $f : U \rightarrow \mathbb{R}^k$  suffit à décrire une sous-variété. C'est le cas pour tous les exemples ci-dessous, sauf le 3ème.

(2)  $X$  est "localement fermée" dans  $\mathbb{R}^n$ :  $U_{x_0} \cap X$  est fermé dans  $U_{x_0}$  car  $f^{-1}(0)$  est un fermé.

(3) Dans la condition (2) de la définition de sous-variété il suffirait de supposer que  $df_{x_0}$  est surjective, car alors  $df_x$  sera surjective pour  $x$  dans un ouvert  $U'_{x_0}$  de  $x_0$ , que l'on peut substituer à  $U_{x_0}$ .



**Figure II.15** – Le tore

**4.7 Exemples.**

(1) Le cercle  $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 - 1 = 0\}$  est une sous-variété de dimension et codimension 1 de  $\mathbb{R}^2$ . En effet, on peut prendre un même ouvert valable pour tout  $x_0 \in S^1$ :  $U = U_{x_0} = \mathbb{R}^2 \setminus \{0\}$ . Puisque  $df_{(x,y)} = (2x, 2y) \neq 0$  si  $(x, y) \in U$ , les conditions (1) et (2) de la définition de sous-variété sont satisfaites.

Plus généralement, la  $n - 1$ -sphère  $S^{n-1}$  est la sous-variété de  $\mathbb{R}^n$  de dimension  $n - 1$  et codimension 1, définie par:

$$S^{n-1} = \left\{ x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^2 - 1 = 0 \right\}$$

et son équation  $f(x_1, \dots, x_n) = \sum_{i=1}^n x_i^2 - 1$  est de rang maximum, car son gradient vaut  $2(x_1, \dots, x_n)$ , qui est non nul sur  $\mathbb{R}^n \setminus \{0\}$ .

(2) Le tore est la figure de l'espace obtenue en faisant tourner autour de l'axe  $Oz$  un cercle de rayon  $r$  placé dans le plan  $yz$ , centré en  $(0, R, 0)$ , avec  $r < R$ . Il a pour équation:

$$(x^2 + y^2 + z^2 + R^2 - r^2)^2 - 4R^2(x^2 + y^2) = 0$$

et on vérifie que la dérivée de cette équation est non nulle sur les points du tore.

(3) Soit  $M(3, 3, \mathbb{R}) \simeq \mathbb{R}^9$  l'ensemble des  $3 \times 3$  matrices à coefficients réels et  $\Sigma^1 \subset M(3, 3, \mathbb{R})$  le sous-ensemble des matrices de rang 1. Si  $A \in M(3, 3, \mathbb{R})$ , désignons par  $A_{(i_1, i_2), (j_1, j_2)}$ , où  $1 \leq i_1 < i_2 \leq 3$  et  $1 \leq j_1 < j_2 \leq 3$ , le  $2 \times 2$  mineur correspondant à ces suites :

$$A_{(i_1, i_2), (j_1, j_2)} = \det \begin{pmatrix} a_{i_1, j_1} & a_{i_1, j_2} \\ a_{i_2, j_1} & a_{i_2, j_2} \end{pmatrix} = a_{i_1, j_1} a_{i_2, j_2} - a_{i_1, j_2} a_{i_2, j_1} \quad .$$

On a:

$$\Sigma^1 = \{A \in M(3, 3, \mathbb{R}) \mid A_{(i_1, i_2), (j_1, j_2)} = 0, 1 \leq i_1 < i_2 \leq 3, 1 \leq j_1 < j_2 \leq 3\}$$

ce qui nous donne une description de  $\Sigma^1$  à l'aide de 9 équations. Si  $A^0 = (a_{i,j}^0)_{i,j=1,\dots,3} \in \Sigma^1$ , l'un de ses coefficients sera non nul; supposons que ce soit  $a_{1,1}$  et posons  $U_{A^0} = \{A \mid a_{1,1} \neq 0\}$ ; c'est un ouvert contenant  $A^0$  et

$$\Sigma^1 \cap U_{A^0} = \{A \mid A_{(1, i_2), (1, j_2)} = 0, i_2, j_2 = 2, 3\}$$

car l'annulation de ces 4 mineurs entraîne que la 2-ème et 3-ème colonne sont multiples de la première, et donc  $A$  est de rang 1.

La dérivée de  $A_{(1, i_2), (1, j_2)} = a_{1,1} a_{i_2, j_2} - a_{1, j_2} a_{i_2, 1}$  par rapport à  $a_{i_2, j_2}$  vaut  $a_{1,1} \neq 0$  et donc la dérivée de l'application  $(A_{(1, i_2), (1, j_2)})_{i_2, j_2=2,3} : U_{A^0} \rightarrow \mathbb{R}^4$  est de la forme :

$$\left( \begin{array}{cccc|c} a_{1,1} & 0 & 0 & 0 & \\ 0 & a_{1,1} & 0 & 0 & \\ 0 & 0 & a_{1,1} & 0 & \\ 0 & 0 & 0 & a_{1,1} & \end{array} \star \right)$$

où les 4 premières colonnes correspondent aux dérivées par rapport aux variables  $a_{2,2}, a_{2,3}, a_{3,2}, a_{3,3}$ . Puisque  $a_{1,1}$  est non nul, on voit que cette dérivée est surjective. Les  $A_{(1, i_2), (1, j_2)}$ ,  $i_2, j_2 = 2, 3$  forment donc un système d'équation locales de  $\Sigma^1$  au voisinage de  $A^0$ . Donc  $\Sigma^1$  est une sous-variété de  $M(n, n, \mathbb{R})$  de codimension 4, de dimension 5. Au départ, on a décrit  $\Sigma^1$  avec 9 équations, mais on a vu que localement 4 équations suffisent; on peut montrer qu'il n'est pas possible de décrire  $\Sigma^1$  à l'aide de 4 équations globales (i.e. définies sur un ouvert de  $M(3, 3, \mathbb{R})$  contenant  $\Sigma^1$ ) avec dérivée surjective en tout point).

(4) Le sous-ensemble  $X$  de  $\mathbb{R}^2$  constitué par la réunion des 2 axes de coordonnées admet comme équation:

$$X = \{(x, y) \in \mathbb{R}^2 \mid x \cdot y = 0\} \quad .$$

Le gradient de  $x \cdot y$  est  $(y, x)$ , et il s'annule en  $(0, 0)$ . Cela ne prouve pas encore que  $X$  n'est pas une sous-variété. Mais si  $f : U \rightarrow \mathbb{R}$ ,  $U$  ouvert contenant  $(0, 0)$ , s'annule sur  $U \cap X$ ,  $f|_{Ox \cap U} = f|_{Oy \cap U} = 0$  et donc  $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$ . Il n'est donc pas possible de décrire  $X$  au voisinage de  $(0, 0)$  par une équation dont la dérivée est surjective (c'est-à-dire non nulle dans ce cas). Ce sous-ensemble de  $\mathbb{R}^2$  n'est donc pas une sous-variété au voisinage de  $(0, 0)$ .

(5) Considérons la courbe

$$X = \{(x, y) \in \mathbb{R}^2 \mid y^2 - x^3 = 0\}$$

appelée parabole semi-cubique.



La dérivée de  $y^2 - x^3$  est nulle en  $(0, 0)$ , mais (comme dans l'exemple (4) ci-dessus) cela ne suffit pas à montrer que  $X$  n'est pas une sous-variété, même si intuitivement on voit très bien que  $X$  ne ressemble pas à une sous-variété en  $(0, 0)$ . Il faut encore se convaincre que pour toute fonction  $f : U \rightarrow \mathbb{R}$  s'annulant sur  $U \cap X$ ,  $U$  ouvert contenant  $(0, 0)$ , on a que  $df_0 = 0$ . Or  $X$  admet une paramétrisation (globale):

$$X = \{(t^2, t^3) \in \mathbb{R}^2 \mid t \in \mathbb{R}\}$$

et donc  $f(t^2, t^3) = 0$  pour tout  $t$  assez petit. On en déduit en dérivant que  $\frac{\partial f}{\partial x}(t^2, t^3) \cdot 2t + \frac{\partial f}{\partial y}(t^2, t^3) \cdot 3t^2 = 0$ , d'où en divisant par  $t$ :  $\frac{\partial f}{\partial x}(t^2, t^3) \cdot 2 + \frac{\partial f}{\partial y}(t^2, t^3) \cdot 3t = 0$ , et en évaluant en  $t = 0$  on trouve que  $\frac{\partial f}{\partial x}(0, 0) = 0$ . D'autre part,  $0 = f(t^2, t^3) - f(t^2, -t^3) = \frac{\partial f}{\partial y}(t^2, -t^3 + \theta(t)2t^3) 2t^3$ , avec  $0 < \theta(t) < 1$ , d'où  $\frac{\partial f}{\partial y}(t^2, \theta(t)2t^3) = 0$ , et en faisant tendre  $t$  vers 0 on en tire que  $\frac{\partial f}{\partial y}(0, 0) = 0$ .

(6) Soient  $R, r > 0$  et considérons les équations  $f_1(x, y, z) = x^2 + y^2 - R^2 = 0$ ,  $f_2(x, y, z) = x^2 + z^2 - r^2 = 0$ , et posons  $f = (f_1, f_2)$ . Alors :

$$df = \begin{pmatrix} 2x & 2y & 0 \\ 2x & 0 & 2z \end{pmatrix}$$

et les  $2 \times 2$ -mineurs de  $df$  sont  $xy$ ,  $xz$  et  $yz$ . Pour qu'ils soient tous les trois nuls, il faut que  $x = y = 0$ , ou bien  $x = z = 0$ , ou bien  $y = z = 0$ . Si de plus  $f_1(x, y, z) = f_2(x, y, z) = 0$ , alors la seule possibilité est que  $y = z = 0$ , ce qui implique que  $x^2 = r^2$  et  $x^2 = R^2$ . Donc  $f$  définit une sous-variété, à condition que  $r \neq R$ ; c'est une courbe lisse, intersection des 2 cylindres  $f_1 = 0$  et  $f_2 = 0$ .

**4.8 Proposition.** Soit  $X \subset \mathbb{R}^n$ . On a équivalence entre les conditions suivantes:

- (1)  $\forall x_0 \in X$ ,  $\exists$  un ouvert  $U_{x_0} \ni x_0$  de  $\mathbb{R}^n$  et  $f : U_{x_0} \rightarrow \mathbb{R}^{n-k}$ , telle que  $X \cap U = f^{-1}(0)$  et  $df_x$  est surjective,  $\forall x \in U_{x_0}$   
(i.e.  $X$  est une sous-variété de dimension  $k$  de  $\mathbb{R}^n$  - on appelle  $f$  une équation locale régulière de  $X$ , ou des équations locales régulières si l'on se réfère aux composantes de  $f$ ).
- (2)  $\forall x_0 \in X$ ,  $\exists$  un ouvert  $V \subset \mathbb{R}^k$ , un ouvert  $U_{x_0} \ni x_0$  et  $h : V \rightarrow U_{x_0}$  telle que
  - $h : V \xrightarrow{\cong} U_{x_0} \cap X$  est une bijection.
  - $dh_t : \mathbb{R}^k \rightarrow \mathbb{R}^n$  est injective  $\forall t \in V$   
(on dira que  $h$  est une paramétrisation locale régulière de  $X$ ).
- (3)  $\forall x_0 \in X$ ,  $\exists$  un ouvert  $U_{x_0} \ni x_0$  et un difféomorphisme  $H : U_{x_0} \rightarrow \Omega$  sur l'ouvert  $\Omega$  de  $\mathbb{R}^n$  tel que

$$H(X \cap U_{x_0}) = \Omega \cap \mathbb{R}^k \times \{0\}$$

( $H$  est un "aplatissement local" de  $X$ )

(voir figure II.16.)

En résumé, cette proposition nous donne 3 manières équivalentes de décrire localement une sous-variété  $X$  de dimension  $k$  de  $\mathbb{R}^n$  : par des équations locale régulières, par une paramétrisation locale régulière, ou par un difféomorphisme qui identifie le couple (ouvert de  $X$ ,  $\mathbb{R}^n$ ) avec le couple (ouvert de  $\mathbb{R}^k$ ,  $\mathbb{R}^n$ ); régulier signifie que la dérivée est de rang maximum.

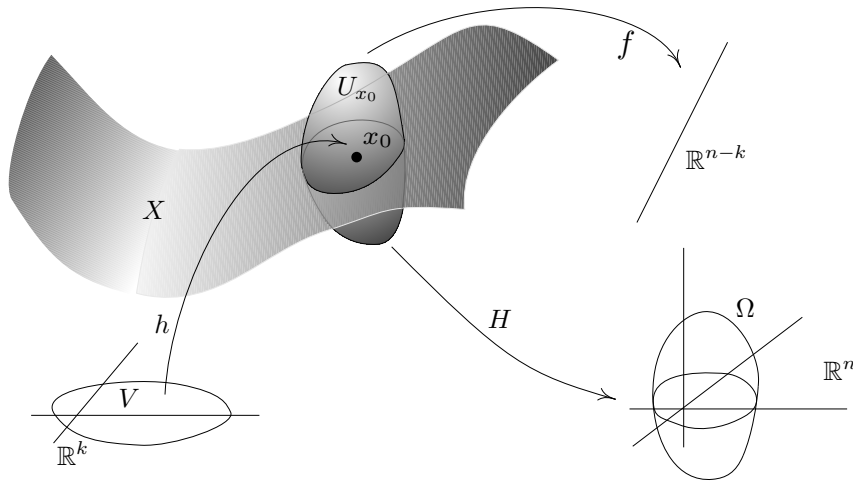
*Preuve:* (1) $\Rightarrow$ (2): par le théorème des fonctions implicites 2.2.

(2) $\Rightarrow$ (3): par le théorème du rang 4.4. Celui-ci nous fournit en effet un difféomorphisme local  $\widehat{H}$  tel que  $\widehat{H}^{-1} \circ h(x_1, \dots, x_k) = (x_1, \dots, x_k, 0, \dots, 0)$ ; il suffit de poser  $H = \widehat{H}^{-1}$ .

(3) $\Rightarrow$ (1): on pose  $f_i(x) = H_{k+i}(x)$ ,  $i = 1, \dots, n - k$ .

*q.e.d.*

Dans la pratique, on rencontre des objets  $X \subset \mathbb{R}^n$  définis par des équations ou des paramétrisations locales, qui satisfont les conditions de la définition de sous-variété en la plupart des points, mais pas partout.



**Figure II.16** – Diverses façons de donner une description locale d'une sous-variété

On appelle points réguliers ou lisses ceux qui admettent des équations locales régulières sur un voisinage, points singuliers les autres; on dit que  $X$  est lisse si tous ses points sont réguliers.

Prenons par exemple le cône  $C$ , décrit par l'équation  $f(x, y, z) = x^2 + y^2 - z^2 = 0$ . Puisque  $df_{(x,y,z)} = (2x, 2y, -2z)$ ,  $df$  est partout de rang maximum (ce qui équivaut à dire non nulle ici) sauf en  $(0, 0, 0)$ . D'autre part, si  $U$  est un ouvert contenant  $(0, 0, 0)$  et  $g : U \rightarrow \mathbb{R}$ ,  $g|_{U \cap C} \equiv 0$ , alors  $g$  s'annule sur toute les génératrices du cône. Alors  $dg_{(0,0,0)}$  s'annule aussi sur ces génératrices, et puisqu'elles engendrent tout l'espace  $\mathbb{R}^3$ ,  $dg_{(0,0,0)} = 0$ . En conclusion, le cône  $C$  est une sous-variété de  $\mathbb{R}^3$ , sauf en  $(0, 0, 0)$ , qui est un point singulier.

**4.9 Définition.** Soit  $X \subset \mathbb{R}^n$  une sous-variété de dimension  $k$ ,  $x_0 \in X$ , et soit  $f : U_{x_0} \rightarrow \mathbb{R}^{n-k}$  une équation locale de  $X$ . L'espace tangent à  $X$  en  $x_0$  est défini par:

$$T_{x_0}X = \{v \in \mathbb{R}^n \mid df_{x_0}(v) = 0\} \quad .$$

Puisque  $df_{x_0}$  est surjective,  $T_{x_0}X$  est de dimension  $k$ .

A priori cette définition dépend de l'équation locale  $f$ . Mais si  $h : V \rightarrow X$ ,  $V \subset \mathbb{R}^k$ ,  $h(t_0) = x_0$  est une paramétrisation locale de  $X$ , on a que  $g \circ h(x) = 0$ , pour tout  $x$  dans un ouvert  $V' \subset V$ ,  $V' \ni t_0$  quelque soit l'équation locale  $g$  et donc  $\text{Im}(dh_{t_0}) \subset \text{Ker}(df'_{x_0})$ ; mais puisque ces 2 espaces ont même dimension, ils coïncident, quelque soit l'équation locale.

Puisque  $f(x) = \underbrace{f(x_0)}_{=0} + df_{x_0}(x - x_0) + r(x - x_0)$ , avec  $\lim_{x \rightarrow x_0} r(x - x_0)/\|x - x_0\| = 0$ , on voit que, près de  $x_0$ , la variété  $X$  est approchée par  $\{x \in \mathbb{R}^n \mid df_{x_0}(x - x_0) = 0\}$ , qui est le translaté de  $T_{x_0}X$  par le vecteur  $x_0$ .

## 5. Singularités d'applications, contours apparents, enveloppes

**5.1 Définition – points singuliers d'une application.** Soit  $U \subset \mathbb{R}^n$  un ouvert et  $f : U \rightarrow \mathbb{R}^n$  une application de classe  $\mathcal{C}^\infty$ ,  $k \geq 1$ . L'ensemble des points singuliers de  $f$  est défini par :

$$\Sigma(f) = \left\{ x \in \mathbb{R}^n \mid df_x \text{ est de rang } < n \right\} = \left\{ x \in \mathbb{R}^n \mid \det \left( \frac{\partial f_i}{\partial x_j}(x) \right)_{i,j=1,\dots,n} = 0 \right\} ;$$

les points de  $f(\Sigma(f)) \subset \mathbb{R}^p$  sont appelés valeurs singulières de  $f$ .

Par exemple, considérons l'application  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $f(x, y) = (x^2, y)$ . On a :

$$df_{(x,y)} = \begin{pmatrix} 2x & 0 \\ 0 & 1 \end{pmatrix} , \quad \Sigma(f) = \{(0, y) \mid y \in \mathbb{R}\} , \quad f(\Sigma(f)) = \{(0, y) \mid y \in \mathbb{R}\} .$$

Cette application est appelée "le pli".

Considérons maintenant l'application  $f(x, y) = (x, y^3 - xy)$ . On a :

$$df_{(x,y)} = \begin{pmatrix} 1 & 0 \\ -y & 3y^2 - x \end{pmatrix} , \quad \Sigma(f) = \{(x, y) \in \mathbb{R}^2 \mid 3y^2 - x = 0\} , \quad f(\Sigma(f)) = \{(3y^2, -2y^3) \in \mathbb{R}^2 \mid y \in \mathbb{R}\} .$$

Cette application est appelée "la fronce" (voir figure II.17).

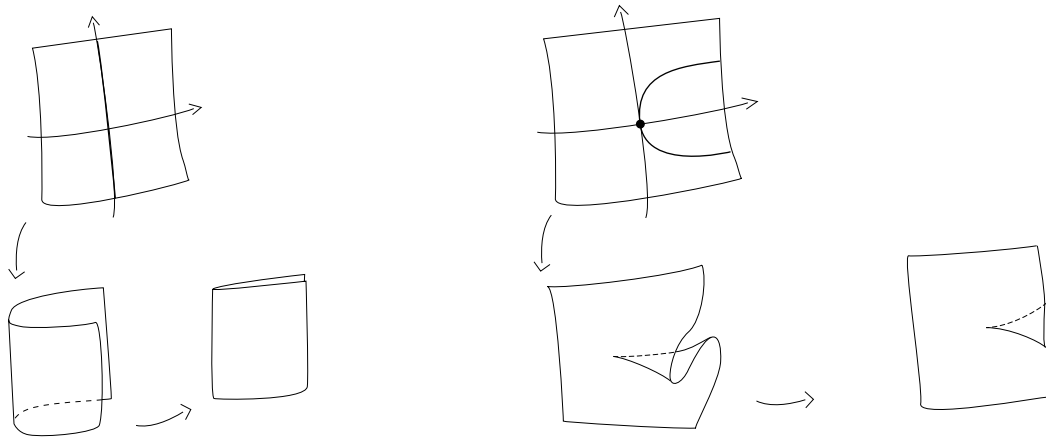


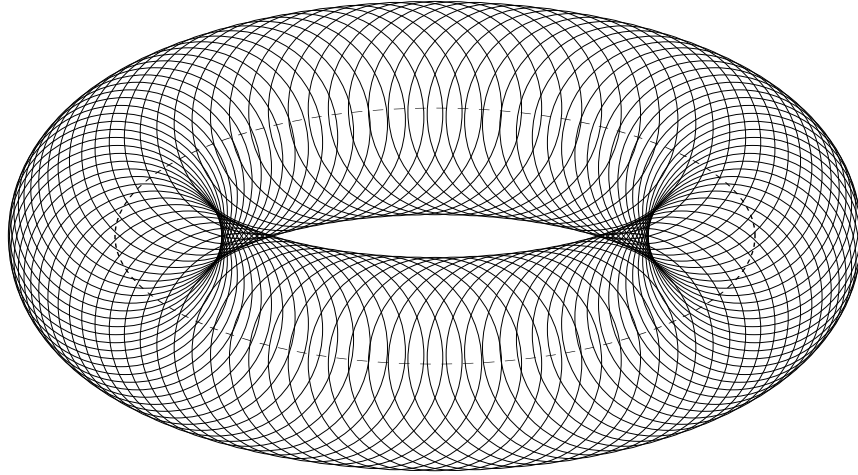
Figure II.17 – Le pli et la fronce

Ces deux applications jouent le rôle de prototype local pour toute application stable de  $\mathbb{R}^2$  dans  $\mathbb{R}^2$ , comme nous le verrons plus loin.

On voit sur ces 2 exemples comment le lieu singulier et son image aident à comprendre l'allure d'une application. En particulier, le nombre de points dans l'image inverse d'un point  $y \in \mathbb{R}^2$  est déterminé par la position de  $y$  par rapport à  $f(\Sigma(f))$ .

Soit  $S \subset \mathbb{R}^3$  une surface,  $\pi \subset \mathbb{R}^3$  un 2-plan et  $p : S \rightarrow \pi$  la restriction à  $S$  de la projection orthogonale de  $\mathbb{R}^3$  sur  $\pi$ . On peut définir les points singuliers de  $p$  de manière analogue au cas d'une application  $U \rightarrow \mathbb{R}^2$ ,  $U \subset \mathbb{R}^2$  un ouvert.

**5.2 Définition – points singuliers, contour apparent.** Soit  $S \subset \mathbb{R}^3$  une surface lisse,  $x_0 \in S$  et  $h : D \rightarrow S$  une paramétrisation locale régulière de  $S$  en  $x_0$ ,  $h(u_0) = x_0$ . On dit que  $x_0$  est un point singulier de la projection orthogonale  $p : S \rightarrow \pi$  sur le plan  $\pi \subset \mathbb{R}^3$  si  $u_0$  est un point singulier de  $p \circ h$ , c'est-à-dire si le rang de la dérivée  $d(p \circ h)_{u_0}$  est plus petit ou égal à 1. On vérifie que cela ne dépend pas de la paramétrisation locale régulière choisie.



**Figure II.18** – Contour apparent de la projection sur le plan  $x + 0.4z = 0$  du tore, représenté comme enveloppe de sphères de rayon  $r$  centrées sur un cercle de rayon  $R$ .

Puisque l'image de la dérivée de  $h$  en  $u_0$  est l'espace tangent à  $S$  au point  $x_0$ , dire que  $x_0$  est singulier c'est dire que  $TS_{x_0}$  est orthogonal à  $\pi$ .

On note par  $\Sigma(p)$  l'ensemble des points singuliers de  $p$  et on appelle  $p(\Sigma(p))$  le contour apparent de  $p$ .

Lorsqu'on dessine la projection d'une surface sur un plan  $\pi$ ,  $p(\Sigma(p))$  contient la frontière de l'image de la projection, d'où le nom "contour apparent".

Lorsque la surface  $S$  est donnée par une équation  $f(x, y, z) = 0$ , les points singuliers de la projection de  $S$  sur le plan  $OXY$  sont décrits par les équations :

$$f(x, y, z) = 0 \quad , \quad \frac{\partial f}{\partial z}(x, y, z) = 0 \quad .$$

Si on arrive à éliminer  $z$  de ce système d'équations, on a l'équation du contour apparent de la projection de  $S$  sur le plan  $OXY$ . Dans le cas de la figure 29, le contour apparent est une courbe de degré 12, donc difficilement traitable.

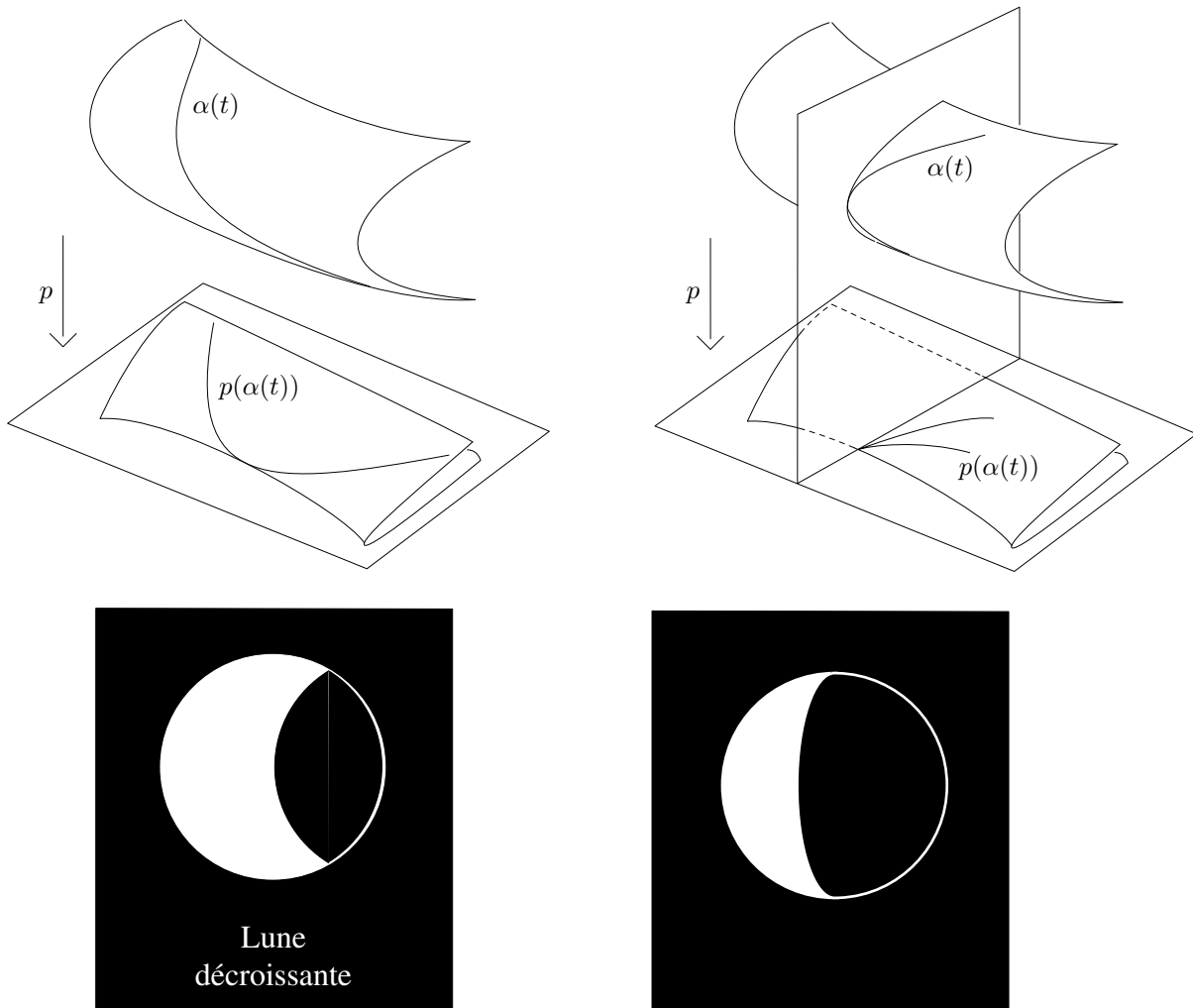
La proposition suivante donne quelques informations utiles sur l'allure du contour apparent. D'après le théorème qu'on énoncera par la suite, les hypothèses en sont presque toujours satisfaites.

**5.3 Proposition.** *Soit  $S \subset \mathbb{R}^3$  une surface lisse,  $P = (x_0, y_0, z_0) \in S$ ,  $f : U \rightarrow \mathbb{R}$  une équation locale de  $S$ ,  $U \ni P$  et soit  $p : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  la projection  $p(x, y, z) = (x, y)$ . Supposons que*

$$\frac{\partial f}{\partial z}(P) = 0 \quad , \quad \frac{\partial^2 f}{\partial z^2}(P) \neq 0 \quad .$$

*Alors  $(x_0, y_0)$  est un point régulier du contour apparent  $p(\Sigma(p) \cap U)$  de  $p$ . De plus, si  $\alpha : I \rightarrow S$  est une courbe tracée sur la surface  $S$ , avec  $\alpha(t_0) = P$ , la projection  $p(\alpha(t)) = (\alpha_1(t), \alpha_2(t))$  est tangente au contour apparent en  $(x_0, y_0)$  et  $p(\alpha(t))$  est située d'un même côté du contour apparent, pour  $t$  proche de  $t_0$  (voir figure II.19).*

*Preuve:* Puisque  $\frac{\partial^2 f}{\partial z^2}(P) \neq 0$ , on peut résoudre explicitement par rapport à  $z$  au voisinage de  $P$  dans l'équation  $\frac{\partial f}{\partial z}(P) = 0$  : soit  $z = g(x, y)$ ,  $g(x_0, y_0) = z_0$  la solution explicite. Alors, pour un ouvert  $U$  assez



**Figure II.19** – Contour apparent et projection d’une courbe tracée sur une surface. Le dessin de la lune à gauche est erroné, celui de droite est correct.

petit contenant  $P$ ,  $p(\Sigma(p) \cap U)$  a pour équation  $f(x, y, g(x, y)) = 0$  au voisinage de  $(x_0, y_0)$ ; cette équation est régulière au point  $(x_0, y_0)$  car :

$$\begin{aligned} \frac{\partial f(x, y, g(x, y))}{\partial x}(P) &= \frac{\partial f}{\partial x}(P) + \underbrace{\frac{\partial f}{\partial z}(P)}_{=0} \frac{\partial g}{\partial x}(x_0, y_0) \\ \frac{\partial f(x, y, g(x, y))}{\partial y}(P) &= \frac{\partial f}{\partial y}(P) + \underbrace{\frac{\partial f}{\partial z}(P)}_{=0} \frac{\partial g}{\partial y}(x_0, y_0) \end{aligned}$$

et  $df_P = (\frac{\partial f}{\partial x}(P), \frac{\partial f}{\partial y}(P), 0)$  est non nul, puisque  $P$  est régulier sur  $S$ .

Posons  $\varphi(t) = f(\alpha_1(t), \alpha_2(t), g(\alpha_1(t), \alpha_2(t)))$ ; il s’agit de montrer que  $\varphi(t)$  ne change pas de signe pour  $t$  proche de  $t_0$ . Remarquons que puisque  $\alpha$  est tracée sur  $S$ , on a que  $f(\alpha_1(t), \alpha_2(t), \alpha_3(t)) \equiv 0$ . En appliquant la formule de Taylor par rapport à la variable  $z$ , on a :

$$f(x, y, z') = f(x, y, z) + \frac{\partial f}{\partial z}(z)(z' - z) + \frac{1}{2} \frac{\partial^2 f}{\partial z^2}(z)(z' - z)^2 + r_2$$

et en prenant  $(x, y, z') = (\alpha_1(t), \alpha_2(t), \alpha_3(t))$ ,  $z = g(\alpha_1(t), \alpha_2(t))$  et en posant  $P(t) = (\alpha_1(t), \alpha_2(t), g(\alpha_1(t), \alpha_2(t)))$ , on obtient :

$$\underbrace{f(\alpha_1(t), \alpha_2(t), \alpha_3(t))}_{=0} = \varphi(t) + \underbrace{\frac{\partial f}{\partial z}(P(t))(\dots)}_{=0} + \frac{\partial^2 f}{\partial z^2}(P(t)) \left( \alpha_3(t) - g(\alpha_1(t), \alpha_2(t)) \right)^2 + r_2$$

et il s'en suit que  $\varphi(t)$  à le même signe que  $-\frac{\partial^2 f}{\partial z^2}(P(t_0))$  pour  $t$  proche de  $t_0$

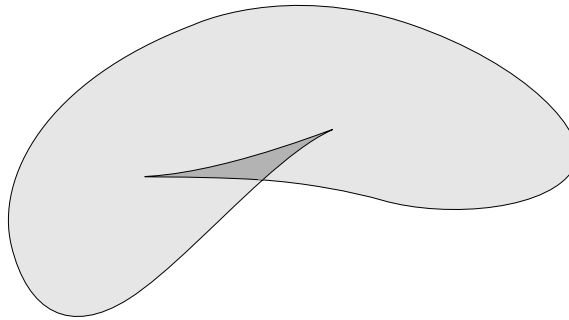
*q.e.d.*

Cette proposition montre que l'image de la lune décroissante de la figure **II.19** est erronée. En effet, la limite de la zone d'ombre est la projection d'une courbe (pratiquement, un grand cercle) tracée sur la lune, donc elle devrait être tangente au contour apparent de la lune. Notons que si  $\alpha_1'(t_0) = \alpha_2'(t_0) = 0$ , le vecteur tangent à  $p(\alpha(t))$  en  $t_0$  est nul, donc évidemment tangent à  $p(\Sigma(p))$ . Néanmoins, visuellement on a l'impression que  $p(\alpha(t))$  s'approche du contour apparent par une direction qui ne lui est pas tangente, avant de rebrousser chemin (voir figure **II.19**).

Voici un théorème fondamental qui donne une information qualitative sur l'allure possible du contour apparent. C'est un cas particulier d'un résultat de H. Whitney, qui remonte à 1940, dont la preuve est assez élaborée.

**5.4 Théorème.** *Pour presque tous les plans  $\pi \subset \mathbb{R}^3$ , le contour apparent de la projection orthogonale de la surface lisse  $S \subset \mathbb{R}^3$  est une courbe ayant comme seules singularités possibles des points doubles ordinaires et des cusps ordinaires. De plus, ces singularités subsistent si la projection subit de petites perturbations. ■*

L'expression "pour presque tous les plans" signifie que l'affirmation du théorème est vraie quitte à remplacer le plan  $\pi$  par un plan  $\pi'$  proche de  $\pi$ . Le fait que les singularités subsistent malgré des petites perturbations de la projection s'exprime en disant qu'elles apparaissent de manière stable.

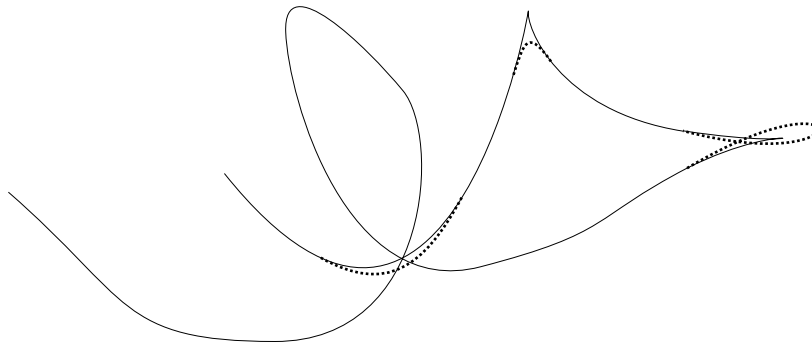


**Figure II.20** – Contour apparent de la projection d'une surface sur un plan

**5.5 Remarque.** En fait Whitney a montré plus précisément que localement, dans des coordonnées locales convenables, la projection se met sous la forme  $(x, y) \mapsto (x^2, y)$  (le pli), ou bien  $(x, y) \mapsto (x, y^3 - xy)$  (la fronce).

---

On a un résultat analogue pour les projections orthogonales d'une courbe lisse  $X \subset \mathbb{R}^3$  sur un plan  $\pi$  : pour presque tous les plans, l'image de  $X$  sera une courbe plane ayant au pire des points doubles ordinaires comme singularités. On peut expérimenter cette affirmation en regardant un fil de fer dans l'espace, ce qui revient à le projeter sur notre plan de vision : si des points triples ou des points cuspidaux apparaissent, une petite perturbation du fil de fer les remplace par des points doubles ordinaires ou les fait disparaître (figure **II.21**).



**Figure II.21** – Application stable d’une courbe dans le plan : après déformation, les singularités non stables cèdent la place à des singularités stables (en pontillé)

5.1 ENVELOPPES

Soit  $S \subset \mathbb{R}^3$  une surface lisse. On va la regarder comme famille de courbes planes : notons par  $(x, y, \lambda)$  un point de  $\mathbb{R}^3$ ; alors, pour tout  $\lambda \in \mathbb{R}$  fixé, l’ensemble :

$$X_\lambda = \{(x, y) \in \mathbb{R}^2 \mid (x, y, \lambda) \in S\}$$

est l’intersection du plan  $z = \lambda$  avec  $S$ , et il a toutes les chances d’être une courbe plane. On appelle enveloppe  $E$  de cette famille de courbes le contour apparent de la projection de  $S$  sur le plan  $(x, y)$ . Si  $S$  est décrite par une équation régulière  $F(x, y, \lambda) = 0$ , on a :

$$(x, y) \in E \iff \exists \lambda \text{ t.q. } F(x, y, \lambda) = 0 \quad , \quad \frac{\partial F}{\partial \lambda}(x, y, \lambda) = 0 \quad .$$

**5.6 Proposition.** Soit  $F(x, y, \lambda) = 0$  une famille de courbes. Soit  $P_0 = (x_0, y_0, \lambda_0) \in \mathbb{R}^3$  et supposons que :

$$F(P_0) = 0 \quad , \quad \frac{\partial F}{\partial \lambda}(P_0) = 0 \quad , \quad \left( \frac{\partial F}{\partial x}(P_0), \frac{\partial F}{\partial y}(P_0) \right) \neq (0, 0) \quad , \quad \frac{\partial^2 F}{\partial \lambda^2}(P_0) \neq 0$$

alors  $(x_0, y_0)$  est un point régulier de  $X_{\lambda_0}$  et aussi de l’enveloppe  $E$  de la famille. Les tangentes de ces deux courbes au point  $(x_0, y_0)$  coïncident et les points de  $X_{\lambda_0}$  dans un voisinage de  $(x_0, y_0)$  sont d’un même côté de  $E$ .

*Preuve:* Puisque  $\left( \frac{\partial F}{\partial x}(P_0), \frac{\partial F}{\partial y}(P_0) \right) \neq (0, 0)$ ,  $(x_0, y_0)$  est un point régulier de  $X_{\lambda_0}$ . Il suffit ensuite d’appliquer la proposition **5.3**

*q.e.d.*

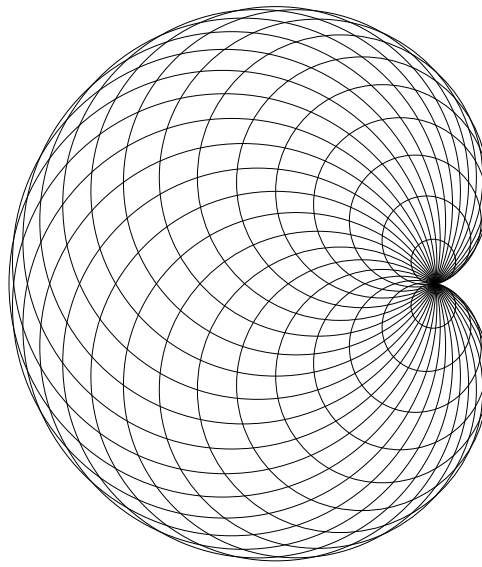
L’affirmation de cette proposition peut se vérifier sur les figures **II.22** et **II.23**.

La proposition **5.6** justifie la définition intuitive de l’enveloppe d’une famille de courbes, qui dit que l’enveloppe est “la courbe tangente à chaque courbe de la famille”. Ainsi exprimée, cette notion peut se généraliser aux familles de surfaces; un exemple d’enveloppe d’une famille de surfaces est représenté sur la figure **II.18**.

**5.7 Exemples.**

(1) Soit  $f(x, y, \lambda) = (x - \lambda)^2 + y^2 - 1 = 0$  la famille des cercles de rayon 1 centrés en  $(\lambda, 0)$ . La système d’équations :

$$\begin{cases} f(x, y, \lambda) = (x - \lambda)^2 + y^2 - 1 = 0 \\ \frac{\partial f}{\partial \lambda}(x, y, \lambda) = -2(x - \lambda) = 0 \end{cases}$$



**Figure II.22** – Enveloppe de la famille de cercles centrés sur un cercle donné, passant par un point fixe du cercle donné

a pour solutions  $x = \lambda$ ,  $y = \pm 1$ . On trouve donc les deux droites horizontales à hauteur  $\pm 1$ .

(2) Considérons la famille des droites qui sont coupées par les axes  $OX$  et  $OY$  selon un intervalle de longueur 1; en prenant comme paramètre l'angle fait par la droite et le côté négatif de  $OX$  cette famille s'écrit :

$$f(x, y, \alpha) = \frac{x}{\cos(\alpha)} + \frac{y}{\sin(\alpha)} - 1 = 0$$

ou encore, en se débarrassant des dénominateurs :

$$g(x, y, \alpha) = x \sin(\alpha) + y \cos(\alpha) - \sin(\alpha) \cos(\alpha) = 0 \quad .$$

Du système d'équations :

$$\begin{cases} g(x, y, \alpha) = x \sin(\alpha) + y \cos(\alpha) - \sin(\alpha) \cos(\alpha) = 0 \\ \frac{\partial g}{\partial \alpha}(x, y, \alpha) = x \cos(\alpha) - y \cos(\alpha) - \cos(\alpha)^2 + \sin(\alpha)^2 = 0 \end{cases}$$

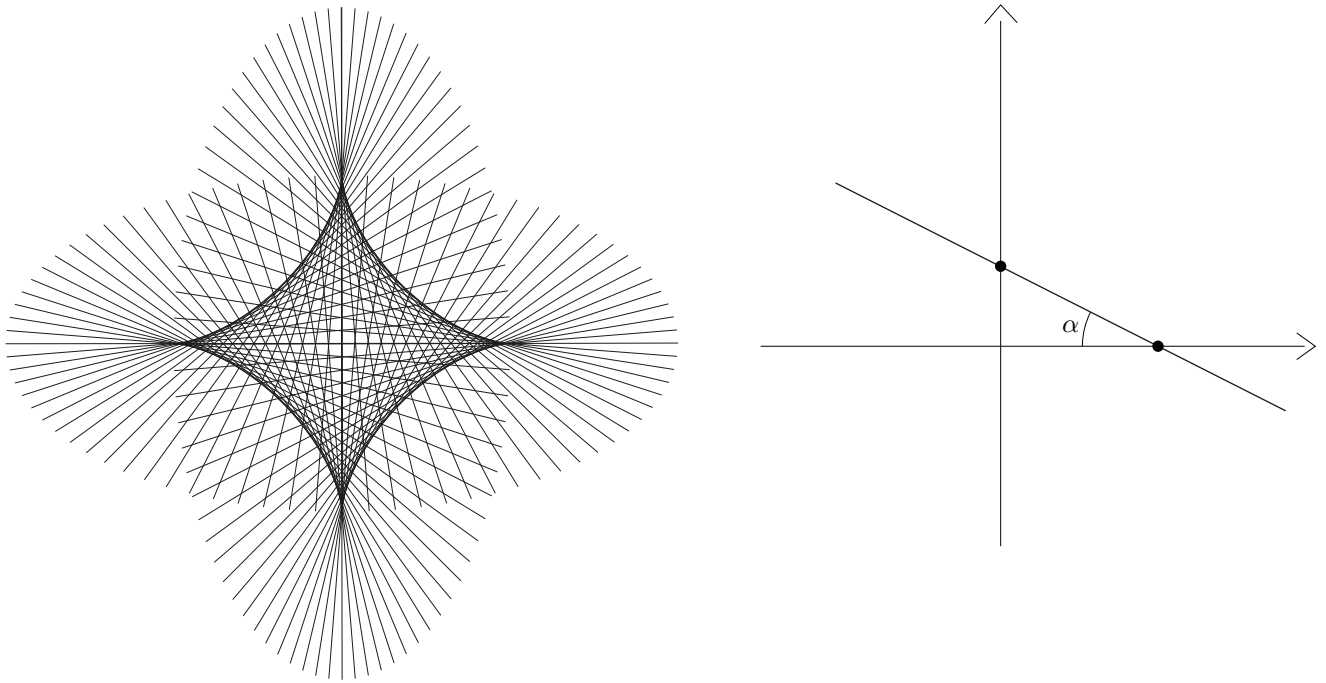
on tire que  $y = \cos(\alpha)^3$ ,  $y = \sin(\alpha)^3$ , d'où l'équation de l'enveloppe :  $x^{2/3} + y^{2/3} = 1$  (voir figures **II.23**, **II.24** et **II.25**)

Regardons maintenant le cas d'une famille de courbes paramétriques planes :

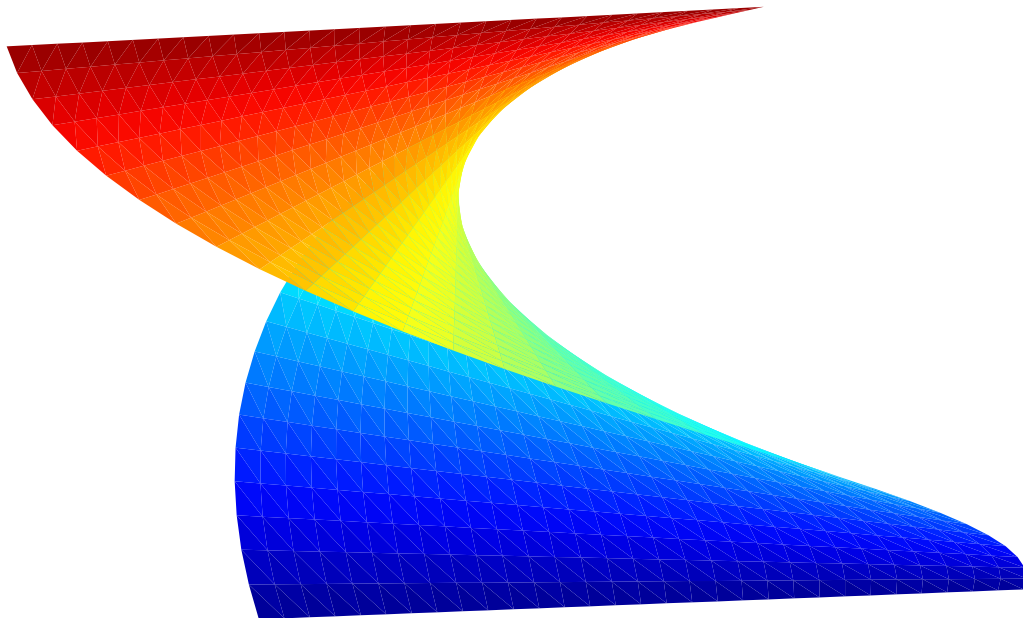
$$\Phi : I \times \Lambda \rightarrow \mathbb{R}^2 \quad , \quad I, \Lambda \subset \mathbb{R} \text{ des intervalles.}$$

Pour tout  $\lambda \in \Lambda$  fixé,  $t \mapsto \Phi_\lambda(t) = \Phi(t, \lambda)$  est une courbe paramétrique du plan. On se ramène au cas précédent en prenant pour  $S$  la surface paramétrique  $(t, \lambda) \mapsto (\Phi(t, \lambda), \lambda)$ . L'intersection de  $S$  avec le plan  $z = \lambda$  est bien la courbe  $\Phi_\lambda$ . Le contour apparent de la projection de  $S$  sur les 2 premières coordonnées coïncide avec le lieu singulier de  $\Phi$ , c'est-à-dire  $\Phi(\Sigma(\Phi))$ .





**Figure II.23** – Enveloppe des droites coupées par les axes  $OX$  et  $OY$  selon un segment de longueur 1

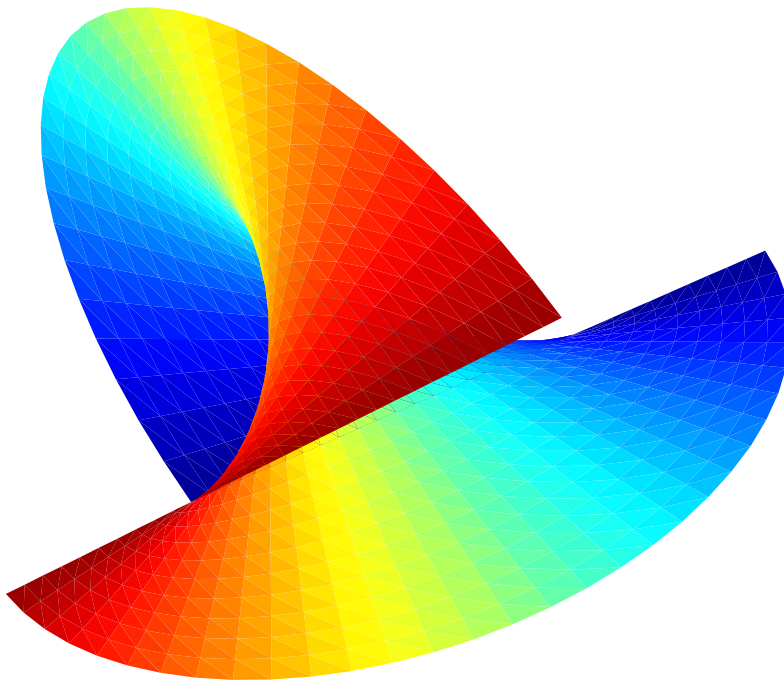


**Figure II.24** – La surface  $S$  associée à la famille de droites précédente, vue de côté

### 5.8 Exemple.

Soit  $\alpha : I \rightarrow \mathbb{R}^2$  une courbe paramétrique régulière, que l'on suppose paramétrée par la longueur d'arc. On appelle droite normale à la courbe en un point  $\alpha(\lambda)$  la droite passant par  $\alpha(\lambda)$  perpendiculaire à la tangente à la courbe en  $\alpha(\lambda)$ ; elle aura pour représentation paramétrique :

$$t \mapsto \Phi(t, \lambda) = \alpha(\lambda) + t\nu(\lambda) = (\alpha_1(\lambda) - t\alpha_2'(\lambda), \alpha_2(\lambda) + t\alpha_1'(\lambda))$$



**Figure II.25** – La même surface d'avant vue d'en haut

où  $\nu = (-\alpha'_2(\lambda), \alpha'_1(\lambda))$ . La matrice jacobienne de  $\Phi$  s'écrit :

$$d\Phi_{(t,\lambda)} = \begin{pmatrix} \alpha'_1(\lambda) + t(-\alpha''_2(\lambda)) & \alpha'_2(\lambda) + t\alpha''_1(\lambda) \\ -\alpha'_2(\lambda) & \alpha'_1(\lambda) \end{pmatrix}$$

et donc

$$\Sigma(\Phi) = \{(t, \lambda) \mid \det(d\Phi_{(t,\lambda)}) = t(-\alpha'_1\alpha''_2 + \alpha''_1\alpha'_2) + \alpha'_1(\lambda)^2 + \alpha'_2(\lambda)^2 = 0\}$$

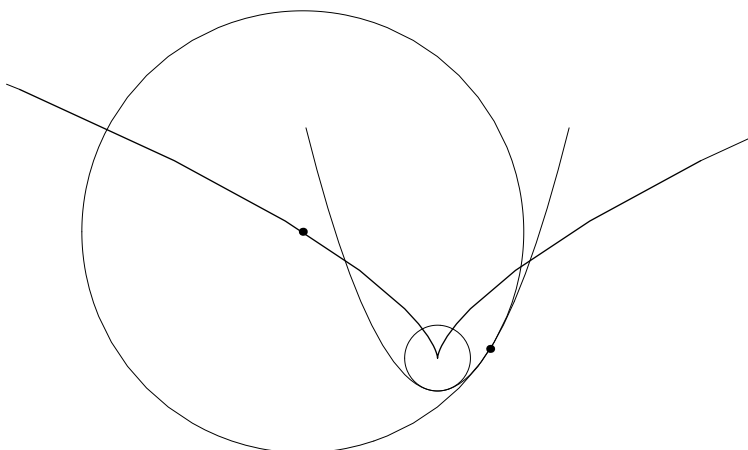
d'où l'on tire que  $(t, \lambda) \in \Sigma(\Phi)$  équivaut à  $t = \frac{1}{\langle \nu(\lambda), \alpha''(\lambda) \rangle}$ , et donc  $\phi(t, \lambda)$  est le centre du cercle osculateur à  $\alpha$  au point  $\alpha(\lambda)$ . L'enveloppe de droites normales est donc le lieu des centres des cercles osculateurs.



**Figure II.26** – Une caustique dans la nature

On déduit du théorème de Whitney qu'en général les enveloppes de familles de courbes ont pour singularités uniquement des cusps ordinaires et des points doubles ordinaires, qui subsistent après une petite

déformation. Cette stabilité explique pourquoi on peut observer dans la nature des courbes présentant des cusps, comme par exemple la caustique constituée par l'enveloppe des rayons de soleil réfléchis dans une tasse de café, alors qu'en général une ficelle posée sur un plan présentera au pire des points doubles à tangentes distinctes. Cela explique aussi pourquoi, en général, le lieu des centres des cercles osculateurs d'une courbe présente des points cuspidaux (voir figure **II.27**).



**Figure II.27** – Le lieu des centres des cercles osculateurs à une parabole