Université Claude Bernard Lyon 1 **Licence Mathématiques et économie**

Analyse pour l'économie 2

Lorenzo Brandolese

1 Espaces métriques complets

1.1 Quelques rappels sur les espaces métriques

- Un *espace métrique* est un couple (X, d) formé par un ensemble et une distance (mais on écrira souvent seulement X pour simplifier). Les éléments de X s'appellent *points*. Par définition, un *distance* (ou *métrique*) sur X est une application $d\colon X \times X \to \mathbb{R}^+$, telle que, pour tout x, y et $z \in X$,

$$\text{i) } d(x,y) = 0 \iff x = y, \qquad \text{ii) } d(x,y) = d(y,x), \qquad \text{iii) } d(x,y) \le d(x,z) + d(z,y).$$

- Pour $r > 0$ et $x \in X$, la boule centrée en x et de rayon $r > 0$ est l'ensemble $B(x,r) = \{y \in X : d(x,y) < r\}$. Dans un espace métrique, un ensemble U est dit *ouvert* si, pour tout $x \in U$, il existe $r > 0$ tel que. $B(x,r) \subset U$. Un ensemble est dit *fermé* si sont complémentaire dans X est u ensemble ouvert.

- Une suite dans un espace métrique X est une application $\mathbb{N} \to X$. Elle est notée généralement $(x_n)_{n\in\mathbb{N}}$ ou simplement (x_n).

  Soit (x_n) une suite d'un espace métrique (X,d) et $x \in X$. On dit la suite (x_n) converge vers x, et on écrit $x_n \to x$ ou encore $\lim_{n\to+\infty} x_n = x$ si, pour tout $\epsilon > 0$ il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \ge n_0$ on a $x_n \in B(x,\epsilon)$ (autrement dit, $d(x_n, x) < \epsilon$). Si la suite ne converge vers aucun point, on dit qu'elle *diverge*. Dans le cas général on voit que $x_n \to x \iff d(x_n, x) \to 0$.

  Le "théorème d'unicité de la limite" affirme que si on a une suite telle que $x_n \to x$ et $x_n \to y$, alors $x = y$.

- L'adhérence $\overline{A}$ d'une partie A d'un espace métrique X est le plus petit fermé contenant A. Un point $x \in \overline{A}$ si et seulement s'il existe une suite $(x_n) \subset A$ telle que $x_n \to x$.

- Une application $f\colon (X, d_X) \to (Y, d_Y)$ est continue en x si et seulement si, pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que : $d_X(x,x') < \delta \Rightarrow d_Y(f(x), f(x')) < \epsilon$. La continuité peut se caractériser par les suites : f est continue en x si et seulement si pour toute suite (x_n) convergente vers x on a $f(x_n)$ convergente vers $f(x)$.

- Une application $f\colon (X, d_X) \to (Y, d_Y)$ entre deux espaces métriques est dite k-lipschitzienne (où $k \ge 0$) si et seulement si, pour tout $x, x' \in X$ on a $d_Y(f(x), f(x')) \le k\, d_X(x,x')$. Elle est dite lipschitzienne s'il existe $k \ge 0$ telle qu'elle est k-lipschitzienne. Les applications lipschitziennes sont continues.

1.2 Suites de Cauchy et espaces complets

Définition 1.1. *Soit (X,d) un espace métrique. Une suite $(x_n) \subset X$ est* **de Cauchy** *si $\forall\, \epsilon > 0$, il existe $n_0 \in \mathbb{N}$ tel que pour tout $n, m \ge n_0$ on a $d(x_n, x_m) < \epsilon$.*

Proposition 1.1. *1. Si (x_n) converge, alors (x_n) est une suite de Cauchy.*

2. Toute suite de Cauchy est bornée.

Dém. 1.) Si $x = \lim_{n\to\infty} x_n$ et $\epsilon > 0$, il existe un n_0 tel que $d(x_n, x) < \epsilon/2$ si $n \geq n_0$. Si $m, n \geq n_0$, on trouve alors $d(x_n, x_m) \leq d(x_n, x) + d(x_m, x) < \epsilon$.

2.) Pour la démonstration de la seconde affirmation, il suffit d'applique la définition de suite de Cauchy avec $\epsilon = 1$. On trouve qu'il existe n_0 tel que, pour tout $m \geq n_0$ on a $d(x_m, x_{n_0}) \leq 1$. Mais alors, pour tout $m \in \mathbb{N}$, $d(x_m, x_{n_0}) \leq R$, où $R = \max_{0\leq i\leq n_0} d(x_i, x_{n_0})\} + 1$. □

Exemple 1.1. Dans l'espace métrique $(\mathbb{Q}, d)$, où $d(x, y) = |x - y|$ est la distance euclidienne, considérons la suite (x_n) définie par $x_n = E(2^n\sqrt{2})/2^n$ (où $E(\alpha)$ désigne la partie entière du nombre réel α, c'est à dire le plus grand entier inférieur ou égale à α). Il s'agit bien d'une suite de nombres rationnels. Cette suite converge vers l'irrationnel $\sqrt{2}$. On conclut que la suite (x_n) est de Cauchy dans $(\mathbb{Q}, d)$, et qu'elle est divergente dans $(\mathbb{Q}, d)$.

Définition 1.2. *Un espace métrique (X, d) est* **complet** *si et seulement si toute suite de Cauchy $(x_n) \subset X$ est convergente dans X.*
Un espace normé $(E, \|\cdot\|)$ est **de Banach** *si et seulement si E est complet pour la distance associée à $\|\cdot\|$.*

Exemple 1.2. $\mathbb{Q}$ muni de la distance usuelle dans $\mathbb{R}$ n'est pas complet, car il existe dans $\mathbb{Q}$ une suite de Cauchy non convergente. L'intervalle $]0, +\infty[$ est un autre exemple d'espace métrique non complet (considérer la suite $(1/n)_{n\in\mathbb{N}^*}$, qui est de Cauchy, pour le voir).

Théorème 1.2. *$\mathbb{R}$ est complet.*

Dém. * Soit (x_n) une suite réelle de Cauchy. Soient $A_n = \{x_n, x_{n+1}, \dots, \}$, $a_n = \inf A_n$, $b_n = \sup A_n$. On a $a_n, b_n \in \mathbb{R}$, car A_n est borné. Clairement, $a_n \leq b_n$, (a_n) est croissante, (b_n) décroissante. Soit $\epsilon > 0$. Il existe un n_0 tel que $|x_n - x_m| < \epsilon/2$ si $n, m \geq n_0$. Pour $n \geq n_0$, on a donc $A_n \subset [x_{n_0} - \epsilon/2, x_{n_0} + \epsilon/2]$, ce qui implique $x_{n_0} - \epsilon/2 \leq a_n \leq b_n \leq x_{n_0} + \epsilon/2$; d'où $b_n - a_n \leq \epsilon$. Il s'ensuit que les suites (a_n), (b_n) sont adjacentes. Par conséquent, il existe un $a \in \mathbb{R}$ tel que $a_n \to a$, $b_n \to a$. Comme $a_n \leq x_n \leq b_n$, on trouve $x_n \to a$. □

Proposition 1.3. *Soit $k \in \mathbb{N}$. Alors $\mathbb{R}^k$, muni de la distance euclidienne, est complet.*

Dém. Soit $(x^n)_{n\in\mathbb{N}}$ une suite de Cauchy de vecteurs de $\mathbb{R}^k$. On a $x^n = (x_1^n, \dots, x_k^n)$. Mais, pour $j = 1, \dots, k$, on a $d(x_j^n, x_j^m) \leq \|x^n - x^m\|_2$. Fixons l'indice j: il s'ensuit que $(x_j^n)_{n\in\mathbb{N}}$ est une suite de Cauchy dans $\mathbb{R}$, qui est complet et qu'alors cette suite converge vers une limite x_j. Mais alors $x_j^n \to x_j$ pour tout $j = 1, \dots, n$. Posons $x = (x_1, \dots, x_k)$. La suite de vecteurs x^n converge composante par composante vers le vecteur $x \in \mathbb{R}^k$. Mais alors $x^n \to x$ dans $\mathbb{R}^k$. □

1.3 Relation entre espaces complets et fermés

Soient (X, d) un espace métrique et $A \subset X$.

Proposition 1.4. *a) Si (A, d) est complet, alors A est un fermé de X.*
b) Si (X, d) est complet et A est un fermé de X, alors (A, d) est complet.

Dém. a) Soit $a \in \overline{A}$. Il existe $(x_n) \subset A$ telle que $x_n \to a$. Alors (x_n) est une suite de Cauchy, donc convergente (dans A, puisque A est complet) vers un $b \in A$. L'unicité de la limite (dans X) implique $a = b \in A$. Il s'ensuit que $\overline{A} \subset A$, d'où A fermé.
b) Soit (x_n) une suite de Cauchy dans A. Alors il existe un $a \in X$ tel que $x_n \to a$. Il s'ensuit que $a \in A$, et donc (x_n) converge dans A. □

Corollaire 1.5. *Dans un espace métrique complet, A complet si et seulement si A fermé.*

*Cette démonstration suppose connu le fait que deux suites réelles adjacentes convergent. Par définition, deux suites (a_n) et (b_n) sont adjacentes lorsque $a_n \leq b_n$, (a_n) croissante, (b_n) décroissante et $b_n - a_n \to 0$.

1.4 L'espace des fonctions continues et bornées

Définition 1.3. *Soit (X, d) un espaces métrique. Une fonction $f : X \to \mathbb{R}$ est dite* bornée *si son image $f(X) = \{f(x) \colon x \in X\}$ est une partie bornée de $\mathbb{R}$. On désigne*

$$C_b(X, \mathbb{R}) = \{f : X \to \mathbb{R} \; ; f \text{ continue et } \text{ bornée}\}.$$

Observons que si K est compact, alors toutes les fonctions continues sont bornées par le théorème de Weierstrass. Donc $C_b(K, \mathbb{R}) = C(K, \mathbb{R})$.

La somme de deux fonctions continues et bornée est une fonction continue et bornée. Et si on multiplie une fonction continue et bornée par un nombre réel on obtient une autre fonction continue et bornée. Donc $C_b(X, \mathbb{R})$ est un espace vectoriel.

Nous définissons la "norme du sup" sur $C_b(X, \mathbb{R})$ par

$$\|f\|_\infty = \sup_{x \in X} |f(x)|.$$

Il n'est pas difficile de voir que $C_b(X, \mathbb{R})$ est un espace vectoriel normé, pour la norme du sup.

On définit une distance δ sur l'ensemble $C_b(X, \mathbb{R})$ (dite "distance du sup"), par

$$\forall\, f, g \in \mathbb{C}_b(X, \mathbb{R}) \colon \qquad \delta(f, g) = \sup_{x \in X} |f(x) - g(x)|.$$

La norme $\| \cdot \|_\infty$ induit bien entendu la distance δ par la relation usuelle $\delta(f, g) = \|f - g\|_\infty$.

Exemple 1.3. Si $X = [0, 1]$, les fonctions $f(x) = \exp(x)$ et $g(x) = \exp(2x)$ appartiennent à $C_b([0, 1], \mathbb{R}) = C([0, 1], \mathbb{R})$. Calculons la distance entre ces deux fonctions f et g : on a

$$\delta(f, g) = \sup_{x \in [0,1]} |e^x - e^{2x}| = \max_{x \in [0,1]} (e^{2x} - e^x) = e^2 - e.$$

Si une suite de fonctions $(f_n) \subset C_b(X, \mathbb{R})$ converge vers $f \in C_b(X, \mathbb{R})$, c'est à dire $f_n \xrightarrow{\delta} f$, on dit que (f_n) converge *uniformément* vers f. Plus explicitement, cela signifie que

$$\sup_{x \in X} |f_n(x) - f(x)| \to 0.$$

Exemple 1.4. Soit $X = [0, 1]$ et $f_n(x) = e^{x/n}$. Pour tout $x \in [0, 1]$, $f_n(x) \to 1$. Soit f la fonction constante égal à 1. On a, pour tout $x \in [0, 1]$, $|f_n(x) - f(x)| = |e^{x/n} - 1| \le e^{1/n} - 1$. Donc $\sup_{x \in [0,1]} |f_n(x) - f(x)| \le e^{1/n} - 1 \to 0$. On conclut que $\delta(f_n, f) \to 0$.

Observons que, parfois, l'on a $\lim_{n \to +\infty} f_n(x) = f(x)$ pour tout $x \in X$ (autrement dit, on a *convergence simple* de (f_n) vers f) sans qu'il y ait convergence uniforme.

Proposition 1.6.

1. *La limite uniforme de fonctions continues et bornées est continue et bornée : autrement dit, si $(f_n) \subset C_b(X, \mathbb{R})$ et $f \colon X \to \mathbb{R}$ est telle que $f_n \xrightarrow{\delta} f$, alors $f \in C_b(X, \mathbb{R})$.*

2. *$C_b(X, \mathbb{R})$ est un espace métrique complet pour la distance du* sup. *Il s'agit donc d'un espace de Banach.*

Dém. (1). Soit $f_n \xrightarrow{\delta} f$ et $x \in X$. Pour démontrer que f est continue en x on considère $\epsilon > 0$. On sait alors qu'il existe n_0 tel que $\delta(f, f_n) < \epsilon/3$ pour tout $n \ge n_0$. On a

$$\begin{aligned} |f(x) - f(x')| &\le |f(x) - f_n(x)| + |f_n(x) - f_n(x')| + |f_n(x') - f(x')| \\ &\le 2\delta(f, f_n) + |f_n(x) - f_n(x')| \\ &\le 2\epsilon/3 + |f_n(x) - f_n(x')|. \end{aligned}$$

En appliquant l'inégalité ci-dessus avec $n = n_0$ et le fait que f_{n_0} est continue en x, on trouve qu'il existe $\eta > 0$ tel que, si $d(x, x') < \eta$, alors $\delta(f, f_n) < \epsilon/3$ et donc $|f(x) - f(x')| < \epsilon$. Ceci assure que $f : X \to \mathbb{R}$ est bien continue.

De plus, en appliquant la définition de convergence avec $\epsilon = 1$, on voit qu'il existe $n_0 \in \mathbb{N}$ tel que

$$\begin{aligned} |f(x)| &\leq |f(x) - f_{n_0}(x)| + |f_{n_0}(x)| \\ &\leq \delta(f, f_{n_0}) + |f_{n_0}(x)| \\ &\leq 1 + |f_{n_0}(x)|. \end{aligned}$$

Si l'on passe au sup,

$$\sup_{x \in X} |f(x)| \leq 1 + \sup_{x \in X} |f_{n_0}(x)| < \infty$$

puisque f_{n_0} est une fonction bornée et donc f est elle même bornée.

(2). Si (f_n) est une suite de Cauchy dans $C_b(X, \mathbb{R})$, alors, pour tout $x \in X$, $(f_n(x))$ est une suite de Cauchy dans $\mathbb{R}$. On pose $f(x) = \lim_{n \to \infty} f_n(x)$. Soit $\epsilon > 0$. Il existe un n_0 tel que, si $n, m \geq n_0$, alors $\delta(f_n, f_m) < \epsilon/2$. Pour tout $x \in X$, on a donc alors $|f_n(x) - f_m(x)| < \epsilon/2$ si $n \geq n_0$. Passons à la limite dans cette inégalité pour $m \to +\infty$. Ceci donne $|f_n(x) - f(x)| \leq \epsilon/2$ pour $n \geq n_0$. Mais alors $\delta(f_n, f) \leq \epsilon/2 < \epsilon$ pour $n \geq n_0$. Il s'ensuit $f_n \xrightarrow{\delta} f$. □

Comme application immédiate de la proposition précédente (avec $X = [a, b]$), nous avons que l'ensemble $C([a, b], \mathbb{R})$ des applications continues sur un intervalle $[a, b]$ et à valeurs réelles est complet pour la distance du sup.

Une autre application intéressante est fournie par la proposition suivante.

Proposition 1.7. *On désigne avec ℓ^∞ l'espace vectoriel de toutes les suites réelles bornées : $\ell^\infty = \{x = (x_n) \subset \mathbb{R}\ ; (x_n)$ bornée$\}$. On munit ℓ^∞ de la norme $\|x\|_\infty = \sup_{n \in \mathbb{N}} |x_n|$. Alors ℓ^∞ est un espace de Banach pour la distance induite de la norme $\| \cdot \|_\infty$.*

Dém. On a $\ell^\infty = C_b(\mathbb{N}, \mathbb{R})$: en effet, on peut voir une suite (bornée) comme une fonction $: \mathbb{N} \to \mathbb{R}$ (bornée) et réciproquement. D'autre part, toute fonction $: \mathbb{N} \to \mathbb{R}$ est continue (il suffit d'appliquer la définition de continuité avec $\epsilon < 1/2$ pour s'en convaincre). Mais alors le résultat de cette proposition est une conséquence immédiate de celle de la proposition précédente avec $X = \mathbb{N}$. □

1.5 Contractions et le théorème du point fixe

Définition 1.4. *Une application $f : (X, d) \to (Y, D)$ est* contractante *s'il existe un $k < 1$ tel que f soit k-lipschitzienne.*

Exemple 1.5. L'application $f : \mathbb{R} \to \mathbb{R}$ telle que $f(x) = \arctan(\frac{x}{2})$ est contractante. En effet, pour tout $x, x' \in \mathbb{R}$ (on pourra supposer $x < x'$) on a, par l'inégalité des accroissements finis, $|f(x) - f(x')| \leq \sup_{\xi \in [x, x']} |f'(\xi)| |x - x'| \leq \sup_{\xi \in [x, x']} \frac{1}{2} \cdot \frac{1}{1 + (x/2)^2} |x - x'| \leq \frac{1}{2} |x - x'|$. Cette fonction est alors $\frac{1}{2}$-lipschitzienne sur $\mathbb{R}$.

Définition 1.5. *Si $f : X \to X$, un* point fixe *de f est une solution de l'équation $f(x) = x$.*

Théorème 1.8 (du point fixe de Picard)**.** *Soient (X, d) un espace métrique* **complet** *et $f : (X, d) \to (X, d)$* **contractante***. Alors :*
a) f possède exactement un point fixe a ;
b) pour tout $x_0 \in X$, la suite (x_n), $x_n = \underbrace{f \circ f \circ \ldots \circ f}_{n \text{ fois}}(x_0)$, converge vers a ;

Dém. a) Soit $0 < k < 1$ tel que f soit k-lipschitzienne. Montrons que f a au plus un point fixe : si, par l'absurde, a et b sont des points fixes et $a \neq b$, on aboutit à la contradiction $0 < d(a,b) = d(f(a), f(b)) \leq kd(a,b) < d(a,b)$.
L'existence de a suit de b) : si la suite (x_n) converge et si a est tel que $x_n \to a$, alors $x_{n+1} = f(x_n) \to f(a)$ (puisque toute fonction lipschitzienne est continue), d'où $f(a) = a$.
b) On a, pour tout n, $d(x_{n+1}, x_n) \leq k^n d(x_1, x_0)$ (par récurrence sur n). Par conséquent, si $m \geq n$, alors

$$(1) \qquad d(x_m, x_n) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \ldots + d(x_{m-1}, x_m) \leq \frac{k^n}{1-k} d(x_1, x_0) = Ck^n.$$

Comme $Ck^n \to 0$, pour tout $\epsilon > 0$ il existe un n_0 tel que $Ck^n < \epsilon$ si $n \geq n_0$. Il s'ensuit que $d(x_m, x_n) < \epsilon$ si $m, n \geq n_0$. La suite (x_n) étant de Cauchy, elle converge vers un $a \in X$. De ce qui précède, a est l'unique point fixe de f.

□

Exemple 1.6. Trouver le nombre des solutions de l'équation $\cos x = x$.

On a $\cos x = x \Rightarrow x \in [-1, 1]$. Soit $f : X = [-1, 1] \to X$, $f(x) = \cos x$. $[-1, 1]$ est complet (avec la distance usuelle), car fermé dans $\mathbb{R}$. Par ailleurs, on a $|f'(x)| \leq \sin 1 < 1$, $x \in X$. Le théorème des accroissements finis implique $|f(x) - f(y)| \leq \sin 1 |x - y|$, $x, y \in X$. Il s'ensuit que l'équation $\cos x = x$ a exactement une solution.

2 Équations différentielles

2.1 Motivations et premiers exemples

Les équations différentielles sont les équations dont l'inconnue est une fonction, qu'il faut déterminer à partir de relations sur ses dérivées. Le cas le plus simple est celui des équations de la forme

$$u'(t) = f(t) \qquad (f \text{ donnée}, \ u \text{ à déterminer}),$$

dont les solutions, sur un intervalle $[a,b]$, sont les primitives de la fonction f, c'est-à-dire les fonctions $u(t) = \int_a^t f(s)\,\mathrm{d}s + c$, où c est une constante arbitraire. Dans toutes les sciences on a affaire à des problèmes qui peuvent se modéliser par des équations différentielles.

Exemple 2.1 (Désintégration du neutron)**.** Le neutron est une particule instable. Soit p la probabilité qu'un neutron se désintègre en une seconde. On souhaite calculer $N(t)$, le nombre de neutrons présents à l'instant t. On s'attend que le nombre de neutrons qui se désintègrent dans l'intervalle de temps $[t, t+h]$ soit $N(t) - N(t+h) = p\,h\,N(t)$. Si on divise par h et on prend $h \to 0$ on trouve

$$N'(t) = -pN(t).$$

Divisons par $N(t)$ (on verra plus loin que pour ce type d'équations ceci est légitime, c'est à dire que $N(t) \neq 0$). On trouve $N'(t)/N(t) = -p$ et, en passant aux primitives, $\ln(N(t)) = -pt + c$. Ou encore,

$$N(t) = Ae^{-pt},$$

où $A = e^c = N(0)$. Si on connait $N(0)$ (le nombre de neutrons à l'instant initial), par cette formule on peut alors prédire le nombre de neutrons présents à aux instants t successifs.

Exemple 2.2 (Un problème géométrique)**.** Trouver l'équation de la courbe dans le quadrant $x > 0$ et $y > 0$, passant par le point $P(3;4)$ et dont la pente de la droite tangente en tout point (x, y) est donnée par y/x.

Solution. Soit $y = y(x)$ la fonction à déterminer. Il s'agit de résoudre l'équation différentielle $y'(x) = y(x)/x$. On réécrit l'équation sous la forme $y'(x)/y(x) = 1/x$. Ensuite on passe aux primitives terme-à-terme : $\ln(y(x)) = \ln(x) + c$, où c est une constante arbitraire. Passons aux exponentiels : $y(x) = e^{\ln x + c} = Ax$, où $A = e^c$. Imposons la condition de passage par le point $P(3;4)$: il faut que $4 = y(3) = 3A$, ce qui permet de calculer A. Donc la courbe cherchée est le graphe de la fonction $y(x) = \frac{4}{3}x$.

Les équations différentielles d'ordre supérieur sont très fréquentes dans les applications : par exemple, en mécanique newtonienne, la position $u(t)$ d'un point materiel qui se déplace le long d'une droite ($u(t)$ exprime alors l'abscisse du point sur cette droite) et soumis à une force f, est donnée par la loi fondamentale $f = ma$, où m est la masse du point et a l'acceleration. La vitesse du point et son accélération à l'instant t sont données par $u'(t)$ et $u''(t)$ respectivement. La force qui agit sur le point est donnée par une fonction connue, dépendente en général du temps, de la position $u(t)$ et de sa vitesse $u'(t)$. La loi de la mécanique $f = ma$ se traduit alors par l'équation différentielle

$$u''(t) = \frac{1}{m} f(t, u(t), u'(t)).$$

Exemple 2.3. On suspend une masse m à un ressort. D'après la loi de Hooke, la force de rappel du ressort qui subit une élongation ℓ par rapport à la position d'équilibre vaut $\kappa\ell$, où κ est une constante que l'on calcule expérimentalement : à l'état de repos, la force de rappel compense la force de gravité mg, donc $mg = \kappa\ell$, qui permet de calculer κ.

On fait osciller le système verticalement : l'équation différentielle associée à ce problème est

$$u''(t) = -\frac{\kappa}{m} u(t).$$

Ici, l'inconnue $u(t)$, représente l'écart à l'équilibre. Le signe négatif est dû au fait que si $y(t)$ est positif (ressort comprimé) la force pointe à l'intant t vers le bas et réciproquement. Nous verrons que la "solution générale" de cette équation différentielle est

$$u(t) = A\cos(\sqrt{\kappa/m}\,t) + B\sin(\sqrt{\kappa/m}\,t), \qquad A, B \in \mathbb{R}.$$

Autrement dit, toutes les fonctions ci-dessus vérifient l'équations différentielle (la vérification est facile) et il n'y en a pas d'autres. Observer que si on connait $u(0) = u_0$ (la position de la masse à l'instant initiale) et $u'(0) = u_1$ (la vitesse de la masse à l'instant initial) on peut déterminer $A = u_0$ et $B = \sqrt{\kappa/m}\,u_1$. La solution du problème posé devient alors unique.

2.2 Le problème de Cauchy

Nous allons étudier notamment les équations différentielles de la forme

$$u'(t) = f(t, u(t)), \qquad \text{(on écrira souvent, en abrégé, } u' = f(t,u))$$

où l'inconnue est la fonction dérivable $u\colon I \to \mathbb{R}$ (où I est un intervalle de $\mathbb{R}$) et la fonction f de deux variables est donnée. Une telle équation différentielle est dite *scalaire* (puisque l'inconnue u est à valeurs dans $\mathbb{R}$), *du premier ordre* (parce qu'elle ne contient que des dérivées d'ordre 1) et *en forme normale* (parce que l'on peut isoler la dérivée à gauche de l'égalité).

Les exemple 2.1 et 2.2 sont bien deux exemples d'équations du type $u' = f(t,u)$: dans le premier cas, la fonction f est donnée par $f(t,u) = -pu$, dans le second cas, on a $f(t,u) = u/t$.

Une équations différentielle possède en général une infinité de solutions. Souvent, on cherche parmi les solutions celle vérifiant des conditions supplémentaires.

Définition 2.1 (Solution locale d'un problème de Cauchy)**.** *Soit f une fonction définie au voisinage de $(t_0, u_0) \in \mathbb{R}^2$. On dit que le* problème de Cauchy

$$\begin{cases} u' = f(t,u) \\ u(t_0) = u_0. \end{cases} \tag{2.1}$$

admet une solution locale *s'il existe une fonction dérivable u sur un voisinage de t_0, telle que $u(t_0) = u_0$ et, pour tout t dans ce voisinage, on a*

$$u'(t) = f(t, u(t)).$$

Il est parfois utile de ramener l'étude d'un problème de Cauchy à une équation intégrale:

Proposition 2.1. *Soit $(t_0, u_0) \in \mathbb{R}^2$ et f une fonction continue dans un voisinage de (t_0, u_0), à valeurs réelles. Une fonction $u(t)$ est une solution de classe C^1 du problème de Cauchy (2.1) dans un voisinage de t_0 si et seulement si, dans ce voisinage, u est une solution continue de l'équation intégrale*

$$u(t) = u_0 + \int_{t_0}^{t} f(s, u(s))\,\mathrm{d}s. \tag{2.2}$$

Dém. À compléter. □

L'existence d'une solution du problème de Cauchy est garantie dès que f est continue. En effet, nous avons le théorème suivant, que nous admettons:

Théorème 2.2 (Peano)**.** *Soit f une fonction continue au voisinage de (t_0, u_0). Alors le problème de Cauchy* (2.1) *possède au moins une solution locale.*

Le théorème principal de ce chapitre est le théorème suivant. Sous une hypothèse un peu plus restrictive sur f que la seule continuité, il est possible de garantire à la fois l'existence et l'unicité d'une solution locale au problème de Cauchy.

Théorème 2.3 (Cauchy-Lipschitz)**.** *Soit $I = [t_0 - r_1, t_0 + r_1]$ un intervalle fermé contenant t_0 et $J = [u_0 - r_2, u_0 + r_2]$ un intervalle fermé contenant u_0 à leur intérieur. On suppose que f vérifie les hypothèses suivantes :*

a) f est continue dans un voisinage de (t_0, u_0), contenant le rectangle $I \times J$.

b) Il existe une constante $L > 0$ telle que, pour tout $s \in I$ et tout $v, w \in J$, on a

$$|f(s, v) - f(s, w)| \leq L|v - w|.$$

Alors le problème de Cauchy (2.1) *possède une unique solution locale.*

Plus précisément, il existe une et une seule solution u, de classe C^1, définie au moins sur l'intervalle $I_0 = [t_0 - r_0, t_0 + r_0]$, où :

$$0 < r_0 < \min\{r_1, r_2/M, 1/L\},$$

et $M = \max_{I \times J} |f|$.

Dém. L'idée de la démonstration est la suivante : u est solution du problème de Cauchy si et seulement si u est un point fixe dans pour une fonction Φ à déterminer. L'existence et l'unicité de la solution découlera du théorème des contractions.

Introduisons l'espace

$$X = \{u \colon I_0 \to \mathbb{R}, : u \text{ continue et } \|u - u_0\|_\infty \leq r_2\}.$$

Observons que X est une partie fermée de $C(I_0, \mathbb{R})$. Cet espace étant complet, X est lui même un espace métrique complet.

Considérons l'application $\Phi \colon X \to X$, définie de la manière suivante : si $v \in X$ (donc $v = v(t)$ est une fonction continue de la variable t), par définition la fonction $\Phi(v) \colon X \to X$ est la fonction qui à t associe

$$\Phi(v)(t) := u_0 + \int_{t_0}^{t} f(s, v(s))\, \mathrm{d}s.$$

Vérifions qu'on a bien $\Phi(v) \in X$. Tout d'abord, $t \mapsto \Phi(v)(t)$ est bien continue. De plus, pour tout $t \in I_0$,

$$\begin{aligned} |\Phi(v)(t) - u_0| &= \Big| \int_{t_0}^{t} f(s, v(s))\, \mathrm{d}s \Big| \leq \int_{t_0}^{t} |f(s, v(s))|\, \mathrm{d}s \\ &\leq M r_0 < r_2. \end{aligned}$$

En passant au sup sur $t \in I_0$ on trouve $\|\Phi(v) - u_0\|_\infty \leq r_2$. Autrement dit, pour tout $v \in X$, on a bien $\Phi(v) \in X$, donc $\Phi \colon X \to X$ comme initialement affirmé.

Démontrons maintenant que $\Phi \colon X \to X$ est une contraction. Soient $v, w \in X$ et estimons $\|\Phi(v) - \Phi(w)\|_\infty = \sup_{t \in I_0} |\Phi(v)(t) - \Phi(w)(t)|$.

On a

$$
\begin{aligned}
|\Phi(v)(t)-\Phi(w)(t)| &= \Big|u_0+\int_{t_0}^t f(s,v(s))\,\mathrm{d}s-u_0-\int_{t_0}^t f(s,v(s))\,\mathrm{d}s\Big| \\
&= \Big|\int_{t_0}^t\Big(f(s,v(s))-f(s,w(s))\Big)\,\mathrm{d}s\Big| \\
&\leq \int_{t_0}^t\Big|f(s,v(s))-f(s,w(s))\Big|\,\mathrm{d}s \\
&\leq L\int_{t_0}^t\Big|v(s)-w(s)\Big|\,\mathrm{d}s \\
&\leq L\int_{t_0}^t\|v-w\|_\infty\,\mathrm{d}s \\
&\leq L\,r_0\|v-w\|_\infty.
\end{aligned}
$$

Si on passe au $\sup_{t\in I_0}$ dans ces inégalités on trouve

$$\|\Phi(v)-\Phi(w)\|_\infty\leq Lr_0\|v-w\|_\infty.$$

Mais nos hypothèses impliquent $Lr_0<1$, ce qui assure que Φ est bien une application contractante dans X. D'après le théorème de point fixe de Picard (théorème 1.8), on en déduit qu'il existe une et une seule fonction $u\in X$ telle que

$$u=\Phi(u).$$

Cette égalité de point fixe signifie que

$$u(t)=\Phi(u)(t)=u_0+\int_{t_0}^t f(s,u(s))\,\mathrm{d}s., \qquad \text{pour tout } t\in I_0.$$

Mais cette équation n'est rien d'autre que l'équation intégrale (2.2), comme on le voit d'après la définition de Φ.

Or, on sait déjà (proposition 2.1) que l'équation intégrale est équivalente au problème de Cauchy. En conclusion, le problème de Cauchy possède une et une seule solution dans l'intervalle I_0. □

La condition *b)* dans le théorème de Cauchy-Lipschitz exprime le fait que la fonction f est lipschitzienne par rapport à la seconde variable (avec une constante de Lipschitz indépendante de la première variable). Dans la plupart de cas, pour vérifier cette condition b), on fait appel au corollaire suivant.

Corollaire 2.4. *Soit f une fonction continue dans un voisinage de (t_0,u_0), telle que $\frac{\partial f}{\partial u}$ est continue au au voisinage de (t_0,u_0). Alors le problème de Cauchy* (2.1) *possède une unique solution locale.*

Dém. En effet, par le théorème de accroissements finis, appliqué à la fonction $u\mapsto f(t,u)$, il existe ξ, compris entre v et w tel que

$$f(t,v)-f(t,w)=\frac{\partial f}{\partial u}(t,\xi)(v-w).$$

Mais comme $\frac{\partial f}{\partial u}$ est continue, nous pouvons poser $L=\max_{I\times J}|\frac{\partial f}{\partial u}|$, et on trouve

$$\forall\,t\in I,\ v,w\in J,\qquad |f(t,v)-f(t,w)|\leq L|v-w|.$$

La condition b) du théorème de Cauchy–Lipschitz est alors satisfaite. □

Exemple 2.4. Pour quelles valeurs de $(t_0, u_0) \in \mathbb{R}^2$ le problème de Cauchy

$$\begin{cases} u' = te^u - \ln u \\ u(t_0) = u_0. \end{cases}$$

possède-t-il une unique solution solution locale ? Réponse : la fonction $f(t,u) = te^u - \ln u$ est définie et continue pour tout $t \in \mathbb{R}$ et pour $u > 0$. De plus, $\frac{\partial f}{\partial u} = te^u - 1/u$ et cette fonction est continue pour $t \in \mathbb{R}$ et $u > 0$. D'après le corollaire, pour tout $t_0 \in \mathbb{R}$ et tout $u_0 > 0$, le problème de Cauchy possède une unique solution locale. D'autre part, pour $u_0 \leq 0$, il n'y a pas de solution locale (parce que le terme à droite dans l'équation différentielle n'a pas de sens pour $t = t_0$).

Exemple 2.5. Considerons le problème de Cauchy

$$\begin{cases} u' = 3u^{2/3} \\ u(0) = 0. \end{cases}$$

Ce problème possède deux solutions locales distinctes : $u(t) = 0$ et $u(t) = t^3$. Observer que dans ce cas la fonction $f(t,u) = 3u^{2/3}$ est bien continue au voisinage de $(0,0)$, donc le théorème d'existence de solutions de Peano s'applique. Mais cette fonction ne vérifie pas la condition (b) du théorème de Cauchy–Lipschitz. Noter que $\frac{\partial f}{\partial u}(t,u) = 2u^{-1/3}$ qui est discontinue en $u = 0$: le corollaire ne s'applique pas non plus.

2.3 Prolongement des solutions

Dans toute cette section nous supposons que f vérifie en tout point des conditions comme dans le corollaire précédent. Plus précisément, nous supposons que* :

a') $f = f(t,u)$ est continue dans un ouvert U de $\mathbb{R}^2$.

b') $\frac{\partial f}{\partial u}$ est continue dans U.

Dans ce cas, les graphes de deux solutions d'une équation différentielle ne peuvent pas s'intersecter. En effet, nous avons le théorème suivant :

Proposition 2.5. *Supposons que f vérifie les hypothèses (a') et (b'). Si v et w sont deux solutions de l'equation différentielle $u' = f(t,u)$ définies sur un même intervalle $]a,b[$, qui coïncident en un point $t_0 \in]a,b[$, alors v et w coïncident sur tout l'intervalle $]a,b[$.*

Dém. Par contradiction, supposons que $v \not\equiv w$ sur $[t_0, b[$. Posons $t_1 = \inf\{t \geq t_0 \colon v(t) \neq w(t)\}$. On a $v(t_1) = w(t_1)$ par la continuité de v et w. Mais en appliquant le théorème de Cauchy-Lipschitz au problème de Cauchy avec la condition initiale centrée en t_1, nous voyons qu'il existe $r_0 > 0$ tel que les deux solutions v et w restent égales pour $t \leq t_1 + r_0$. Cela contredit la définition de t_1. Donc v et w coïncident sur $[t_0, b[$. De la même manière on prouve que les solutions coïncident sur $]a, t_0[$. □

*La condition b') pourrait être affaiblie en demandant seulement à f d'être, localement, lipschitzienne en u. Plus précisément, il s'agit de supposer qu'en tout point (t_0, u_0) de U, il existe une constante $L > 0$ (pouvant dépendre de (t_0, u_0) et un voisinage $V \subset U$ de (t_0, u_0) tel que,

$$|f(s,v) - f(s,w)| \leq L|v - w|$$

pour tout (s,v) et (s,w) dans V.

Définition 2.2 (prolongement, solutions maximales, courbes intégrales). *On dit qu'une solution $v\colon I \to \mathbb{R}$ d'une équation différentielle $u' = f(t,u)$, prolonge une autre solution $w\colon J \to \mathbb{R}$ (où I et J sont deux intervalles ouverts) de la même équation si $I \supset J$ et $v(t) = w(t)$ pour tout $t \in J$.*

On dit que v est une solution maximale si elle n'admet pas d'autres prolongements possibles autre que par la solution v elle même.

Le graphe d'une solution maximale d'une équation différentielle s'appelle courbe intégrale.

Remarque 2.6. La proposition précédente implique que deux courbes intégrales ne peuvent pas s'intersecter.

Exemple 2.7. Les courbes intégrales de l'équation différentielle $u' = -2tu$ sont les graphes des fonctions $t \mapsto Ce^{-t^2}$, définies pour tout $t \in \mathbb{R}$, où $C \in \mathbb{R}$.

Étant donnée une solution d'une équation différentielle, quand est-ce que celle-ci peut être prolongée ? Il y a une situation assez fréquente où le prolongement existe :

Proposition 2.6. *Supposons que $u\colon]a,b[\to \mathbb{R}$ soit une solution de $u' = f(t,u)$, où f vérifie (a') et (b'). S'il existe la limite $\lim_{t\to b} u(t) = u_0$, et si $(b,u_0) \in U$, alors la solution $u(t)$ se prolonge au delà de b. De même, s'il existe la limite $\lim_{t\to a} u(t) = u_0$, et si $(a,u_0) \in U$, alors la solution $u(t)$ se prolonge au delà de a.*

Dém. En effet, en appliquant le théorème de Cauchy-Lipschitz on trouve, dans un intervalle $[b, b+\delta[$ une solution v du problème de Cauchy $u'(t) = f(t,u(t))$ avec condition initiale $u(b) = u_0$. Posons $w(t) = u(t)$ si $t \in]a,b[$ et $w(t) = v(t)$ si $t \in [b, b+\delta[$. Par construction, w est continue, dérivable sur $]a,b[$ et sur $]b, b+\delta[$ et $\exists \lim_{t\to b} w'(t) (= f(b,u_0))$. Ces conditions impliquent que w est dérivable en b, avec $w'(b) = \lim_{t\to b} w'(t)$. En effet, si $b < t < b+\delta$, d'après le théorème des accroissements finies, il existe $\xi \in (b,t)$ tel que $w(t) - w(b) = w'(\xi)(t-b)$. En divisant par $t-b$, on voit que la limite $t \to b$ existe et vaut $\lim_{\xi\to b} w'(\xi)$. Autrement dit w est dérivable à droite en b. De la même manière w et dérivable à gauche en b et les valeurs de dérivées à droite et à gauche sont les mêmes. □

Le cas des fonctions $f\colon]a,b[\times \mathbb{R} \to \mathbb{R}$.

Si la fonction f, toujours supposée vérifier (a') et (b') est définie sur $]a,b[\times \mathbb{R}$ (donc $f(t,u)$ est définie pour tout $u \in \mathbb{R}$), bien souvent (mais pas toujours) les solutions maximales sont définies sur tout l'intervalle $]a,b[$. Dans ce cas on dit la solution maximale est *globale.* Dans le cas contraire on peut démontrer que les solutions possèdent une asymptote verticale, on dit alors que la solution *explose* à l'infini.

Exemple 2.8. Considérons le problème de Cauchy

$$\begin{cases} u' = u^2 \\ u(0) = -1. \end{cases}$$

Ici, $f(t,u) = u^2$, qui est définie sur $\mathbb{R} \times \mathbb{R}$. Ce problème possède comme solution maximale la fonction $u(t) = -1/(t+1)$, définie sur $]-1, +\infty[$, qui possède une asymptote verticale en -1. Noter qu'il s'agit là d'un exemple de solution maximale, qui n'est pas une solution globale, mais qui "explose".

Théorème 2.7 (Un résultat d'existence globale)**.** *Soit f une fonction vérifiant (a') et (b'), avec $U =]a, b[\times\mathbb{R}$. Supposons que pour tout intervalle compact $K \subset]a, b[$ il existe deux constantes A et B (éventuellement dépendentes de K, telles que, pour tout $(t, u) \in K \times \mathbb{R}$ on a*

$$|f(t, u)| \leq A + B|u|.$$

Alors toute solution de l'équation différentielle $u' = f(t, u)$ se prolonge à une solution "globale", c'est à dire, définie sur tout l'intervalle $]a, b[$.

Dém. Admis. □

Exemple 2.9. Toutes les solutions de l'équation différentielle

$$u' = (1 + 2t + 3tu)\cos(t^2 + u^2)$$

sont globales (définies pout tout $\mathbb{R}$). En effet, appliquons le théorème précédent avec $]a, b[= \mathbb{R}$. D'abord, $f(t, u) = (1+2t+3tu)\cos(t^2+u^2)$ est clairement de classe C^1 dans $\mathbb{R}\times\mathbb{R}$: cela suffit pour garantire la validité des hypothèses (a') et (b'). De plus pour tout intervalle $K \subset \mathbb{R}$, il existe $c > 0$ tel que $K \subset [-c, c]$. Mais $|f(t, u)| \leq 1 + |2t| + |3tu| \leq 1 + 2c + 3c|u|$. L'inégalité du théorème précédent est satisfaite avec $A = 1 + 2c$ et $B = 3c$.

2.4 Systèmes différentiels

Soient f_1, f_2,..., f_n des fonctions de $n+1$ variables réelles, t, u_1, $\ldots u_n$, définies sur un ouvert U de $\mathbb{R}^{n+1}$. Un système d'équations différentielles (d'ordre 1 et en forme normale) est un système de la forme

$$\begin{cases} u_1' = f_1(t, u_1, \ldots u_n) \\ \quad\vdots \\ u_n' = f_n(t, u_1, \ldots, u_n), \end{cases} \tag{2.3}$$

d'inconnues les fonctions $u_1(t)$,$\ldots u_n(t)$.

Définition 2.3. *On dit que les n fonctions $u_1(t)$,$\ldots u_n(t)$, définies sur un même intervalle $I \subset \mathbb{R}$ sont une solution du système (2.3) si, pour tout $t \in I$, le point $(t, u_1(t), \ldots, u_n(t))$ appartient à U et l'on a*

$$\begin{cases} u_1'(t) = f_1(t, u_1(t), \ldots u_n(t)) \\ \qquad\ldots \\ u_n'(t) = f_n(t, u_1, (t) \ldots, u_n(t)). \end{cases} \tag{2.4}$$

Le problème de Cauchy associé est le système ci-dessus complété par les conditions initiales

$$\begin{cases} u_1(t_0) = u_{0,1} \\ \vdots \\ u_n(t_0) = u_{0,n}. \end{cases}$$

Dans ce cas, il convient de considérer les fonctions vectorielles $u = (u_1, \ldots, u_n)$ et $f = (f_1, \ldots f_n)$. Par définition, la dérivée d'une fonction vectorielle se fait composante par composante : $u' = (u_1', \ldots, u_n')$. Donc le système (2.3) équivaut à la seule équation *vectorielle*

$$u' = f(t, u). \tag{2.5}$$

Le problème de Cauchy pour le systèmes différentiels du premier-ordre, ou problème de Cauchy vectoriel, consiste alors à résoudre

$$\begin{cases} u' = f(t, u) \\ u(t_0) = u_0. \end{cases} \tag{C}$$

Ici $u_0 = (u_{0,1}, \dots u_{0,n}) \in \mathbb{R}^n$ est donné, $f\colon U \to \mathbb{R}^n$ est une fonction définie dans un voisinage de U de $(t_0, u_0) \in \mathbb{R}^{1+n}$, et l'inconnue est la fonction vectorielle $u = (u_1, \dots, u_n)$.

Le théorème de Cauchy-Lipschitz reste valable dans le cas vectoriel. On l'énonce sous une forme un peu moins précise que dans le cas des équations scalaire:

Théorème 2.8 (de Cauchy-Lipschitz pour les systèmes différentiels)**.** *Soit $(t_0, u_0) \in \mathbb{R}^{1+n}$ et f une fonction continue au voisinage U de (t_0, u_0) et telle qu'il existe une constante $L > 0$ telle que, pour tout (t, v) et tout (t, w) dans U on a*

$$\|f(t, v) - f(t, w)\| \leq L\|v - w\|,$$

où on a noté par $\|\cdot\|$ la norme euclidienne de $\mathbb{R}^n$ (mais on peut prendre aussi la norme du maximum des composantes). Alors il existe un intervalle ouvert I_0, contenant t_0, tel que le problème de Cauchy (C) possède une unique solution $u = (u_1, \dots, u_n)$ définie dans I_0.

Remarque 2.10. Si f est de classe C^1, sur U alors l'hypothèse ci-dessus est automatiquement satisfaite. En fait, il suffit que f soit continue dans U et qu'elle soit de classe C^1 dans U en les variables $u_1, \dots, u_n$. Cette affirmation se justifie en appliquant l'inégalité des accroissement finis à la fonction vectorielle de n variables $u \mapsto f(t, u)$.

Par exemple, la fonction $f(t, u_1, u_2) = (|t| + u_1 + e_2^u, t - u_2)$ vérifie bien les hypothèses du théorème de Cauchy Lipschitz en tout point. Donc le problème de Cauchy $u' = f(t, u)$, avec condition initiale avec $u(t_0) = (u_1, u_2)$ possède une et une seule solution locale.

Théorème 2.9 (Existence globale pour le systèmes différentiels)**.** *avec les notations et les hypotheses du théorème précédent : supposons de plus que f pour tout intervalle compact $K \subset]a, b[$ il existe deux constantes A et B (éventullement dépendentes de K, telles que, pour tout $(t, u) \in K \times \mathbb{R}$ on a*

$$\|f(t, u)\| \leq A + B\|u\|.$$

Alors toute solution de l'équation différentielle $u' = f(t, u)$ se prolonge à une solution "globale", c'est à dire, définie sur tout l'intervalle $]a, b[$.

En général, on peut réduire une équation d'ordre supérieur à un système du premier ordre. Voici un exemple de la démarche :

Exemple 2.11. Considérons l'équation scalaire d'ordre 3

$$u''' = 3t(u'')^2 + e^{u \sin t}.$$

On introduit la fonction vectorielle $U = (U_1, U_2, U_3) := (u, u', u'')$. Avec ces notations on voit que l'équation donnée équivaut au système

$$\begin{cases} U_1' = U_2 \\ U_2' = U_3 \\ U_3' = 3tU_3^2 + e^{U_1 \sin t} \end{cases}$$

Ce système s'écrit sous forme vectorielle comme

$$U' = F(t, U).$$

Ici, la fonction vectorielle $F = (F_1, F_2, F_3)$ est une fonction des 4 variables t, U_1, U_2 et U_3. Ses composantes sont données par :

$$F_1(t, U_1, U_2, U_3) = U_2, \qquad F_2(t, U_1, U_2, U_3) = U_3 \quad \text{et} \quad F_3(t, U_1, U_2, U_3) = 3tU_3^2 + e^{U_1 \sin t}.$$

2.5 Équations différentielles linéaires

Définition 2.4. *Soient* $a_0(t)$, $a_1(t)$, $\dots a_{k-1}(t)$ *et* $f(t)$ *des fonctions continues sur un intervalle* $I \subset \mathbb{R}$. *Une équation différentielle* linéaire *est une équation de la forme*

$$u^{(k)} + a_{k-1}(t)u^{(k-1)} + \cdots + a_1(t)u' + a_0(t)u = f(t). \tag{2.6}$$

Si le terme à droite est identiquement nul sur I : $f(t) = 0$, *alors ; équation linéaire est dite* homogène.

Posons

$$E(u) = u^{(k)} + a_{k-1}(t)u^{(k-1)} + \cdots + a_1(t)u' + a_0(t)u.$$

Observons que l'application $u \mapsto E(u)$ est une application linéaire, au sens que, pour toute fonction u et v de classe C^k dans l'intervalle I, et pour tout scalaire λ et μ, on a

$$E(\lambda u + \nu v) = \lambda E(u) + \mu E(\nu).$$

Une conséquence de ceci est que si v et w sont deux solutions de (2.6), et donc $E(v) = f$, $E(w) = f$, alors $E(v - w) = 0$. Autrement dit, la différence de deux solutions d'une équation différentielle linéaire est la solution de l'équation différentielle *homogène* $E(u) = 0$.

Notons par $\mathcal{V}_f$ l'ensemble de toutes les solutions de l'équation (2.6) et par $\mathcal{V}_0$ l'ensemble de toutes les solutions de l'équation différentielle homogène qui lui est associée. Soit $u_f \in \mathcal{V}_f$ (u_f est donc une solution particulière de l'équation différentielle (2.6)).

- Si $v \in \mathcal{V}_f$, alors $v - u_f \in \mathcal{V}_0$.
- si $v \in V_0$, alors $u_f + v \in \mathcal{V}_f$.

Les deux affirmations précédentes se résument par l'égalité d'ensembles :

$$\mathcal{V}_f = u_f + \mathcal{V}_0.$$

On exprime cette égalité avec la terminologie suivante : *la solution générale de l'équation différentielle avec second membre est donnée par une solution particulière plus la solution générale de l'équation homogène associée.*

Équations d'ordre 1 et méthode de variation de la constante
Considérons l'équation linéaire d'ordre 1, sur un intervalle I.

$$u' + a(t)u = f(t). \tag{E}$$

L'équation homogène associée est

$$u' + a(t)u = 0. \tag{H}$$

Soit $A(t)$ une primitive sur I de $a(t)$. En multipliant cette équation par $e^{A(t)}$ on trouve $(e^{A(t)}u)' = 0$. Donc $e^{A(t)u} = c$ est constante sur I. Mais alors, la solution générale de l'équation homogène (H) est $u(t) = ce^{-A(t)}$. Pour trouver une solution particulière de l'équation (E), on peut faire appel à la méthode de variations des constantes : il s'agit de chercher une solution de (E) parmi les fonctions de la forme

$$u(t) = c(t)e^{-A(t)}.$$

Un petit calcul montre qu'il faut que $c'(t) = f(t)e^{A(t)}$. En conclusion, la solution générale de l'équation (E) est

$$u(t) = c(t)e^{-A(t)} + c\, e^{-A(t)}, \qquad \text{où } c(t) = \int f(t)e^{A(t)}\, \mathrm{d}t,\ A(t) = \int a(t)\, \mathrm{d}t, \text{ et } c \in \mathbb{R}.$$

Exemple 2.12. La solution générale sur $]0, +\infty[$ de l'équation différentielle

$$u' + u/t = e^t$$

est, d'après l'application de la méthode ci-dessus,

$$u(t) = \frac{(t-1)e^t}{t} + c\,\frac{1}{t}, \qquad c \in \mathbb{R}.$$

Théorème 2.10. *Pour tous réels $u_0,\ldots,\ u_{k-1}$, et $t_0 \in I$ il existe une et une seule solution $u\colon I \to \mathbb{R}$ de l'équation* (2.6)*, vérifiant les conditions initiales*

$$u(t_0) = u_0,\ u'(t_0) = u_1, \ldots,\ u^{(k-1)}(t_0) = u_{k-1}.$$

Toute solution maximale de l'équation différentielle (2.6) *est globale, c'est-à-dire définie sur l'intervalle I tout entier.*

Dém. En effet, introduisons les fonctions vectorielles $v\colon I \to \mathbb{R}^k$ et $F\colon I \times \mathbb{R}^k \to \mathbb{R}^k$, où

$$v = \begin{pmatrix} u \\ u' \\ \vdots \\ u^{(k-1)} \end{pmatrix}, \qquad \text{et} \qquad F(t,v) = \begin{pmatrix} v_2 \\ \vdots \\ v_k \\ -\big(a_{k-1}(t)v_k + \cdots + a_1(t)v_2 + a_0(t)v_1\big) + f(t). \end{pmatrix}$$

Observons qu'alors l'équation scalaire d'ordre k (2.6) équivaut à l'équation vectorielle d'ordre 1

$$v' = F(t,v).$$

Les conditions initiales se traduisent par la condition vectorielle $v(t_0) = v_0$, où $v_0 = (u_0, \ldots u_{k-1})$. Appliquons le théorème 2.8 : il conviendra de travailler avec la norme sur $\mathbb{R}^k$, $\|\alpha\|_1 = |\alpha_1| +$

$\cdots |\alpha_k|$ qui est équivalente à la norme euclidienne. La continuité de F au voisinage est assurée par l'hypothèse de continuité des coefficients $a_j(t)$ et $f(t)$. La condition de Lipshitz sur F découle du fait que, pour tout intervalle compact $K \subset I$, l'on peut trouver une constante C telle que, pour tout $t \in K$, on a $\max_{j=1,\dots k-1} |a_j(t)| \le L$. Mais alors,

$$\begin{aligned}\|F(t,v) - F(t,w)\|_1 &= \left\| \begin{pmatrix} (v-w)_2 \\ \vdots \\ (v-w)_k \\ -\big(a_{k-1}(t)(v-w)_k + \cdots + a_1(t)(v-w)_2 + a_0(t)(v-w)_1\big) \end{pmatrix} \right\|_1 \\ &\le (1+L)\|v-w\|_1.\end{aligned}$$

Les conditions du théore de Cauchy–Lipschitz vectoriel sont alors satisfaites. Posons $D = \max_{x \in K}\{|f(t)|$. Mais alors,

$$\begin{aligned}\|F(t,v)\|_1 &\le \|v\|_1 + \Big|a_{k-1}(t)u^{(k-1)} + \cdots + a_1(t)u' + a_0(t)u\Big|_1 \\ &\le (1+C)\|v\|_1 + D.\end{aligned}$$

Ceci permet d'appliquer le théorème d'existence globale (2.9). $\square$

Remarque 2.13. L'équation différentielle scalaire d'ordre k s'écrit aussi sous la forme matricielle suivante :

$$v' = M(t)v + B(t).$$

Ici,

$$M(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & & & & \\ -a_0(t) & -a_1(t) & \cdots & \cdots & -a_{k-1}(t) \end{pmatrix}, \qquad B(t) = \begin{pmatrix} 0 \\ \vdots \\ f(t) \end{pmatrix}.$$

Théorème 2.11. *L'ensemble $\mathcal{V}_0$ des solutions de l'équation différentielle homogène* (2.6)*, avec $f \equiv 0$, est un espace vectoriel de dimension k.*

Dém. En effet, pour $j = 0, \dots k-1$, considérons le problème de Cauchy

$$\begin{cases} E(u) = 0 \\ u^{(i)}(t_0) = \delta_{i,j}, \quad i = 0, \dots, k-1 \end{cases} \tag{C_j}$$

où $\delta_{i,j} = 1$ si $i = j$ et 0 si $i \neq j$ est le symbole de Kronecker. Le problème (C_j) possède une unique solution $u_j(t)$, définie sur tout l'intervalle I. Ces solutions $u_0(t)$, $u_1(t)$,$\dots$ $u_{k-1}(t)$ sont linéairement indépendentes dans l'espace $\mathcal{V}_0$. En effet, si

$$\lambda_0 u_0 + \lambda_1 u_1 + \cdots + \lambda_{k-1} u_{k-1} = 0,$$

comme la fonction $u = \lambda_0 u_0 + \cdots + \lambda_{k-1} u_{k-1} = 0$ est solution du problème de Cauchy

$$\begin{cases} E(u) = 0 \\ u(t_0) = \lambda_0 \\ \cdots \\ u(t_{k-1}) = \lambda_{k-1}, \end{cases}$$

on doit avoir

$$\lambda_0 = \lambda_1 = \ldots = \lambda_{k-1} = 0.$$

D'autre part, toute solution w de l'équation (2.6) s'écrit

$$w(t) = w(t_0)u_0(t) + \cdots w^{(k-1)}(t_0)u_{k-1}(t).$$

En effet, le terme à droite et la fonction w sont solutions d'un même problème de Cauchy et donc elle doivent être la même solution. Autrement dit, $(u_0, \dots u_{k-1})$ est un système de générateurs pour $\mathcal{V}_0$. $\square$

2.6 Équations différentielles linéaires à coefficients constants

Une équation différentielle linéaire *homogène* à coefficients constants est une équation de la forme

$$u^{(k)} + a_{k-1}u^{(k-1)} + \cdots + a_1u' + a_0u = 0, \tag{H}$$

où $a_{k-1}, \dots a_0$ sont des constantes réelles.

Introduisons le polynôme caractéristique de cette équation, qui par définition est le polynôme

$$P(\lambda) = \lambda^k + a_{k-1}\lambda^{k-1} + \cdots + a_1\lambda + a_0.$$

Exemple 2.14. L'équation différentielle homogène d'ordre 2

$$u'' - 3u' + 2u = 0$$

a pour polynôme caractéristique $P(\lambda) = \lambda^2 - 3\lambda + 2$, qui possède les deux racines réelles $\lambda_1 = 1$ et $\lambda_2 = 2$. Observons que e^t et e^{2t} sont deux solutions linéarement indépendantes de l'équation différentielle. Donc l'équation a pour solution générale

$$u(t) = c_1e^t + c_2e^{2t}, \qquad c_1, c_2 \in \mathbb{R}.$$

Exemple 2.15. L'équation différentielle homogène d'ordre 2

$$u'' - 2u' + u = 0$$

a pour polynôme caractéristique $P(\lambda) = \lambda^2 - 2\lambda + 1$, qui possède une racine double $\lambda = 1$. Observons que e^t est bien une solution de l'équation différentielle. Mais cela ne suffit pas pour décrire la solution générale $\mathcal{V}_0$, qui est un espace de dimension 2. Observons cependant que te^t est une autre solution de l'équation, indépendente de la précédente. Donc l'équation a pour solution générale

$$u(t) = c_1e^t + c_2\, te^t, \qquad c_1, c_2 \in \mathbb{R}.$$

Proposition 2.12. *Si λ_1, λ_2, …, λ_k sont les k racines (éventuellement complexes) du poynôme caractéristique, alors les k fonctions*

$$e^{\lambda_1 t},\ e^{\lambda_2 t},\ \dots,\ e^{\lambda_k t}$$

sont des solutions de l'équation homogène (H).
- Si toutes les racines sont distinctes, ces solutions engendrent la solution générale de l'équation (H).
- S'il y a une racine λ avec multiplicité $m \geq 1$, alors les fonctions

$$e^{\lambda t},\ te^{\lambda t},\ \dots,\ t^{m-1}e^{\lambda t}$$

sont également des solutions de l'équation (H).

Si l'une de ces racines est complexe, par exemple $\lambda = \alpha + i\beta$, alors $\bar{\lambda} = \alpha - i\beta$ est aussi une racine. Mais

$$\begin{aligned} e^{\lambda\, t} &= e^{\alpha\, t}(\cos(\beta t) + i\sin(\beta\, t)) \\ e^{\bar{\lambda}\, t} &= e^{\alpha\, t}(\cos(\beta t) - i\sin(\beta\, t)). \end{aligned}$$

Ces fonctions sont alors deux solutions à valeurs complexes de l'équation (H). On construit cependant deux solutions à valeurs réelles en prenant

$$\begin{aligned} \frac{e^{\lambda\, t} + e^{\lambda\, t}}{2} &= e^{\alpha\, t}\cos(\beta\, t) \\ \frac{e^{\lambda\, t} - e^{\lambda\, t}}{2i} &= e^{\alpha\, t}\sin(\beta\, t). \end{aligned}$$

Exemple 2.16. L'équation

$$u''' - 2u'' + 2u' = 0$$

admet le polynôme caractéristique $P(\lambda) = \lambda^3 - 2\lambda^2 + 2\lambda$ dont les racines sont $\lambda = 0$, $\lambda = 1 + i$ et $\lambda = 1 - i$. La solution générale est alors

$$u(t) = c_1 + c_2 e^t \cos t + c_2 \sin t, \qquad c_1, c_2, c_3 \in \mathbb{R}.$$

Plus en général, on peut considérer les équations différentielles linéaires à coefficients constants avec second membre.

$$u^{(k)} + a_{k-1}u^{(k-1)} + \cdots + a_1 u' + a_0 u = f(t). \tag{E}$$

Les techniques vu précedemment fournissent la solution générale de l'équation homogène associée. Pour trouver la solution générale de l'équation (E) il ne reste qu'à trouver une solution particulière de (E). Pour ce faire, on peut chercher d'abord des solutions qui "ressemblent" à la fonction $f(t)$ *. Si on n'en trouve pas, il peut être utile d'appliquer la méthode de variations des constantes.

2.7 Inégalités de Gronwall

Les inégalités de Gronwall sont des inégalités différentielles ou intégrales. Il en existe plusieurs variantes, qu'il n'est pas indispensable de connaitre par coeur.

Lemme 2.13 (version différentielle)**.** *Soient $a(t)$ et $b(t)$ deux fontions continues dans un intervalle I de $\mathbb{R}$. Si u est une fonction de classe C^1 dans I vérifiant*

$$\forall\, t \in I, :\ u'(t) \leq a(t)u'(t) + b(t),$$

et $t_0 \in I$, alors

$$u(t) \leq u(t_0)e^{\int_{t_0}^t a(\tau)\,\mathrm{d}\tau} + \int_{t_0}^t e^{\int_s^t a(\tau)\,\mathrm{d}\tau}\, b(s)\,\mathrm{d}s.$$

Dém. On raisonne comme dans le cas d'égalité et on multiplie terme-à-terme par $e^{-\int_{t_0}^t a(\tau)\,\mathrm{d}\tau}$. Ceci est légitime puisque on multiplie là l'inégalité par une expression positive. Ensuite il suffit d'intégrer terme-à-terme. □

*Si $f(t)$ est de la forme $P(t)e^{\lambda t}$, avec $P(t)$ polynôme, on cherchera une solution de la forme $Q(t)e^{\lambda t}$ avec $Q(t)$ polynôme du même degré que $P(t)$. Si $f(t)$ est de la forme $\sin(\lambda t)$, ou $\cos(\lambda t)$, on cherchera une solution de la forme $A\cos t + B\sin t$. Cette méthode ne fonctionne pas si la solution particulière que l'on cherche de l'équation avec second membre s'avère être une solution de l'équation homogène associée.

La version intégrale est plus souvent plus utile :

Lemme 2.14 (version intégrale)**.** *Soit $a(t)$ une fonction continue sur $[0,T]$, où $T>0$, avec $a(t)\geq 0$ pour tout $t\in[0,T]$. Si $u(t)$ est continue et vérifie l'inégalité intégrale*

$$\forall\, t\in[0,T]\colon\ u(t)\leq C+\int_0^t a(s)u(s)\,\mathrm{d}s,$$

avec $C\geq 0$ indépendente de t, alors

$$u(t)\leq Ce^{\int_0^t a(s)\,\mathrm{d}s}.$$

Dém. Posons

$$f(t)=\frac{C+\int_0^t a(s)u(s)\,\mathrm{d}s}{\exp\Big(\int_0^t a(s)\,\mathrm{d}s\Big)}.$$

Alors,

$$f'(t)=a(t)\frac{u(t)-C-\int_0^t a(s)u(s)\,\mathrm{d}s}{\exp\Big(\int_0^t a(s)\,\mathrm{d}s\Big)}\leq 0.$$

Donc $f(t)\leq f(0)=C$. Mais

$$u(t)\leq C+\int_0^t a(s)u(s)\,\mathrm{d}s=f(t)\exp\Big(\int_0^t a(s)\,\mathrm{d}s\Big)$$

et l'inégalité de Gronwall en découle.

□

Il existe de nombreuses variantes de l'inégalités de Gronwall, qui généralisent les résultats des lemmes précédents.

3 Chapitre vide

4 Intégrales multiples

Voir les notes de Vincent Borrelli

5 Courbes paramétrées

5.1 Courbes paramétrées

Dans tout ce chapitre $\|x\|$ désigne la norme euclidienne du vecteur $x\in\mathbb{R}^n$.

Une courbe dans $\mathbb{R}^n$ est un ensemble $\Gamma\subset\mathbb{R}^n$ tel qu'il existe une application continue $\varphi\colon[a,b]\to\mathbb{R}^n$, où $[a,b]$ est un intervalle de $\mathbb{R}$, telle que

$$\Gamma=\{\varphi(t)\colon t\in[a,b]\}.$$

On dit aussi que φ est une *paramétrisation* de Γ, ou aussi que Γ est le *support* de la courbe paramétrée φ. Une paramétrisation permet de définir une *orientation* (c'est-à-dire un sens de parcours) sur Γ : le parcours allant du point $\varphi(a)$ au point $\varphi(b)$.

Interpretation cinématique. Si $x(t)$, $y(t)$ et $z(t)$ sont les coordonnées d'un point materiel mobile à l'instant t, la *loi horaire* du mouvement, c'est à dire l'application $t \mapsto (x(t), y(t), z(t))$ définit une courbe paramétrée φ dans $\mathbb{R}^3$. Dans cet interpretation cinématique, le support Γ de la courbe est la *trajectoire* du point matériel. La *vitesse* instantanée du point est donnée par le vecteur $\varphi'(t) = (x'(t), y'(t), z'(t))$.

Définition 5.1. - *On dit qu'une paramétrisation* $\varphi\colon [a,b] \to \mathbb{R}^n$ *d'une courbe* Γ *est* régulière *si* φ *est de classe* C^1 *(c'est à dire que ses composantes* $\varphi_1, \ldots, \varphi_n$ *sont toutes des fonctions de classe* C^1*) et si, pour tout* $t \in [a,b]$*, le vecteur* $\varphi(t) \in \mathbb{R}^n$ *n'est pas le vecteur* $(0, \ldots, 0)$*. S'il existe une partition finie* $a = a_0 \leq a_1 \leq a_2 \leq \ldots \leq a_N = b$ *de l'intervalle* $[a,b]$ *telle que* φ *est régulière dans tout intervalle* $[a_i, a_{i+1}]$ *on dit que* φ *est* régulière par morceaux. *Un point où* $\varphi'(t_0)$ *s'annule est dit* point singulier.

- *On dit qu'une courbe paramétrée* $\varphi\colon [a,b] \to \mathbb{R}^n$ *est* fermée *si* $\varphi(a) = \varphi(b)$.

- *On dit qu'une courbe paramétrée* $\varphi\colon [a,b] \to \mathbb{R}^n$ *est* simple *si* $\varphi(t) \neq \varphi(t')$*, pour tout* $t \neq t'$ *et* $a < t < b$*,* $a < t' < b$*. Cela signifie que la courbe n'a pas d'intersection avec elle même.*

Exemple 5.1. - Si $f\colon [a,b] \to \mathbb{R}$ est une fonction de classe C^1, on peut toujours lui associer une courbe paramétrée dans $\mathbb{R}^2$ régulière, simple, non fermée, en prenant son graphe. Cela revient à poser $\varphi(t) = (t, f(t))$, $t \in [a,b]$.

- La fonction $\varphi(t) = (\cos(t), \sin(t))$, où $t \in [0, 2\pi]$ définit une paramétrisation régulière, simple, fermée. Le support de cette courbe paramétrée est le cercle de centre O et de rayon 1.

- La courbe paramétrée $\varphi(t) = (t^2, t^3)$, $t \in [-2, 2]$ est simple non régulière. En effet, observons que le vecteur dérivée $\varphi'(t) = (2t, 3t^2)$ s'annule pour $t = 0$. Pour dessiner cette courbe, on part du système $\begin{cases} x = t^2 \\ y = t^3 \end{cases}$ et on obtient, en éliminant t (il faut distinguer $t \geq 0$ et $t < 0$) $y = \pm x^{2/3}$, avec $0 \leq x \leq 4$. On voit alors sans peine que le support Γ de cette courbe présente un cusp à l'origine.

- Soit $a \in \mathbb{R}^n$ et v un vecteur de $\mathbb{R}^n$. L'équation paramétrique de la droite d passant par le point $a = (a_1, \ldots a_n)$ et parallèle au vecteur $v = (v_1, \ldots, v_n)$ est l'application

$$\varphi(t) = a + tv \qquad t \in \mathbb{R}$$

Si $t_0 \in \mathbb{R}$, la même droite d peut aussi être paramétrée par

$$\varphi(t) = a + (t - t_0)v \qquad t \in \mathbb{R}.$$

- Les bords des polygônes sont des exemples typiques de courbes qui admettent des paramétrisations régulières par morceaux, mais pas de paramétrisation régulières.

Droite et verseur tangent. Soit φ une paramétrisation régulière, définie sur un intervalle $[a,b]$, et soit t_0 et t_1 deux points de cet intervalle. La droite passant par les points $\varphi(t_0)$ et $\varphi(t_1)$ s'identifie à la droite passant par $\varphi(t_0)$ et parallèle au vecteur $v = \frac{\varphi(t_1) - \varphi(t_0)}{t_1 - t_0}$. Ainsi, cette droite est paramétrée par l'application

$$t \mapsto \varphi(t_0) + (t - t_0)\frac{\varphi(t_1) - \varphi(t_0)}{t_1 - t_0}, \qquad t \in \mathbb{R}.$$

En passant à la limite pour $t_1 \to t_0$ on trouve

$$\varphi(t_0) + \varphi'(t_0)(t - t_0), \qquad t \in \mathbb{R}. \tag{5.1}$$

qui est l'équation paramétrique de la *droite tangente* au support de φ au point $\varphi(t_0)$. Observons que cette droite est parallèle au vecteur $\varphi'(t_0)$. Le vecteur

$$\vec{T}(t_0) = \frac{\varphi'(t_0)}{\|\varphi'(t_0)\|}$$

qui est de norme unitaire, est donc le *verseur tangent* à la courbe au point $\varphi(t_0)$.

5.2 Longueur d'une courbe

Pour définir la longueur d'une courbe paramétrée $\varphi\colon [a,b] \to \mathbb{R}^n$, considérons la famille $\mathcal{P}$ de toutes les partitions finies possibles $a = a_0 < a_1 < \ldots < a_N = b$ de l'intervalle $[a,b]$. À partir d'une telle partition, on considère les points $\varphi(a_i)$, $i = 0,\ldots,N$ appartenant au support de la courbe et les ségments de $\mathbb{R}^n$, $[\varphi(a_i), \varphi(a_{i+1})] = \{(1-t)\varphi(a_i) + t\varphi(a_{i+1})\colon t \in [0,1]\}$. Si $P \in \mathcal{P}$ est une telle partition, la quantité

$$L(P) = \sum_{i=0}^{N-1} \|\varphi(a_i) - \varphi(a_{i+1})\|$$

n'est rien d'autre que la longueur de la ligne affine par morceaux qui relie les points $\varphi(a_i)$. Il est alors naturel de définir la longueur de la courbe comme

$$\mathrm{Long}(\varphi) = \sup\{L(P)\colon P \in \mathcal{P}\}.$$

On peut démontrer que si φ est une courbe régulière par morceaux, alors la quantité ci-dessus est toujours finie.

Théorème 5.1. *Pour toute courbe paramétrée régulière par morceaux, on a*

$$\mathrm{Long}(\varphi) = \int_a^b \|\varphi'(t)\|\,\mathrm{d}t.$$

La démonstration de ce théorème est hors programme.

Exemple 5.2 (Longueur du cercle)**.** En paramétrant le cercle centré en O et de rayon R par $\varphi(t) = (R\cos t, R\sin t)$, $t \in [0,2\pi]$ on voit que sa longueur est $\int_0^{2\pi} \sqrt{(-R\sin t)^2 + (R\cos t)^2}\,\mathrm{d}t = 2\pi R$.

5.3 Paramétrisations équivalentes et abscisse curviligne

Une même courbe Γ peut être paramétrée de plusieurs manière différentes.

Définition 5.2. *On dit que deux paramétrisations régulières par morceaux $\varphi\colon [a,b] \to \mathbb{R}^n$ et $\psi\colon [c,d] \to \mathbb{R}^n$ d'une même courbe sont équivalentes s'il existe un difféomorphisme $h\colon [a,b] \to [c,d]$ (h est donc une application bijective et h et h^{-1} sont de classe C^1) telle que, pour tout $t \in [a,b]$, on a*

$$\varphi(t) = \psi(h(t)).$$

Dans ce cas on écrit $\varphi \sim \psi$

La relation $\sim$ est une relation d'équivalence. Observons que pour un difféomorphisme $h\colon [a,b] \to [c,d]$ il y a deux possibilité : soit h est strictement croissante, soit h est strictement décroissante (c'est une conséquence de la bijectivité, de la continuité, et du théorème des valeurs intermédiaires). Dans le premier cas on dit que φ et ψ ont la même orientation.

Exemple 5.3. Un quart de cercle peut être paramétrée par $t \mapsto (R\cos t, R\sin t)$, $t \in [0, \pi/2]$, mais aussi, en tant graphe d'une fonction, par $x \mapsto (x, \sqrt{R^2 - x^2})$, $x \in [0, R]$. On construit le difféomorphisme $h\colon [0, \pi/2] \to [0, R]$ en posant $x = h(t) = \mathbb{R}\cos t$. Les deux paramérisations sont alors équivalentes. Observons que le sens de parcours n'est pas le même dans ce deux cas. En effet la fonction h est décroissante.

Remarque 5.4. Si $\varphi\colon [a, b] \to \mathbb{R}^n$ et $\psi\colon [c, d] \to \mathbb{R}^n$, avec $\varphi \sim \psi$ par un difféomorphisme h, et si $s_0 = h(t_0)$, alors le verseur tangent au point $\varphi(t_0) = \psi(s_0)$ est

$$\vec{T}(t_0) = \frac{\varphi'(t_0)}{\|\varphi(t_0)\|} = \frac{\psi'(h(t_0))h'(t_0)}{\|\psi'(h(t_0)\| \, |h'(t_0)|} = \pm \frac{\psi'(s_0)}{\|\psi(s_0)\|},$$

avec le signe $+$ si φ et ψ ont la même orientation et le signe $-$ dans le cas contraire. Ce calcul montre que la droite tangente à une courbe Γ ne dépend pas de la paramétrisation choisie.

Définition 5.3. *Une* paramétrisation normale*, ou* paramérisation par abscisse curviligne *d'une courbe* Γ*, est une paramétrisation* $\psi\colon [0, \mathrm{Long}(\Gamma)] \to \mathbb{R}^n$ *telle que* $\|\psi'(s)\| = 1$ *pour tout* $0 \leq s \leq \mathrm{Long}(\Gamma)$.

Proposition 5.2. *Toute courbe* Γ *régulière par morceaux admet une paramétrisation normale équivalente*

Dém. À partir d'une paramétrisation régulière par morceaux $\varphi\colon [a, b] \to \mathbb{R}^n$, construisons une paramétrisation normale $\psi \sim \varphi$. Pour cela il suffit de définir $h\colon [a, b] \to [0, \mathrm{Long}(\varphi)]$, avec

$$h(t) = \int_a^t \|\varphi'(\tau)\| \,\mathrm{d}\tau.$$

Cette application h définit bien un difféomorphisme entre $[a, b]$ et $0, \mathrm{Long}(\varphi)]$. Ensuite il suffit de poser $\psi(s) = \varphi(h^{-1}(s))$ (ou $\varphi(t) = \psi(h(t))$, de manière équivalente). Observons que $\psi'(s) = \varphi'(h^{-1}(s))((h^{-1})'(s))$. Mais, pour tout s, $h(h^{-1}(s)) = s$ et en dérivant terme-à-terme cette relation on trouve $(h^{-1})'(s) = 1/h'(h^{-1}(s))$ (c'est la formule usuelle de la dérivée de l'application inverse). Comme $h'(t) = \|\varphi(t)\|$, on trouve

$$\psi'(s) = \varphi'(h^{-1}(s)) \, \frac{1}{\|\varphi'(h^{-1}(s))\|}.$$

Mais alors $\|\psi'(s)\| = 1$.

□

L'abscisse curviligne (c'est-à-dire, le paramètre d'une paramétrisation normale) d'une courbe Γ est traditionnellement notée avec la lettre s. Observons que la différentielle $\mathrm{d}s = \|\varphi'(t)\| \,\mathrm{d}t$ s'interprete comme un "élément infinitésimal de longueur de la courbe".

Définition 5.4. *Soit* $f\colon \Gamma \to \mathbb{R}$ *une fonction continue, définie sur le support* Γ *d'une courbe régulière par morceaux, paramétrée par* $\varphi\colon [a, b] \to \mathbb{R}$*. L'intégrale curviligne de* f *le long* Γ *est définie par*

$$\int_\Gamma f \,\mathrm{d}s = \int_a^b f(\varphi(t)) \|\varphi'(t)\| \,\mathrm{d}t.$$

Observons que dans le cas particulier où Γ est un intervalle $[a, b]$, en le paramétrant par $\varphi(t) = t$, on trouve que $\int_\Gamma f \,\mathrm{d}s = \int_a^b f(t) \,\mathrm{d}t$. L'intégrale curviligne est dont une généralisation de l'intégrale usuelle.

Il est facile, à l'aide du théorème de changement de variables, de prouver que $\int_\Gamma f \,\mathrm{d}s$ est indépendant de la paramétrisation choisie pour Γ. En particulier (en prenant la fonction constante $f = 1$) on voit que la longueur est indépendante du choix de la paramétrisation.

Exemple 5.5. Une paroi verticale est limitée par le haut par la courbe d'équation $t \mapsto (t, t^2, 6t)$, $t \in [0,1]$. Quelle est l'aire de la surface de cette paroi ? Réponse. Le profil au sol de la paroi est la courbe Γ de $\mathbb{R}^2$ paramétrée par $t \mapsto (t, t^2)$, $t \in [0,1]$. Considérons ensuite la fonction $f(x,y) = 6x$. Il s'agit de calculer l'intégrale curviligne $\int_\Gamma f \, \mathrm{d}s = \int_0^1 \sqrt{1+4t^2}\, 6t \, \mathrm{d}t = [(1+4t^2)^{3/2}]_{t=0}^{t=1} = 5^{3/2} - 1$.

5.4 La courbure

Soit $\psi \colon [0, \mathrm{Long}(\Gamma)] \to \mathbb{R}^2$ la **paramétrisation normale** d'une courbe Γ. On note par s l'abscisse curviligne. On suppose que φ est de classe C^2.

On sait que $\|\psi'(s)\| = 1$, donc

$$\vec{T}(s) = \psi'(s),$$

est le verseur tangent à la courbe au point $\psi(s)$. Soit $\vec{N}(s)$ le verseur normal à $\vec{T}(s)$, de manière que $(\vec{T}(s), \vec{N}(s))$ soit une base orthonormé directe du plan*.

Proposition 5.3. *Il existe une fonction continue $s \mapsto \bar{\kappa}(s)$ telle que $\psi''(s) = \bar{\kappa}(s)\vec{N}(s)$. En particulier, si la courbe est paramétrée à l'aide de l'abscisse curviligne, alors la direction normale à la courbe au point $\varphi(s)$ est donnée par le vecteur $\varphi''(s)$.*

Dém. En effet $\psi'(s) \cdot \psi'(s) = \vec{T}(s) \cdot \vec{T}(s) = \|\vec{T}(s)\|^2 = 1$. En dérivant par rapport à s on trouve $2\psi''(s) \cdot \psi'(s) = 0$. Mais alors $\psi''(s)$ est orthogonal à $\psi'(s)$, et donc parallèle à $\vec{N}(s)$. □

Définition 5.5. *La courbure algébrique $\bar{\kappa}(s)$ de ψ au point $\psi(s)$ est le nombre réel*

$$\bar{\kappa}(s) = \psi''(s) \cdot \vec{N}(s) = \vec{T}'(s) \cdot \vec{N}(s).$$

La courbure de φ au point $\varphi(s)$ est le nombre ≥ 0

$$\kappa(s) = |\bar{\kappa}(s)| = \|\psi''(s)\|.$$

Le rayon de courbure est $R(s) = \frac{1}{\kappa(s)}$.

Si $\kappa(s) \neq 0$, on dit que le point $\psi(s)$ est birégulier.

Exemple 5.6. Le cercle de rayon R admet la paramétrisation normale $\psi(s) = (R\cos\frac{s}{R}, R\sin\frac{s}{R})$. La courbure vaut $\kappa(s) = \|\psi''(s)\| = \frac{1}{R}$. La courbure d'un cercle de rayon R est donc constante et égale à $1/R$. Le rayon de courbure du cercle est alors constant égale à R.

Étant donnée une courbe Γ, il n'est pas aisé d'expliciter la paramétrisation normale (en effet, l'intégrale de longueur est souvent difficile à calculer). À partir d'une paramétrisation quelconque $t \mapsto \varphi(t)$, on peut trouver la courbure à l'aide de la proposition suivante :

Proposition 5.4. *Soit $\varphi \colon [a,b] \to \mathbb{R}^2$ une courbe paramétrée régulière de classe C^2. Alors pour tout $t \in [a,b]$ on a*

$$\bar{\kappa}(t) = \frac{\det(\varphi'(t), \varphi''(t))}{\|\varphi'(t)\|^3}, \qquad \text{et} \qquad \kappa(t) = \frac{|\det(\varphi'(t), \varphi''(t))|}{\|\varphi'(t)\|^3}.$$

Dém. Considérons la paramétrisation normale $\psi(s) = \varphi(t(s))$ de Γ, où $t(s) = h^{-1}(s)$, h étant le difféomorphisme introduit dans la démonstration de la Proposition 5.2. Nous avons (voir la démonstration de la Proposition 5.2):

$$\psi'(s) = \frac{\varphi'(t(s))}{\|\varphi'(t(s))\|}.$$

*Rappelons que si $\vec{T}$ et $\vec{N}$ sont deux vecteurs dans le plan orthogonaux, de norme unitaire, alors le déterminant de la matrice 2×2 $(\vec{T}, \vec{N})$ est de déterminant ± 1. Affirmer que $(\vec{T}(s), \vec{N}(s))$ est une "base directe" signifie que le déterminant est positif

En dérivant par rapport à s on trouve

$$\psi''(s) = \frac{\varphi''(t(s))}{\|\varphi'(t(s))\|^2} + \lambda(s)\varphi'(t(s)),$$

où $\lambda(s)$ est la dérivée de la fonction $s \mapsto \frac{1}{\|\varphi'(t(s))\|}$, qu'on n'aura pas besoin de calculer. Donc, en appliquant les propriétés élémentaires du déterminant $\det(\vec{a}, \vec{b} + \lambda\vec{c}) = \det(\vec{a}, \vec{b}) + \lambda \det(\vec{a}, \vec{c})$ et $\det(\vec{c}, \vec{c}) = 0$, on voit que

$$\det(\psi'(s), \psi''(s))) = \frac{\det(\varphi'(t), \varphi''(t))}{\|\varphi'(t)\|^3}.$$

D'autre part,

$$\det(\varphi'(s), \psi''(s)) = \det(\vec{T}(s), \bar{\kappa}(s)\vec{N}(s)) = \bar{\kappa}(s).$$

□

5.5 Étude locale des courbes planes

Considérons une paramétrisation normale $\psi\colon [0, \mathrm{Long}(\Gamma)] \to \mathbb{R}^2$ d'une courbe Γ, et supposont ψ de classe C^2.

Au voisinage de $s_0 \in [0, \mathrm{Long}(\Gamma)] \to \mathbb{R}^2$, en écrivant un développement limité d'ordre 2 on trouve

$$\psi(s) = \psi(s_0) + (s - s_0)\psi'(s_0) + \frac{(s - s_0)^2}{2}\psi''(s_0) + o((s - s_0)^2).$$

On sait déjà que $s \mapsto \varphi(s_0) + (s - s_0)\varphi'(s_0)$ est la paramétrisation de la droite tangente à la courbe au point $\psi(s_0)$. Le terme suivant du développement $\frac{(s-s_0)^2}{2}\psi''(s_0)$ (son signe, notamment) nous renseigne sur le positionnement de la courbe par rapport à cette droite, au voisinage de $\psi(s_0)$. Plus précisément, réecrivons le développement précédent comme

$$\psi(s) = \psi(s_0) + (s - s_0)\vec{T}(s_0) + \bar{\kappa}(s_0)\frac{(s - s_0)^2}{2}\vec{N}(s_0) + o((s - s_0)^2).$$

Si $\psi(s_0)$ est un point bi-régulier, les vecteurs $\vec{T}(s_0)$ et $\vec{N}(s_0)$ sont bien définis. Dans le repère orthonormée $(\psi(s_0), \vec{T}(s_0), \vec{N}(s_0))$, c'est-à-dire le repère avec pour origine $\psi(s_0)$ et base $(\vec{T}(s_0), \vec{N}(s_0))$, la courbe Γ admet alors pour composantes (en négligeant les termes d'ordre supérieur)

$$(s - s_0, \bar{\kappa}(s_0)\frac{(s - s_0)^2}{2}).$$

Ainsi, dans ce repère la courbe se comporte localement comme une parabole.

Exemple d'étude d'un point singulier Si $\varphi\colon [a, b] \to \mathbb{R}^2$ est une paramétrisation d'une courbe admettant un point singulier $\varphi(t_0)$, la formule pour la construction de la droite tangente (5.1) ne s'applique plus. Certaines courbes admettent cependant une droite tangente aux points singuliers. Si c'est le cas, la pente de cette droite est donnée par la limite

$$\ell = \lim_{t \to t_0} \frac{y(t) - y(t_0)}{x(t) - x(t_0)}.$$

La droite est alors celle d'équation $y = \ell(x - x(t_0)) + y(t_0)$.

Exemple 5.7. Étudions le point singulier de la courbe paramétrée par $x(t) = 1 + t^2$ et $y(t) = -1 + t^2 - t^3$. Le point singulier est $(x(0), y(0)) = (1, -1)$. La limite $\ell = \lim_{t\to 0} \frac{y(t)-y(0)}{x(t)-x(0)}$ existe et vaut $\ell = 1$. La droite tangente au point singulier est donc la droite d'équation $y = x - 2$. Pour connaître l'allure de la courbe au voisinage du point singulier, étudions la position de la courbe par rapport au point $(1, -1)$: pour cela observons que $x(t) - 1 = t^2 \geq 0$ et $y(t) + 1 = t^2 - t^3 \geq 0$ au voisinage de $t = 0$, la courbe reste alors à la droite et en haut par rapport au point $(1, -1)$, au voisinage de celui-ci. Étudions ensuite la position de la courbe par rapport à sa droite tangente : on a $y(t) - x(t) + 2 = -t^3$, qui change de signe en $t = 0$. Ainsi la courbe est à la fois au dessus et en dessous de sa droite tangente au voisinage du point $(1, -1)$. Ces renseignements nous permettent de dire que la courbe présente en $(1, -1)$ une singularité de type “cusp”.

6 Nappes paramétrées et surfaces de $\mathbb{R}^3$

6.1 Rappel sur le produit vectoriel

Soit $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ la base canonique de $\mathbb{R}^3$. Considérons deux vecteurs $\vec{a}$ et $\vec{b}$ de $\mathbb{R}^3$ non co-linéaires. Rappelons que le vecteur

$$\vec{a} \wedge \vec{b} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} \tag{6.1}$$

est le vecteur orthogonal à $\vec{a}$ et $\vec{b}$, de norme $\|\vec{a} \wedge \vec{b}\| = \|\vec{a}\|\,\|\vec{b}\|\,|\sin\theta|$, où θ est l'angle orienté compris entre $\vec{a}$ et $\vec{b}$ et tel que $(\vec{a}, \vec{b}, \vec{a} \wedge \vec{b})$ forme une base directe de $\mathbb{R}^3$. La norme $\|\vec{a} \wedge \vec{b}\|$ est donc égale à l'aire du parallélogramme engendrée par $\vec{a}$ et $\vec{b}$.

Le produit mixte

$$[\vec{a}, \vec{b}, \vec{u}] = (\vec{a} \wedge \vec{b}) \cdot \vec{u} = \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ u_1 & u_2 & u_3 \end{pmatrix}$$

s'interprète comme le volume (avec signe) du parallélipipède engendré par les vecteurs $\vec{a}$, $\vec{b}$ et $\vec{u}$.

6.2 Surfaces et nappes paramétrées

Une *surface* est une partie $\Sigma \subset \mathbb{R}^3$, de la forme

$$\Sigma = \{\varphi(u, v)\colon u, v \in K\} \subset \mathbb{R}^3.$$

où φ est une application $\varphi\colon K \to \mathbb{R}^3$, et $K = \overline{U}$ est un compact de $\mathbb{R}^2$, égale à l'adhérence d'un ouvert $U \subset \mathbb{R}^2$. L'application φ est une *paramétrisation* de Σ. On dit aussi que φ est une nappe paramétree et que l'ensemble Σ est le support de la nappe.

Exemple 6.1. (paramétrisation d'un plan de $\mathbb{R}^3$) Soient $\vec{a}$ et $\vec{b}$ deux vecteurs de $\mathbb{R}^3$ non colinéaires. Ces deux vecteurs engendrent un plan π passant par l'origine, qui est le sous-espace vectoriel de $\mathbb{R}^3$ donné par $\pi = \{u(a_1, a_2, a_3) + v(b_1, b_2, b_3)\colon u, v \in \mathbb{R}\}$.

Si $\bar{P} = (\bar{x}_1, \bar{x}_2, \bar{x}_3) \in \mathbb{R}^3$, alors le plan passant par $\bar{P}$ et parallèle à π admet comme équation paramétrique

$$\varphi(u, v) = (\bar{x}_1, \bar{x}_2, \bar{x}_3) + u\,(a_1, a_2, a_3) + v\,(b_1, b_2, b_3), \qquad u, v \in \mathbb{R}.$$

Ce même plan consiste en l'ensemble points $P = (x_1, x_2, x_3)$ tels que le vecteur $\overrightarrow{\bar{P}P}$ est orthogonal à $\vec{a} \wedge \vec{b}$. Ainsi,

$$(x_1, x_2, x_3) \in \pi \iff (\vec{a} \wedge \vec{b}) \cdot \overrightarrow{\bar{P}P} = 0 \iff \det \begin{pmatrix} (x - \bar{x})_1 & (x - \bar{x})_2 & (x - \bar{x})_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} = 0.$$

La dernière égalité est l'**équation cartesienne du plan** passant par $\bar{P}$, de vecteurs directeurs $\vec{a}$ et $\vec{b}$.

Exemple 6.2. La sphère centrée en O et de rayon R peut se paramétrer à l'aide de la longitude et de la co-latitude par l'application

$$(\phi, \theta) \mapsto (R \sin\theta \cos\phi, R \sin\theta \sin\phi, R \cos\theta), \qquad \phi \in [0, 2\pi[, \quad \theta \in [0, \pi].$$

Posons $\varphi_v = (\frac{\partial\varphi_1}{\partial u}, \frac{\partial\varphi_2}{\partial u}, \frac{\partial\varphi_3}{\partial u})$ et $\varphi_v = (\frac{\partial\varphi_1}{\partial v}, \frac{\partial\varphi_2}{\partial v}, \frac{\partial\varphi_3}{\partial v})$.

Définition 6.1. *On dit qu'une paramétrisation $\varphi\colon K \to \mathbb{R}^3$ d'une surface Σ est* régulière *si φ est de classe C^1, injective dans $\overset{\circ}{K}$, et, pour tout (u, v) à l'intérieur de K, les deux vecteurs de $\mathbb{R}^3$ $\varphi_u(u, v)$ et $\varphi_v(u, v)$ sont linéarement indépendants.*

Justifions la définition précédente. Fixons $v = v_0$ et faisons varier le paramètre u. L'application $u \mapsto \varphi(u, v_0)$ définit une courbe paramétrée dans $\mathbb{R}^3$, dont le support est contenu sur la surface Σ. De la même manière, si on fixe $u = u_0$ et on fait varier v, on obtient une autre courbe paramétrée $v \mapsto \varphi(u_0, v)$ sur la surface. Ces deux courbes admettent comme vecteurs tangents au point $P_0 = \varphi(u_0, v_0) \in \mathbb{R}^3$ respectivement les vecteurs $\vec{a} = \frac{\partial\varphi}{\partial u}(u_0, v_0)$ et $\vec{b} = \frac{\partial\varphi}{\partial v}(u_0, v_0)$. Sous l'hypothèse de paramétrisation régulière, ces deux vecteurs sont linéairement indépendants et donc ils engendrent un plan dans $\mathbb{R}^3$. Il s'agit du *plan tangent* à Σ au point P_0.

Le plan tangent à Σ au point P_0 a donc pour équation paramétrique

$$(u, v) \mapsto \varphi(u_0, v_0) + u\, \varphi_u(u_0, v_0) + v\, \varphi_v(u_0, v_0), \qquad u, v \in \mathbb{R}.$$

Le *verseur normal* à une surface Σ paramétrée par $(u, v) \mapsto \varphi(u, v)$, au point $\varphi(u, v)$ est alors donnée par la formule

$$\vec{\nu}(u_0, v_0) = \frac{(\varphi_u \wedge \varphi_v)(u_0, v_0)}{\|(\varphi_u \wedge \varphi_v)(u_0, v_0)\|}.$$

Exemple 6.3. Le cylindre Σ de de base le disque $\{(x, y)\colon x^2 + y^2 \leq R^2\}$ et hauteur infinie peut être paramétrée par

$$(\theta, z)\colon (\cos\theta, \sin\theta, z), \qquad \theta \in [0, 2\pi],\ z \in \mathbb{R}.$$

Le plan tangent au cylindre au point $(cos\theta_0, \sin\theta_0, z_0)$ est le plan paramétré par

$$(\theta, z) \mapsto (\cos\theta_0, \sin\theta_0, z_0) + \theta(-\sin\theta_0, \cos\theta_0, 0) + z(0, 0, 1).$$

L'équation cartesienne de ce plan est

$$0 = \det \begin{pmatrix} x - \cos\theta_0 & y - \sin\theta_0 & z - z_0 \\ -\sin\theta_0 & \cos\theta_0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = x\cos\theta_0 + y\sin\theta_0 - 1$$

Exemple 6.4. Le graphe d'une fonction de 2 variables f peut être paramétré par

$$(x, y) \mapsto (x, y, f(x, y)).$$

Le verseur normal au graphe est alors donné par la formule

$$\nu = \frac{1}{\sqrt{1 + |\nabla f|^2}}(-f_x, -f_y, 1).$$

6.3 Aire d'une surface paramétrée et intégrale de surface

Considérons la surface Σ paramétrée par $(u,v) \mapsto \varphi(u,v)$, où $(u,v) \in K$. Considérons un grillage de K, le subdivisant en petits rectangles d'aire $\mathrm{d}u\,\mathrm{d}v$, et soit R_0 l'un de ces rectangles, de sommets (u_0,v_0), $(u_0+\mathrm{d}u, v_0)$, $(u_0, v_0+\mathrm{d}v)$ et $(u_0+\mathrm{d}u, v_0+\mathrm{d}v)$. L'application φ envoie R_0 en une petite portion de la surface $\varphi(R_0)$. Mais en appliquant la formule de Taylor d'ordre 1 (en négligeant le reste),

$$\varphi(u_0+\mathrm{d}u, v_0) \simeq \varphi(u_0,v_0) + \varphi_u(u_0,v_0) \qquad \text{et} \qquad \varphi(u_0, v_0+\mathrm{d}v) \simeq \varphi(u_0,v_0) + \varphi_v(u_0,v_0).$$

Ainsi, la portion de surface $\varphi(R_0)$ est est approximativement le parallélogramme construit à l'aide de vecteurs $\varphi_u(u_0,v_0)$ et $\varphi_v(u_0,v_0)$, pointés en $\varphi(u_0,v_0)$. Ce parallélogramme est d'aire $\|(\varphi_u \wedge \varphi_v)(u_0,v_0)\|\,\mathrm{d}u\,\mathrm{d}v$. Ainsi la quantité

$$\mathrm{d}\sigma \overset{\text{déf}}{=} \|(\varphi_u \wedge \varphi_v)(u,v)\|\,\mathrm{d}u\,\mathrm{d}v$$

s'interpréte comme un "élément infinitesimal d'aire" sur la surface. Cela conduit naturellement à la définition suivante :

$$\mathrm{Aire}(\Sigma) = \int_K \|(\varphi_u \wedge \varphi_v)(u,v)\|\,\mathrm{d}u\,\mathrm{d}v.$$

Exemple 6.5. La formule précédente, appliquée à la sphère de rayon R, donne l'aire $A = 4\pi R^2$.

Définition 6.2. *Soit $f\colon \Sigma \to \mathbb{R}$ une fonction continue, définie sur une surface Σ régulière, paramétrée par $(u,v) \mapsto \varphi(u,v)$, où $(u,v) \in K$. La quantité*

$$\int_K f\,\mathrm{d}\sigma \overset{\text{déf}}{=} \int_K f(\varphi(u,v))\|(\varphi_u \wedge \varphi_v)(u,v)\|\,\mathrm{d}u\,\mathrm{d}v$$

est l'intégrale de surface de f sur Σ.

Pour les nappes paramétrées, tout comme pour les courbes, on peut introduire la notion de paramétrisations équivalentes. On montre alors que l'intégrale de surface est indépendant de la paramétrisation choisie. En particulier (en considérant la fonction constante $f=1$), l'aire d'une surface est indépendante de la paramétrisation choisie.

7 Le théorème des fonctions implicites

7.1 Le cas des fonctions de deux variables

Une courbe dans le plan est souvent donnée à l'aide d'une équation de la forme

$$F(x,y) = 0.$$

Le but de cette section est d'étudier les propriétés de telles courbes. À partir de l'équation $F(x,y)=0$ il est parfois possible d'exprimer une variable en fonction de l'autre, par exemple $y = f(x)$. Dans ce cas la courbe d'équation $F(x,y)=0$ est le graphe de la fonction f. Il peut être plus aisé d'exprimer x comme fonction de y : si l'on peut écrire $x = g(y)$ cela signifie que la courbe est le graphe de la fonction g.

Mais il y a des courbes qui ne sont pas (globalement) de graphes : c'est le cas, par exemple, du cercle centré en O de rayon 1, qui correspond aux solutions de l'équation $x^2+y^2-1=0$. Localement la situation est différente : si on fixe un point du cercle (x_0,y_0) il est possible de trouver un voisinage de ce point où le cercle est bien le graphe d'une fonction $y=f(x)$ et/ou $x = g(y)$: il s'agit de prendre la fonction $y = \sqrt{1-x^2}$ (si $y_0>0$) ou la fonction $y = -\sqrt{1-x^2}$ (si $y_0<0$) ; ou encore $x = \sqrt{1-y^2}$ (si $x_0>0$) ou $x = -\sqrt{1-y^2}$ (si $x_0<0$).

Étant donnée une courbe d'équation $F(x,y)=0$, et un point (x_0,y_0) de cette courbe, quand est-ce qu'on peut dire que la courbe est (au moins localement) le graphe d'une fonction $y=f(x)$? La réponse est fournie par le théorème ci-dessous :

Théorème 7.1 (des fonctions implicites pour les fonctions de 2 variables)**.** *Soit U un ouvert de $\mathbb{R}^2$ et $F\colon U\to\mathbb{R}$ une fonction de classe C^1. Soit $(x_0,y_0)\in U$ tel que*

$$F(x_0,y_0)=0, \qquad \frac{\partial F}{\partial y}(x_0,y_0)\neq 0.$$

Alors il existe un voisinage I de x_0, un voisinage J de y_0 et une fonction $f\colon I\to J$ tels que

$$\forall\,(x,y)\in I\times J:\quad F(x,y)=0 \iff y=f(x).$$

Dans ce cas la fonction $x\mapsto f(x)$ est de classe C^1 et

$$f'(x)=-\frac{F_x(x,f(x))}{F_y(x),f(x))}, \tag{7.1}$$

où $F_x=\frac{\partial F}{\partial x}$ et $F_y=\frac{\partial F}{\partial y}$.

Revenons au cas de l'équation $x^2+y^2-1=0$. Les hypothèse sont satisfaites en tout point (x_0,y_0) du cercle, sauf aux points $(-1,0)$ et $(1,0)$: au voisinage de ces deux points il n'est manifestement pas possible exprimer y comme fonction de x, mais il est possible exprimer x en fonction de y.

Remarque 7.1. En échangeant les rôles des variables x et y on trouve ceci : *Si $(x_0,y_0)\in U$ est un point tel que*

$$F(x_0,y_0)=0, \qquad \frac{\partial F}{\partial x}(x_0,y_0)\neq 0.$$

Alors il existe un voisinage I de x_0 et un voisinage J de y_0 et une fonction $g\colon J\to I$ tels que

$$\forall\,(x,y)\in I\times J:\quad F(x,y)=0 \iff x=g(x).$$

Dans ce cas la fonction $y\mapsto g(y)$ est de classe C^1 et $g'(y)=-\frac{F_y(g(y),y)}{F_x(g(y),y)}$.

Dém. du théorème. Nous pouvons supposer, par exemple $F_y(x_0,y_0)>0$. Par la continuité de F_y, on a $F_y(x,y)>0$ dans un rectangle $I\times J$. Posons $J=[y_0-h,y_0+h]$. La fonction $y\mapsto F(x_0,y)$ étant croissante dans J on a $F(x_0,y_0-h)<0$ et $F(x_0,y_0+h)>0$. Considérons les fonctios continues $x\mapsto F(x_0,y_0-h)$ et $x\mapsto F(x_0,y_0+h)$. Quitte à remplacer I par un intervalle centré en x_0 plus petit, pour $x\in I$ on a : $F(x,y_0-h)<0$ et $F(x,y_0+h)>0$. Par le théorème des valeures intermédiaires, appliqué à la fonction croissante $y\mapsto F(x,y)$ (avec x fixé), il existe un et un seul $y=f(x)$ tel que $F(x,f(x))=0$. La fonction f est ainsi construite.

Démontrons que f est de classe C^1 : Prenons $x,x_1\in I$. En appliquant le théorème des accroissements finis à la fonction

$$\Phi(t)=F((1-t)x_1+tx,(1-t)f(x_1)-tf(x)), \qquad 0\le t\le 1,$$

on trouve

$$\exists\tau\in[0,1]:\quad \Phi(1)-\Phi(0)=\Phi'(\tau).$$

Considérons le point $P_\tau=((1-\tau)x_1+\tau x,(1-\tau)f(x_1)-\tau f(x))$, qui est un point sur le ségment de $(x,f(x))$ à $(x_1,f(x_1))$. L'égalité précédente se réécrit

$$F(x,f(x))-F(x_1,f(x_1))=\Phi'(\tau)=(x-x_1)F_x(P_\tau)+(f(x)-f(x_1))F_y(P_\tau).$$

D'autre part, $F(x, f(x)) = F(x_1, f(x_1)) = 0$. Donc,

$$\frac{f(x) - f(x_1)}{x - x_1} = -\frac{F_x(P_\tau)}{F_y(P_\tau)}.$$

Si l'on prend $x_1 \to x$, on a $P_\tau \to (x, f(x))$. Mais comme F est de classe C^1, on voit que f est dérivable et

$$f'(x) = -\frac{F_x(x, f(x))}{F_y(x, f(x))}.$$

L'expression à droite étant une fonction continue de x, on trouve que f est de classe C^1. □

La démonstration ci-dessus montre que si f est de classe C^k, alors f est de classe C^k.

Corollaire 7.2. *Si F est une fonction de classe C^1 au voisinage de $(x_0, y_0) \in \mathbb{R}^2$ et $\nabla F(x_0, y_0) \neq (0, 0)$, alors la courbe d'équation $F(x, y) = 0$ possède une droite tangente au point (x_0, y_0) et $\nu = \frac{\nabla F(x_0,y_0)}{\|\nabla F(x_0,y_0)\|}$ est le verseur orthogonal la courbe d'équation $F(x, y) = 0$, au point (x_0, y_0).*

Dém. En effet, au voisinage de (x_0, y_0), la courbe est le graphe d'une fonction dérivable de la forme $y = f(x)$, ou éventuellement de la forme $x = g(y)$. Traitons, par exemple, le premier cas. Un point (x, y) appartient à la droite tangente si et seulement si $y - y_0 = f'(x_0)(x - x_0)$, c'est à dire si et seulement si (d'après (7.1)) $(x - x_0)F_x(x_0, y_0) + (y - y_0)F_y(x_0, y_0) = 0$. Cela dit précisément que le vecteur $\nabla F(x_0, y_0)$ est orthogonale à $(x - x_0, y - y_0)$, c'est à dire orthogonale à la droite tangente. □

Corollaire 7.3. *Si $g \colon U \to \mathbb{R}$ est une fonction de classe C^1, où U est un ouvert de $\mathbb{R}^2$, alors en tout point $(x_0, y_0) \in U$, $\nabla g(x_0, y_0)$ est orthogonal à la ligne de niveau de g passant par (x_0, y_0).*

Dém. Si $\nabla g(x_0, y_0) = 0$ il n'y a rien à démontrer (le vecteur nul est orthogonal à tous les vecteurs). Si $\nabla g(x_0, y_0) \neq 0$, considérons la fonction $F(x, y) = g(x, y) - c$ où $c = g(x_0, y_0)$. On a $\nabla F(x_0, y_0) = \nabla g(x_0, y_0)$ et la ligne de niveau de g passant par (x_0, y_0) est l'ensemble d'équation $F(x, y) = 0$. Le résultat est alors une conséquence immédiate du corollaire précédent. □

7.2 Le cas des fonctions de 3 variables

Si $F \colon U \to \mathbb{R}$ est une fonction de 3 variables de classe C^1, définie sur un ouvert $U \subset \mathbb{R}^3$ et telle que, en un point $(x_0, y_0, z_0) \in U$ on a

$$F(x_0, y_0, z_0) = 0, \qquad F_z(x_0, y_0, z_0) \neq 0,$$

alors il est possible, localement, d'exprimer z comme fonction de x et de y. Autrement dit, il existe un voisinage V de (x_0, y_0), un voisinage J de z_0 et une fonction, de classe C^1, $f \colon V \to J$ tels que

$$\forall\, (x, y, z) \in V \times J \colon \quad F(x, y, z) = 0 \iff z = f(x, y)$$

et dans ce cas

$$f_x(x, y) = -\frac{F_x(x, y, f(x, y))}{F_z(x, y, f(x, y))}, \qquad f_y(x, y) = -\frac{F_y(x, y, f(x, y))}{F_z(x, y, f(x, y))}.$$

La démonstration est semblable à celle du théorème précédent et omise ici.

Dans ce cas la surface représentative du graphe de la fonction $z = f(x, y)$ possède, au point $(x, y, f(x, y))$, un verseur normal donné par (voir l'exemple 6.4)

$$\nu = \frac{(-f_x, -f_y, 1)}{\sqrt{1 + \|\nabla f\|^2}},$$

où toutes les dérivées sont calculées au point (x, y). Supposons maintenant, par exemple, que $F_z(x_0, y_0, z_0) > 0$. Compte tenu des formules précédentes pour f_x et f_y, on trouve

$$\nu = \left(\frac{F_x}{\|\nabla F\|}, \frac{F_y}{\|\nabla F\|}, \frac{F_z}{\|\nabla F\|} \right) = \frac{\nabla F}{\|\nabla F\|}. \tag{7.2}$$

Exemple 7.2. La fonction $F(x, y, z) = x^2 e^z + z e^y + y^2$ s'annule à l'origine. De plus $F_z(0, 0, 0) = 1$. Par conséquent, l'équation $F(x, y, z) = 0$ est, au voisinage de l'origine, le graphe d'une fonction $z = f(x, y)$. Cette fonction f est solution du système différentiel $f_x(x, y) = -\frac{2xe^{f(x,y)}}{x^2 e^{f(x,y)} + e^y}$ et $f_y(x, y) = -\frac{f(x,y)e^y + 2f(x,y)}{x^2 e^{f(x,y)} + e^y}$. De plus, on a $f(0, 0) = 0$ et $f_x(0, 0) = 0$, $f_y(0, 0) = 0$.

Observons que la formule (7.2) implique que, pour une fonction de 3 variables, le vecteur gradient est orthogonal, en tout point, à la surface de niveau passant par ce point. Cette affirmation se démontre comme dans le cas des fonctions de 2 variables (voir le corollaire 7.3). En généralisant à n variables, il n'est pas difficile de se convaincre alors de la validité de la proposition suivante :

Proposition 7.4. *Soit Σ un ensemble de niveau d'une fonction g de n nariables de classe C^1, et $x_0 \in \Sigma$ un point tel que $g(x_0) = 0$, $\nabla g(x_0) \neq 0$. Alors le vecteur $\nabla g(x_0)$ est orthogonal à la surface Σ au point x_0.*

Traitons maintenant le cas de deux équations de la forme

$$\begin{cases} F(x, y, z) = 0 \\ G(x, y, z) = 0, \end{cases} \tag{7.3}$$

où les fonctions F et G sont de classe C^1 au voisinage de $P_0 = (x_0, y_0, z_0)$. On se propose d'exprimer deux variables en fonction de la troisième, par exemple $y = f(x)$ et $z = g(x)$. Pour cela, commençons par supposer

$$F(P_0) = G(P_0) = 0, \qquad F_z(P_0) \neq 0.$$

En appliquant ce qui a été vu au début de cette section on peut écrire, au voisinage de P_0,

$$F(x, y, z) = 0 \iff z = \varphi(x, y).$$

Considérons maintenant la fonction

$$\Phi(x, y) = G(x, y, \varphi(x, y)).$$

Au voisinage de P_0 on a

$$\begin{cases} F(x, y, z) = 0 \\ G(x, y, z) = 0 \end{cases} \iff \begin{cases} z = \varphi(x, y) \\ G(x, y, \varphi(x, y)) = 0 \end{cases} \iff \begin{cases} z = \varphi(x, y) \\ \Phi(x, y) = 0 \end{cases}.$$

Observons que

$$\Phi_y = G_y + G_z \varphi_y = G_y - G_z \frac{F_y}{F_z}. \tag{7.4}$$

Maintenant, supposons que $0 \neq \Phi_y(x_0, y_0)$.

Dans ce cas nous pouvons exprimer y en fonction de x et écrire, au voisinage de (x_0, y_0),

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

Cela donne alors, au voisinage de P_0,

$$\begin{cases} F(x,y,z) = 0 \\ G(x,y,z) = 0 \end{cases} \iff y = f(x) \quad \text{et} \quad z = \varphi(x, f(x)) = g(x).$$

Pour arriver à cette conclusion nous avons eu besoin de supposer :

$$F(P_0) = G(P_0) = 0, \qquad F_z(P_0) \neq 0 \quad \text{et} \quad \Phi_y(x_0, y_0) \neq 0. \tag{7.5}$$

Mais la formule précédemment donnée pour Φ_y montre que

$$\Phi_y(x_0, y_0) \neq 0 \iff \det \frac{\partial(F,G)}{\partial(y,z)} \neq 0,$$

où

$$\frac{\partial(F,G)}{\partial(y,z)} = \begin{pmatrix} F_y & F_z \\ G_y & G_z \end{pmatrix}.$$

Observons par allieurs que lorsque $\det \frac{\partial(F,G)}{\partial(y,z)} \neq 0$, on ne peut avoir simultanément $F_z(P_0) = 0$ et $G_z(P_0) = 0$. Or, on a déjà vu que si $F_z(P_0) \neq 0$ on peut exprimer les solutions du système (7.3) par les relations $y = f(x)$ et $z = g(x)$. Si on revanche $F_z(P_0) = 0$ on doit avoir $G_z(P_0) \neq 0$ et en interchageant les rôles de F et G on parvient à la même conclusion. Autrement dit, dans (7.5), la condition $F_z(P_0) \neq 0$ est superflue.

Les discussions précédentes nous conduisent à énoncer le théorème suivant.

Théorème 7.5 (Des fonctions implicites – 3 variables)**.** *Soit F et G deux fonctions de classe C^1 au voisinage d'un point $P_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$, tel que $F(P_0) = G(P_0) = 0$. Si la matrice*

$$\frac{\partial(F,G)}{\partial(x,y,z)} = \begin{pmatrix} F_x(P_0) & F_y(P_0) & F_z(P_0) \\ G_x(P_0) & G_y(P_0) & G_z(P_0) \end{pmatrix}.$$

est de rang 2, alors il existe un voisinage de P_0 où le système (7.3) *définit deux variables en fonction de la troisième. L'ensemble des solutions de* (7.3) *dans ce voisinage est alors le support d'une courbe régulière de $\mathbb{R}^3$.**

Rappelons que les lignes d'une matrice 2×3 de rang 2 sont linéairement indépendantes. Sous les hypothèses du théorème précédent, on voit que les vecteurs $\nabla F(P_0)$ et $\nabla G(P_0)$ ne sont pas co-linéaires. Autrement dit, $\nabla F(P_0) \wedge \nabla G(P_0) \neq 0$. Mais le vecteur $\nabla F(P_0)$ est orthogonal en P_0 à la surface de niveau $F(x,y,z) = 0$. Et le vecteur $\nabla G(P_0)$ est orthogonal en P_0 à la surface de niveau $G(x,y,z) = 0$. L'hypothèse que la matrice est de rang 2 traduit alors le fait que ces deux surfaces n'ont pas le même plan tangent. La conclusion du théorème est que l'intersection des ces deux surfaces est bien une courbe.

7.3 Le théorème de l'inversibilité locale

Si U et V sont deux ouverts de $\mathbb{R}^n$ et $x_0 \in U$. On dit que f est un difféomorphisme local en x_0 s'il existe une voisinage ouvert W de x_0 tel que $f \colon W \to F(W)$ soit un C^1-difféomorphisme, c'est-à-dire une application bijective de classe C^1 telle que l'application inverse est de classe C^1.

*Par exemple, la courbe paramétrée par $x = x$, $y = f(x)$, $z = g(x)$ si les deux dernière colonnes de la matrice sont linéarement indépendentes. Si le deux premières colonnes sont linéairement indépendantes on pourra expliciter x et y en fonction de z. Si la première et dernière colonne sont linéairement indépendantes, on écrira x et z comme fonction de y.

Pour qu'une fonction f soit un difféomorphisme local en x_0 il est necessaire que la matrice jacobienne en x_0 soit inversible (c'est à dire de déterminant non nul, s'agissant d'une matrice carrée) En effet, on nous avons déjà vu que, en posant $y_0 = f(x_0)$,

$$J_{f^{-1}}(y_0) = \Big(J_f(x_0)\Big)^{-1}.$$

Le théorème suivant affirme que cette condition est aussi suffisante.

Exemple 7.3. La fonction $f\colon \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^3$ est bijective et de classe C^1. Mais cette fonction n'est pas un difféomorphisme local en $x = 0$. En effet, la matrice jacobienne est simplement le singleton $f'(0)$ et l'on a $f'(0) = 0$. Observons que l'application inverse est $f^{-1}(y) = y^{1/3}$ et f^{-1} n'est pas de classe C^1 au voisinage de $y = 0$.

Théorème 7.6 (de l'inversibilité locale)**.** *Soient U et V deux ouverts de $\mathbb{R}^n$, $x_0 \in U$ et $f\colon U \to V$. Si $\det\big(J_f(x_0)\big) \neq 0$, alors f est un difféomorphisme local en x_0.*

Donnons juste l'idée de la démonstration pour les fonctions de deux variables : La formule de Taylor au premier ordre affirme que $f(x) - f(x_0)$ est approximativement égale à la différentielle $\mathrm{d}f_{x_0}$ évalué en $(x - x_0)$, ou encore au produit de la matrice Jacobienne $J_{x_0}(f)$ par le vecteur $x - x_0$. Ainsi, et en négligeant le reste dans la formule de Taylor, $f(x) \simeq f(x_0) + \mathrm{d}f_{x_0}(x - x_0) = f(x_0) + J_f(x_0)(x - x_0)$. Mais le terme à droite est inversible si et seulement si $\det(J_f(x_0) \neq 0$. On s'attent alors que dans un voisinage suffisament petit de x_0 on puisse inverser f et que $f^{-1}(y) \simeq y_0 + (J_f(x_0))^{-1}(y - y_0)$.

La démonstration rigoureuse repose sur le théorème des fonctions implicites. Une démonstration alternative fait appel au théorème des contractions.

7.4 Extrema liés. Optimisation sous contrainte

Nous avons déjà étudié les problèmes d'optimisation de type

$$\min\{f(x)\colon x \in U\} \qquad \text{et} \qquad \max\{f(x)\colon x \in U\},$$

où l'ensemble $U \subset \mathbb{R}^n$ était un ouvert.

Dans cette section on s'intéresse aux problèmes analogues d'optimisation sur des ensembles fermés, par exemple,

$$\min\{f(x)\colon x \in \Sigma\}, \qquad \max\{f(x)\colon x \in \Sigma\},$$

où Σ est un fermé de $\mathbb{R}^n$.

Nous considérons le cas particulier important où Σ est une surface de niveau d'une fonction continue $g\colon \mathbb{R}^n \to \mathbb{R}$ (ou une ligne de niveau si $n = 2$). Sans perte de généralité on pourra alors supposer que

$$\Sigma = \{x \in \mathbb{R}^n \colon g(x) = 0\}.$$

Le calcul de

$$\min\{f(x)\colon g(x) = 0\}, \qquad \text{et} \qquad \min\{f(x)\colon g(x) = 0\}, \tag{P}$$

sont deux “problèmes d'optimisation avec contrainte égalité”. Pour simplifier la présentation nous supposerons que g est une fonction C^1. Commençons par établir une propriété géométrique importante de l'ensemble Σ.

La condition $\nabla g(x_0) \neq 0$ exprime qu'en x_0 la contrainte *n'est pas dégénérée.* Voici un exemple de contrainte dégénéré : $xy = 0$ est une contrainte dégénéré à l'origine (observer que cette contrainte n'est pas une courbe mais plutôt l'intersection de deux courbes) à l'origine.

S'il n'y a avait pas la contrainte $g(x) = 0$, les solutions de ces problèmes d'optimisation seraient à chercher parmi les points stationnaires de la fonction f. Dans le cas d'une optimisation avec contrainte, comme c'est le cas pour les problèmes (P), la situation est différente :

Théorème 7.7 (des multiplicateurs de Lagrange). *Soient f et g deux fonctions de classe C^1 au voisinage de $P_0 \in \mathbb{R}^n$. Si P_0 est un point tel que $g(P_0) = 0$, $\nabla g(P_0) \neq 0$ et P_0 est un minimiseur ou un maximiseur pour $f(x)$ sous la contrainte $g(x) = 0$, c'est à dire*

$$f(P_0) = \min\{f(x)\colon g(x) = 0\}, \quad \textit{ou} \quad f(P_0) = \max\{f(x)\colon g(x) = 0\},$$

alors $\nabla f(P_0)$ est parallèle à $\nabla g(P_0)$. Autrement dit,

$$\exists\, \lambda_0 \in \mathbb{R} \textit{ tel que } \nabla f(P_0) + \lambda_0 \nabla g(P_0) = 0.$$

On appelle le réel λ_0 le multiplicateur de Lagrange *associé au problème d'optimisation.*

Remarque 7.4. Dans le cas $\nabla f(P_0) \neq 0$, et pour les fonctions de 2 variables, la conclusion s'interprète ainsi : les lignes de niveau de f (la fonction à optimiser) et de g (la fonction qui donne la contrainte) tangentes au point P_0. En effet, $\nabla f(P_0)$ et $\nabla g(P_0)$ sont, respectectivement, orthogonaux à leurs lignes de niveau. De même, pour les fonctions de trois variables, les surfaces de niveau de f et de g sont tangentes ne P_0.

Dém. Nous donnons la démonstration uniquement pour les fonctions de 2 ou 3 variables.

- Le cas des fonction de deux variables. Soit $(x_0, y_0) \in \mathbb{R}^2$ un point où $g(x_0, y_0) = 0$ et $\nabla g(x_0, y_0) \neq 0$. Supposons, par exemple $\frac{\partial g}{\partial y}(x_0, y_0) \neq 0$ (si $\frac{\partial g}{\partial y}(x_0, y_0) = 0$, alors, par l'hypothèse sur g, $\frac{\partial g}{\partial x}(x_0, y_0) \neq 0$ et l'on raisonne en intervertissant les variables x et y). Par le théorème des fonctions implicites, au voisinage de (x_0, y_0) on a $g(x, y) = 0 \iff y = \phi(x)$, où $\phi\colon I \to J$ et I et J sont des intervalles de $\mathbb{R}$, voisinages de x_0 et y_0 respectivement. En particulier, $\phi(x_0) = y_0$. Si $f(x_0, y_0) = \min\{f(x, y)\colon g(x, y) = 0\}$, ou si $f(x_0, y_0) = \max\{f(x, y)\colon g(x, y) = 0\}$, alors la fonction $\Psi(x) = f(x, \phi(x))$ possède respectivement un minimum ou un maximum local en x_0. Dans les deux cas, $\Psi'(x_0) = 0$ par le principe de Fermat. Mais alors,

  $$0 = \Psi'(x_0) = f_x(x_0, \phi(x_0)) + f_y(x_0, \phi(x_0))\phi'(x_0) = \nabla f(x_0, y_0) \cdot (1, \phi'(x_0)).$$

  Le vecteur $\nabla f(x_0, y_0)$ est donc orthogonal au vecteur $(1, \phi'(x_0))$. Mais le vecteur $(1, \phi'(x_0))$ est tangent en (x_0, y_0) au graphe de la fonction ϕ, c'est-à-dire à la courbe $\{(x, y)\colon y = \phi(x)\} = \{(x, y)\colon g(x, y) = 0\}$. Ainsi, $\nabla f(x_0, y_0)$ est orthogonal à la courbe $\{(x, y)\colon g(x, y) = 0\}$. Mais $\nabla g(x_0, y_0)$ est lui même orthogonale à cette courbe par la proposition 7.4. Ainsi, $\nabla f(x_0, y_0)$ et $\nabla g(x_0, y_0)$ sont deux vecteurs parallèles.

- Le cas des fonctions de 3 variables. Soit $P_0 = (x_0, y_0, z_0)$. L'hypothèse sur g est $\nabla g(P_0) \neq 0$. Supposons, par exemple, $g_z(P_0) \neq 0$. On a $g(x, y, z) = 0$ au voisinage de P_0 si et seulement si $z = \phi(x, y)$ dans ce voisinage, où ϕ est une fonction de classe C^1. La fonction $\Psi(x, y) := f(x, y, \phi(x, y))$ possède alors un extremum libre en (x_0, y_0). Mais (x_0, y_0) est alors un point stationnaire pour Ψ et $\nabla\Psi(x_0, y_0) = 0$. En appliquant à ϕ des formules analogues (7.4) on trouve

  $$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \nabla\Psi(x_0, y_0) = \begin{pmatrix} (f_x - f_z g_x/g_z)(P_0) \\ (f_y - f_z g_y/g_z)(P_0) \end{pmatrix} = \frac{1}{g_z(P_0)} \begin{pmatrix} (f_x g_z - f_z g_x)(P_0) \\ (f_x g_z - f_z g_x)(P_0) \end{pmatrix}.$$

  D'après la formule du produit vectoriel (6.1), cela implique que la première et la deuxième composante du vecteur $(\nabla f \wedge \nabla g)(P_0)$ s'annulent. Maintenant, si $g_x(P_0) = 0$ et $g_y(P_0) = 0$ alors la troisième composante de $(\nabla f \wedge \nabla g)(P_0)$ sera nulle également. Sinon, $g_x(P_0) \neq 0$ ou $g_y(P_0) \neq 0$ et on parvient à la même conclusion en raisonnant comme ci-dessus mais en échangeant les rôles des variables x, y et z. En conclusion $\nabla f \wedge \nabla g$ s'annule au point P_0, ce qui se produit précisément lorsque $\nabla f(P_0)$ et $\nabla g(P_0)$ sont parallèles.

□

Le théorème précédent admet la reformulation suivante : introduisons la Lagrangienne du problème d'optimisation, qui est la fonction de $n+1$ variables

$$\mathcal{L}(x,\lambda) = f(x) + \lambda g(x).$$

On a

$$\begin{cases} \nabla f(P_0) + \lambda_0 \nabla g(P_0) = 0 \\ g(P_0) = 0 \end{cases} \iff \nabla \mathcal{L}(x_0, \lambda_0) = 0,$$

où $\nabla_{x,\lambda}\mathcal{L}$ est le vecteur de $n+1$ composantes $\frac{\partial \mathcal{L}}{\partial x_1}, \dots \frac{\partial \mathcal{L}}{\partial x_n}$ et $\frac{\partial \mathcal{L}}{\partial \lambda}$.

Autrement dit, les solutions $P_0 \in \mathbb{R}^n$ du problème d'optimisation sous contrainte

$$\min\{f(x) \colon g(x) = 0\}, \qquad \text{et} \qquad \min\{f(x) \colon g(x) = 0\}, \tag{P}$$

où f et g sont de classe C^1 au voisinage de P_0 et à valeur dans $\mathbb{R}$, sont à chercher parmi les points P_0 tels qu'il existe $\lambda_0 \in \mathbb{R}$ tel que $(P_0, \lambda_0) \in \mathbb{R}^{n+1}$ est un point stationnaire de la Lagrangienne (vue comme fonction de $n+1$-variables). Les solutions du problème (P) sont aussi à chercher parmi les points où la contrainte $g(x) = 0$ est dégénérée. En conclusion, il s'agit alors de chercher les solutions de

$$\nabla_{x,\lambda}\, \mathcal{L}(x,\lambda) = 0, \qquad \text{ou} \qquad \begin{cases} \nabla\, g(x) = 0 \\ g(x) = 0. \end{cases}$$

Observer que ce théorème ne fournit qu'une condition nécessaire : ce théorème s'avère très utile pour trouver les points P_0 qui sont les bons candidats à être les solutions des problèmes d'optimisation avec contrainte. Mais sous les hypothèses du théorème, il peut arriver que (P_0, λ_0) soit un point stationnaire de $\mathcal{L}$, ou que $g(P_0) = 0$ et $\nabla g(P_0) = 0$, sans que P_0 soit ni un minimum, ni un maximum du problème d'optimisation.

Dispose-t-on de conditions suffisantes pour l'existence d'extrema avec contraintes ? La réponse est affirmative, mais n'insistons pas sur ce point. Bien souvent, on applique le théorème de Weierstrass pour démontrer que le minimum et/ou le maximum existent. Cela est possible notamment si la contrainte Σ définit un ensemble compact. *

Exemple 7.5. Trouvons le maximum et le minimum de la fonction $f(x,y,z) = x + y - \sqrt{6}z$ sur la sphère Σ de centre O et de rayon 1. Tout d'abord, f est une fonction continue et $\Sigma = \{(x,y,z) \colon x^2 + y^2 + z^2 - 1 = 0\}$ est un compact de $\mathbb{R}^3$. Donc, par le théorème de Weierstrass,

---

*Rappelons aussi la variante suivante du théorème de Weierstrass, qui s'applique (parfois) quand Σ n'est pas compact :

Théorème 7.8 (Weierstrass - variante)**.** *Soit $f \colon \Sigma \subset \mathbb{R}^n \to \mathbb{R}$ une fonction continue. Si $\bar{x} \in \Sigma$ et si l'*ensemble de sous-niveau $K = \{x \in \Sigma \colon f(x) \le f(\bar{x})\}$ *est compact, alors le problème de minimisation* $\min_{x \in \Sigma} f(x)$ *possède une solution. Autrement dit, il existe $x^* \in D$ tel que $f(x^*) = \min_{x \in \Sigma} f(x)$.*

Le cas typique d'application est celui d'une fonction $f \colon \mathbb{R}^n \to \mathbb{R}$ telle que $\lim_{\|x\| \to +\infty} f(x) = +\infty$. Une telle fonction a tous les ensembles de sous-niveau bornés (pourquoi ?). Si de plus f est continue, ses ensembles de sous-niveau sont fermés (pourquoi ?) et donc compacts. La fonction possède alors un minimum absolu.

Bien entendu on peut établir un théorème analogue pour l'existence d'un maximum absolu : il s'agit cette fois-ci de supposer que l'ensemble de "sur-niveau" $\{x \colon f(x) \ge f(\bar{x})\}$ est borné. Le cas typique d'application est celui d'une fonction continue telle que $\lim_{\|x\| \to +\infty} f(x) = -\infty$.

les problèmes de minimisation et de maximisation posés possèdent bien des solutions. Ici, $g(x,y,z) = x^2+y^2+z^2-1$. Observons que la contrainte $g(x,y,z)=0$ n'est jamais dégénérée, puisque le seul point où ∇g s'annule est l'origine, mais l'origine ne vérifie pas la contrainte. Ainsi, les points de minimum et maximum sont à chercher parmi les points stationnaires de la Lagrangienne. La Lagrangienne du système est la fonction $\mathcal{L}(x,y,z,\lambda) = x+y-\sqrt{6}z+\lambda(x^2+y^2+z^2-1)$. Ses points stationnaires sont les solutions du système

$$\begin{cases} 1+2\lambda x = 0 \\ 1+2\lambda y = 0 \\ -\sqrt{6}+2\lambda z = 0 \\ x^2+y^2+z^2-1=0. \end{cases}$$

En exprimant x,y,z en fonction de λ, la quatrième équation donne $\lambda = \pm\sqrt{2}$ et ensuite $(x_0,y_0,z_0,\lambda_0) = \pm(1/2\sqrt{2}, 1/2\sqrt{2}, -\sqrt{6}/2\sqrt{2}, \sqrt{2})$. Les solutions des problèmes de minimisations et maximisation (dont l'existence a été établie avant via le théorème de Weierstrass) sont alors à chercher parmi les points $P = (1/2\sqrt{2}, 1/2\sqrt{2}, -\sqrt{6}/2\sqrt{2})$ et $Q = (-1/2\sqrt{2}, -1/2\sqrt{2}, +\sqrt{6}/2\sqrt{2})$. Mais un calcul direct montre que $f(P) > f(Q)$. Ceci permet de dire que f atteint sur la sphère son maximum en P et son minimum en Q.

Même si le théorème des multiplicateurs de Lagrange est un outil puissant, il convient parfois ne pas l'utiliser, notamment s'il est possible d'éliminer la contrainte en exprimant une variable en fonction des autres :

Exemple 7.6. Optimiser $x^2+y^2+z^2$ sous la contrainte $x+2y+3z=1$. Interpreter le résultat. Solution. La lagrangienne du système est $\mathcal{L}(x,y,z,\lambda) = x^2+y^2+z^2+\lambda(x+2y+3z-1)$. Son unique point stationnaire, obtenu avec le multiplicateur $\lambda = -1/7$, est $(1/14, 1/7, 3/14)$. La contrainte n'étant pas compacte le théorème de Weierstrass ne s'applique pas. Appliquons la variante du théorème de Weierstrass. Choisissons un point arbitraire sur la contrainte, par exemple $(0,0,1/3)$. L'ensemble de sous-niveau $\{(x,y,z)\colon x^2+y^2+z^2 \le 1/9\}$ est manifestement compact (c'est la boule fermée de rayon 1 et centre 1/3). Le problème posé possède alors un minimum absolu. Le point $(1/14, 1/7, 3/14)$ est donc l'optimum demandé et il s'agit d'un minimum. Il s'agit du point du plan d'équation $x+2y+3z=1$ à distance minimale de l'origine.

Exemple 7.7. On cherche à construire un caisson de 20m^3. Le matériau pour le fond coûte 3 euros/ m^2, pour le couvercle 2 euros/m^2 et pour les côtés 1 euro/m^2. Quel est le caisson le moins cher ? Et le plus cher ?
Réponse :

- *Modélisation :* Notons x(=longueur), y(=largeur), z(=hauteur) les mesures du casson en mètres. Le coût de construction est $C(x,y,z) = 3xy+2xy+2(xz+yz) = 5xy+2xz+2yz$. Il s'agit de résoudre les problèmes de minimisation et maximisation pour C "avec contrainte" : $\min\{C(x,y,z)\colon xyz=20\}$ et $\max\{C(x,y,z)\colon xyz=20\}$.

- *Élimination de la contrainte et recherche des points stationnaires :* On pourrait introduire la Lagrangienne et en déterminer les points stationnaires (cela conduit à resoudre un système de 4 équations et 4 inconnues). Mais il est plus aisé de poser $z = 20/(xy)$ et d'étudier la fonction

  $$f(x,y) = C(x,y,\tfrac{20}{xy}) = 5xy + 40(\tfrac{1}{y}+\tfrac{1}{x}), \qquad x>0, y>0.$$

  La fonction f étant définie sur un ouvert, il s'agit d'en trouver les points stationnaires : Cela conduit à résoudre le système (de deux équations et 2 inconnues) $\nabla f(x,y) = 0$. Ce système possède une seule solution pour $x>0$ et $y>0$. Elle est donnée par $x=y=2$ (et donc $z=5$).

- *Synthèse :* On construit donc un caisson de mesures $2 \times 2 \times 5$. Ce choix correspond-t-il au caisson de coût minimum, maximum, ou ni l'un ni l'autre ? Le problème de minimisation est-il *bien posé* ? Et celui de maximisation ? Pour répondre à ces questions appliquons la variante du théorème de Weierstrass avec $(\bar{x}, \bar{y}) = (1,1)$ (ce choix est arbitraire). Observons que l'ensemble de sous-niveau $K = \{(x,y)\colon f(x,y) \leq f(1,1)\} = 5xy + 40(\frac{1}{y} + \frac{1}{x}) \leq 85\}$ est compact (en effet, il est manifestement fermé et il est borné, puisque $x \geq 40/85$, $y \geq 40/85$ et $5xy \leq 85 \Rightarrow x \leq 17 \cdot 85/40$ et $y \leq 17 \cdot 85/40$). Mais alors le problème de minimisation de l'exemple 7.7 est bien posé, c'est-à dire que le caisson de coût minimum existe : c'est bien le caisson de mesures $2 \times 2 \times 5$ trouvé avant. Le problème de maximisation est mal posé : le caisson de coût maximum n'existe pas. On le voit en observant que des cassons de mesures $x, x, 20/x^2$ ont un coût qui tend à l'infini si $x \to +\infty$.

Optimisation sous plusieurs contraintes. Le théorème des multiplicateurs de Lagrange se généralise au cas où il y a plusieurs contraintes à satisfaire. Considérons par exemple le problème

$$\min\{f(x)\colon g_1(x) = 0,\ g_2(x) = 0\}, \quad \text{ou} \quad \max\{f(x)\colon g_1(x) = 0,\ g_2(x) = 0\} \quad (x \in \mathbb{R}^n,\ n \geq 3).$$

où f, g_1 et g_2 sont des fonctions de classe C^1. On introduit dans ce cas la Lagrangienne de $(n+2)$-variables $\mathcal{L}(x, \lambda_1, \lambda_2)$, où

$$\mathcal{L}(x, \lambda_1, \lambda_2) = f(x) + \lambda_1 g_1(x) + \lambda_2 g_2(x).$$

On peut démontrer que les points x_0 de minimum ou maximum de f sous les contraintes $g_1(x) = g_2(x) = 0$ sont à chercher parmi les points $x \in \mathbb{R}^n$ tels que $(x, \lambda_1, \lambda_2)$ est un point stationnaires de la Lagrangienne. Ainsi, pour trouver ces points (ou du moins des points candidats à être des solutions du problème d'optimisation), on commence par trouver les solutions du système de $n+2$ équations

$$\nabla_{x,\lambda_1,\lambda_2} \mathcal{L}(x, \lambda_1, \lambda_2) = 0. \tag{7.6a}$$

Mais comme on l'a vu avant, un autre cas de figure est possible : celui où la contrainte $g_1(x) = g_2(x) = 0$ est dégénérée : ainsi les solutions du problèmes d'optimisation sont aussi à chercher parmi les éventuelles solutions des systèmes

$$\begin{cases} \nabla g_1(x) = 0 \\ g_1(x) = g_2(x) = 0 \end{cases} \qquad \text{ou} \qquad \begin{cases} \nabla g_2(x) = 0 \\ g_1(x) = g_2(x) = 0 \end{cases} \tag{7.6b}$$

Dans la plupart des cas la contrainte ne sera pas dégénérée et les systèmes (7.6b) n'ont pas de solution.

Bien entendu, ces considérations se généralisent à un nombre arbitraire de contraintes.