

Introduction à la géométrie différentielle et aux équations différentielles

Irène Waldspurger

waldspurger@ceremade.dauphine.fr

25 mai 2023

Remerciements

La section finale, sur le pendule, est largement inspirée de notes écrites par Jacques Fejoz, que je remercie beaucoup.

Table des matières

1	Rappels de calcul différentiel	7
1.1	Définition de la différentiabilité	7
1.2	Dérivées partielles	8
1.3	Théorème d'inversion locale	12
1.4	Immersion et submersions	16
1.5	Inégalité des accroissements finis	19
2	Sous-variétés de \mathbb{R}^n	21
2.1	Définition	21
2.2	Exemples et contre-exemples	27
2.2.1	Sphère	27
2.2.2	Produit de sous-variétés	28
2.2.3	$O_n(\mathbb{R})$	29
2.2.4	Solutions d'équations et images de fonctions	31
2.2.5	Deux contre-exemples	32
2.3	Espace tangent	35
2.3.1	Définition	35
2.3.2	Exemples	41
2.4	Applications entre sous-variétés	46
3	Géométrie riemannienne	55
3.1	Sous-variétés de dimension 1	56
3.1.1	Arcs paramétrés	56
3.1.2	Démonstration du théorème 3.4	59
3.1.3	Longueur et abscisse curviligne	72
3.2	Sous-variétés de dimension générale	75
3.2.1	Distance et géodésiques	75
3.2.2	Distance et géodésiques sur la sphère	82

3.2.3	Isométries	86
3.2.4	Exemples d'isométries	91
4	Équations différentielles : existence et unicité	97
4.1	Théorème de Cauchy-Lipschitz	97
4.2	Solutions maximales	101
4.3	Théorème des bouts	105
4.4	Régularité en la condition initiale	110
5	Résolution explicite dans des cas particuliers	119
5.1	Équations scalaires autonomes	119
5.2	Équations linéaires scalaires	122
5.3	Équations linéaires en dimension générale	124
5.3.1	Sans terme source	125
5.3.2	Avec terme source	128
5.3.3	Coefficients constants	129
6	Équilibres des équations autonomes	135
6.1	Définitions	135
6.1.1	Flot	135
6.1.2	Portrait de phase	137
6.1.3	Équilibres	139
6.2	Équations linéaires	140
6.2.1	Cas diagonalisable	141
6.2.2	Cas non-diagonalisable	143
6.2.3	Représentation graphique en dimension 2	148
6.3	Équations non-linéaires	150
6.4	L'exemple du pendule	154
6.4.1	Justification de l'équation	155
6.4.2	Équilibres	157
6.4.3	Intégrale première et portrait de phase	157
A	Connexité	165
B	Fonctions régulières à valeurs imposées	167
C	Démonstrations pour la section 3.2	173
C.1	Démonstration de la proposition 3.17	173
C.2	Démonstration de la proposition 3.21	175

C.3	Démonstration de la proposition 3.29	177
D	Compléments pour le chapitre 4	179
D.1	Lemme de Gronwall	179
D.2	Démonstration du lemme 4.11	179
D.3	Démonstration du lemme 4.12	181
D.4	Démonstration de la proposition 4.13	183

Chapitre 1

Rappels de calcul différentiel

1.1 Définition de la différentiabilité

Soient $(E, \|\cdot\|_E), (F, \|\cdot\|_F), (G, \|\cdot\|_G)$ des espaces vectoriels normés. On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires continues de E vers F ¹.

Définition 1.1 : différentiabilité en un point

Soient $U \subset E$ un ouvert et $f : U \rightarrow F$ une fonction.

Si x est un point de U , on dit que f est *différentiable en x* s'il existe $L \in \mathcal{L}(E, F)$ telle que

$$\frac{\|f(x+h) - f(x) - L(h)\|_F}{\|h\|_E} \rightarrow 0 \quad \text{quand } \|h\|_E \rightarrow 0,$$

(ou, de manière équivalente, $f(x+h) = f(x) + L(h) + o(\|h\|_E)$).
On appelle alors L la *différentielle de f en x* et on la note $df(x)$.

Remarque

Si $(E, \|\cdot\|_E) = (\mathbb{R}, |\cdot|)$, alors la différentielle, lorsqu'elle existe, est de la forme

$$h \in \mathbb{R} \quad \rightarrow \quad h z_x \in F,$$

1. Rappelons que, lorsque E est de dimension finie, toutes les applications linéaires de E vers F sont continues. Ce n'est plus vrai si E est de dimension infinie.

pour un certain élément z_x de F . Dans ce cas, on note

$$f'(x) = z_x.$$

On retrouve alors la formule bien connue :

$$f(x+h) = f(x) + f'(x)h + o(h) \quad \text{quand } h \rightarrow 0.$$

Définition 1.2 : fonctions de classe C^n

Soient $U \subset E$ un ouvert et $f : U \rightarrow F$ une fonction.

La fonction f est dite *différentiable sur U* si elle est différentiable en x pour tout $x \in U$.

Elle est *de classe C^1* si elle est différentiable et $df : U \rightarrow \mathcal{L}(E, F)$ est une application continue.

Plus généralement, pour tout $n \geq 1$, elle est *de classe C^n* si elle est différentiable et df est de classe C^{n-1} .

Elle est *de classe C^∞* si elle est de classe C^n pour tout $n \geq 1$.

On ne revient pas sur les propriétés de base liées à la différentiabilité (une somme de fonctions différentiables est différentiable etc.), à part celle qui concerne les fonctions composées.

Théorème 1.3 : composée de fonctions différentiables

Soient $U \subset E, V \subset F$ des ouverts. Soient $f : U \rightarrow V$ et $g : V \rightarrow G$ deux fonctions. Soit $x \in U$.

Si f est différentiable en x et g est différentiable en $f(x)$, alors

- $g \circ f$ est différentiable en x ;
- $d(g \circ f)(x) = dg(f(x)) \circ df(x)$.

1.2 Dérivées partielles

En géométrie différentielle, il est très fréquent de devoir effectuer des calculs explicites impliquant des différentielles de fonctions de \mathbb{R}^n vers \mathbb{R}^m . Pour cela, il est utile de pouvoir représenter les différentielles comme des matrices de taille $m \times n$ (ou des vecteurs, si $m = 1$) dont on sait calculer les

coordonnées. La notion de *dérivées partielles* permet cela.

Définition 1.4 : dérivée partielle

Soit $n \in \mathbb{N}^*$. Soient U un ouvert de \mathbb{R}^n et $f : U \rightarrow \mathbb{R}$ une fonction. Soit $x = (x_1, \dots, x_n) \in U$. Pour tout $i = 1, \dots, n$, on dit que f est *dérivable par rapport à sa i -ème variable* en x si l'application

$$y \rightarrow f(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots)$$

est dérivable en x_i . On note alors la dérivée $\partial_i f(x)$, $\partial_{x_i} f(x)$ ou $\frac{\partial f}{\partial x_i}(x)$.

Remarque

Si f est différentiable en x , alors elle est dérivable en x par rapport à chacune de ses variables. La réciproque n'est pas nécessairement vraie.

Remarque

Plus généralement, si E_1, \dots, E_n, F sont des espaces vectoriels normés, U un ouvert de $E_1 \times \dots \times E_n$ et $f : U \rightarrow F$ une fonction, on peut définir, pour tous $x = (x_1, \dots, x_n) \in U$ et $i = 1, \dots, n$, la dérivée partielle de f par rapport à x_i ,

$$\partial_{x_i} f(x) \in \mathcal{L}(E_i, F).$$

Soient maintenant $n, m \in \mathbb{N}^*$, U un ouvert de \mathbb{R}^n et $f : U \rightarrow \mathbb{R}^m$ une fonction différentiable. Pour tout x , $df(x)$ est une application linéaire de $\mathbb{R}^n \rightarrow \mathbb{R}^m$; on note $Jf(x)$ sa représentation matricielle dans les bases canoniques. Si on identifie \mathbb{R}^n (respectivement \mathbb{R}^m) avec l'ensemble des vecteurs colonnes de taille n (respectivement m), on a alors

$$\forall u \in \mathbb{R}^n, \quad df(x)(u) = Jf(x) \times u.$$

La matrice $Jf(x)$ s'appelle la *matrice jacobienne* de f au point x .

Proposition 1.5

Notons $f_1, \dots, f_m : U \rightarrow \mathbb{R}$ les composantes de f . Alors, pour tout x ,

$$Jf(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \dots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}.$$

Démonstration. Fixons $x = (x_1, \dots, x_n) \in U$. Soit $\nu \in \{1, \dots, n\}$. Notons e_ν le ν -ème vecteur de la base canonique de \mathbb{R}^n (c'est-à-dire le vecteur dont toutes les coordonnées valent 0, sauf la ν -ème qui vaut 1).

D'après la définition de la différentielle,

$$\begin{aligned} f(x_1, \dots, x_{\nu-1}, y, x_{\nu+1}, \dots) &= f(x + (y - x_\nu)e_\nu) \\ &= f(x) + (y - x_\nu)df(x)(e_\nu) + o(y - x_\nu) \\ &\quad \text{quand } y \rightarrow x_\nu. \end{aligned}$$

Pour tout $\mu \in \{1, \dots, m\}$, on a donc

$$\begin{aligned} f_\mu(x_1, \dots, x_{\nu-1}, y, x_{\nu+1}, \dots) &= f_\mu(x) + (y - x_\nu)(df(x)(e_\nu))_\mu + o(y - x_\nu) \\ &\quad \text{quand } y \rightarrow x_\nu. \end{aligned}$$

Ainsi, d'après la définition de la dérivée partielle,

$$\begin{aligned} \partial_\nu f_\mu(x) &= \lim_{y \rightarrow x_\nu} \frac{f_\mu(x_1, \dots, x_{\nu-1}, y, x_{\nu+1}, \dots) - f_\mu(x)}{y - x_\nu} \\ &= (df(x)(e_\nu))_\mu. \end{aligned}$$

Par définition de la matrice jacobienne, $(Jf(x))_{\mu,\nu} = (df(x)(e_\nu))_\mu$, donc

$$(Jf(x))_{\mu,\nu} = \partial_\nu f_\mu(x).$$

□

Exemple 1.6

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ telle que, pour tout $(x_1, x_2) \in \mathbb{R}^2$,

$$f(x_1, x_2) = (x_1 x_2, x_1 + x_2).$$

Elle est différentiable et sa matrice jacobienne est

$$\forall (x_1, x_2) \in \mathbb{R}^2, \quad Jf(x_1, x_2) = \begin{pmatrix} x_2 & x_1 \\ 1 & 1 \end{pmatrix}.$$

Dans le cas particulier où $m = 1$, la matrice jacobienne a une seule ligne :

$$\forall x \in U, \quad Jf(x) = \left(\frac{\partial f}{\partial x_1}(x) \quad \frac{\partial f}{\partial x_2}(x) \quad \dots \quad \frac{\partial f}{\partial x_n}(x) \right).$$

On appelle alors sa transposée le *gradient* :

$$\forall x \in U, \quad \nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}.$$

On observe que, pour tous $x \in U, h = (h_1, \dots, h_n) \in \mathbb{R}^n$,

$$df(x)(h) = Jf(x) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) h_i = \langle \nabla f(x), h \rangle,$$

où la notation « $\langle \cdot, \cdot \rangle$ » désigne le produit scalaire usuel sur \mathbb{R}^n .

Toujours si $m = 1$, supposons maintenant f deux fois différentiable. Sa différentielle seconde peut également se représenter par une matrice. En effet, pour tout x , $d^2 f(x) = d(df)(x)$ appartient à $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$. L'application

$$(h, l) \in \mathbb{R}^n \times \mathbb{R}^n \quad \rightarrow \quad d^2 f(x)(h)(l) \tag{1.1}$$

est donc bilinéaire. Ainsi que l'énonce la propriété qui suit, c'est même une forme quadratique (c'est-à-dire qu'elle est symétrique) et la matrice qui lui est associée dans la base canonique a une expression simple en fonction des dérivées partielles de f .

Proposition 1.7 : matrice hessienne

Soit $x \in U$. L'application définie en (1.1) est une forme bilinéaire symétrique. La matrice qui la représente dans la base canonique est

$$H(f)(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}.$$

On l'appelle *matrice hessienne* de f au point x .

1.3 Théorème d'inversion locale**Définition 1.8 : homéomorphisme**

Soient U, V deux espaces topologiques^a. Une application $\phi : U \rightarrow V$ est un *homéomorphisme* de U vers V si elle vérifie les trois propriétés suivantes :

1. ϕ est une bijection de U vers V ;
2. ϕ est continue sur U ;
3. ϕ^{-1} est continue sur V .

^a. Le lecteur ou la lectrice qui ne serait pas à l'aise avec la notion d'« espace topologique » peut se limiter au cas où U et V sont deux espaces métriques, voire au cas où U et V sont des sous-ensembles, respectivement, de \mathbb{R}^{n_1} et \mathbb{R}^{n_2} pour $n_1, n_2 \in \mathbb{N}$ deux entiers.

Définition 1.9 : difféomorphisme

Soient $n \in \mathbb{N}^*$ un entier, $U, V \subset \mathbb{R}^n$ deux ouverts. Une application $\phi : U \rightarrow V$ est un *difféomorphisme* si elle vérifie les trois propriétés suivantes :

1. ϕ est une bijection de U vers V ;
2. ϕ est de classe C^1 sur U ;

3. ϕ^{-1} est de classe C^1 sur V .

Si, en outre, ϕ et ϕ^{-1} sont de classe C^k pour un entier $k \in \mathbb{N}^*$, on dit que ϕ est un C^k -difféomorphisme.

Théorème 1.10 : inversion locale

Soient $n, k \in \mathbb{N}^*$ des entiers, $U, V \subset \mathbb{R}^n$ deux ouverts et $x_0 \in U$. Soit $\phi : U \rightarrow V$ une application de classe C^k .

Si $d\phi(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ est bijective, alors il existe $U_{x_0} \subset U$ un voisinage ouvert de x_0 et $V_{\phi(x_0)} \subset V$ un voisinage ouvert de $\phi(x_0)$ tels que ϕ réalise un C^k -difféomorphisme de U_{x_0} vers $V_{\phi(x_0)}$.

Pour la démonstration de ce résultat, on pourra se référer à [Paulin, 2009, p. 250].

Une conséquence importante du théorème d'inversion locale est le théorème des fonctions implicites, qui permet de paramétrer l'ensemble des solutions d'une équation.

Théorème 1.11 : fonctions implicites

Soient $n, m \in \mathbb{N}^*$. Soient $U \subset \mathbb{R}^n \times \mathbb{R}^m$ un ouvert, $f : U \rightarrow \mathbb{R}^m$ une application de classe C^k pour un entier $k \in \mathbb{N}^*$, et (x_0, y_0) un point de U tel que

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

Si $\partial_y f(x_0, y_0) \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ est bijective, alors il existe

- un voisinage ouvert $U_{(x_0, y_0)} \subset U$ de (x_0, y_0) ,
- un voisinage ouvert $V_{x_0} \subset \mathbb{R}^n$ de x_0 ,
- une application $g : V_{x_0} \rightarrow \mathbb{R}^m$ de classe C^k

tels que, pour tout $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$((x, y) \in U_{(x_0, y_0)} \text{ et } f(x, y) = 0) \iff (x \in V_{x_0} \text{ et } y = g(x)).$$

Dans ce théorème, il faut interpréter l'expression « $f(x, y) = 0$ » comme une équation dépendant d'un paramètre x , dont l'inconnue est y . Le théorème dit qu'au voisinage de (x_0, y_0) , l'équation a, pour chaque valeur du paramètre x , une unique solution (qui est $g(x)$) et que cette solution dépend de manière

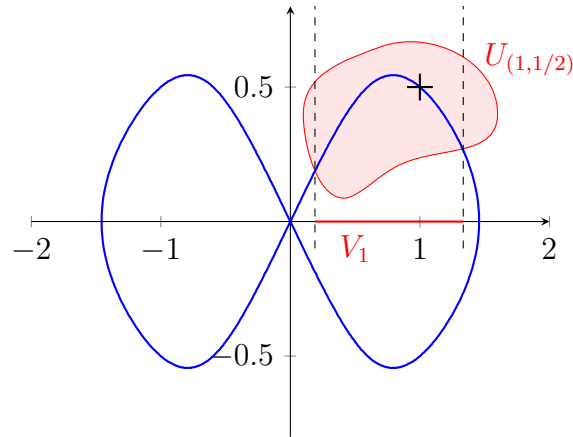


FIGURE 1.1 – En bleu, $\{(x, y) \in \mathbb{R}^2, \cos(\pi x) - \cos(\pi y) + 3x^2y^2 + \frac{x^4}{4} = 0\}$. Cet ensemble n'est pas le graphe d'une fonction. Toutefois, la partie de l'ensemble à l'intérieur de $U_{(1,1/2)}$ coïncide avec le graphe d'une fonction $g : V_1 \rightarrow \mathbb{R}$.

C^k de x .

Exemple 1.12

Il existe un voisinage ouvert $U_{(1,1/2)} \subset \mathbb{R}^2$ de $(1, 1/2)$ et un voisinage ouvert $U_1 \subset \mathbb{R}$ de 1 tels que les solutions de l'équation

$$\cos(\pi x) - \cos(\pi y) + 3x^2y^2 + \frac{x^4}{4} = 0$$

pour $(x, y) \in U_{(1,1/2)}$ sont exactement les points de l'ensemble $\{(x, g(x))\}$ pour une certaine fonction $g : U_1 \rightarrow \mathbb{R}$ de classe C^∞ .

Cela se démontre en appliquant le théorème des fonctions implicites à

$$f : (x, y) \in \mathbb{R} \times \mathbb{R} \rightarrow \cos(\pi x) - \cos(\pi y) + 3x^2y^2 + \frac{x^4}{4} \in \mathbb{R}.$$

L'hypothèse de bijectivité de $\partial_y f(1, 1)$ est bien vérifiée :

$$\partial_y f(1, 1/2) = \pi + 3 \neq 0.$$

L'ensemble des solutions de l'équation est représenté sur la figure 1.1.

Démonstration. Soit

$$\begin{aligned}\phi : U &\rightarrow \mathbb{R}^n \times \mathbb{R}^m \\ (x, y) &\rightarrow (x, f(x, y)).\end{aligned}$$

C'est une application de classe C^k et, pour tout $(h, l) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$\begin{aligned}d\phi(x_0, y_0)(h, l) &= (h, df(x_0, y_0)(h, l)) \\ &= (h, \partial_x f(x_0, y_0)(h) + \partial_y f(x_0, y_0)(l)).\end{aligned}$$

On observe que $d\phi(x_0, y_0)$ est injective. En effet, pour tout $(h, l) \in \mathbb{R}^n \times \mathbb{R}^m$, si $d\phi(x_0, y_0)(h, l) = 0$, on doit avoir

$$h = 0 \text{ et } \partial_y f(x_0, y_0)(l) = 0.$$

Puisque $\partial_y f(x_0, y_0)$ est bijective, cela entraîne $l = 0$.

Ainsi, $d\phi(x_0, y_0)$ est une application injective de $\mathbb{R}^n \times \mathbb{R}^m$ vers $\mathbb{R}^n \times \mathbb{R}^m$. Elle est donc bijective (ses espaces de départ et d'arrivée ont la même dimension).

Appliquons le théorème d'inversion locale au point (x_0, y_0) . Il existe un voisinage ouvert $U_{(x_0, y_0)}$ de (x_0, y_0) , un voisinage ouvert V de $\phi(x_0, y_0) = (x_0, 0)$ tels que ϕ réalise un C^k -difféomorphisme de $U_{(x_0, y_0)}$ vers V . Notons

$$\psi : V \rightarrow U_{(x_0, y_0)}$$

son application réciproque.

Pour tout $(x, y) \in V$, notons $\psi(x, y) = (\psi_1(x, y), \psi_2(x, y)) \in \mathbb{R}^n \times \mathbb{R}^m$. Pour tout $(x, y) \in V$,

$$\begin{aligned}(x, y) &= \phi \circ \psi(x, y) \\ &= \phi(\psi_1(x, y), \psi_2(x, y)) \\ &= (\psi_1(x, y), f(\psi_1(x, y), \psi_2(x, y))).\end{aligned}$$

On a donc nécessairement

$$\psi_1(x, y) = x.$$

Prenons

$$\begin{aligned}V_{x_0} &= \{x \in \mathbb{R}^n, (x, 0) \in V\}; \\ g : x \in V_{x_0} &\rightarrow \psi_2(x, 0) \in \mathbb{R}^m.\end{aligned}$$

Comme voulu, l'ensemble V_{x_0} est un voisinage ouvert de x_0 et g est de classe C^k . Pour tout $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$\begin{aligned} ((x, y) \in U_{(x_0, y_0)} \text{ et } f(x, y) = 0) \\ \iff ((x, y) \in U_{(x_0, y_0)} \text{ et } \phi(x, y) = (x, 0)) \\ \iff ((x, y) \in U_{(x_0, y_0)} \text{ et } (x, 0) \in V \text{ et } (x, y) = \psi(x, 0)) \\ \iff ((x, 0) \in V \text{ et } (x, y) = \psi(x, 0) = (x, \psi_2(x, 0))) \\ \iff (x \in V_{x_0} \text{ et } y = g(x)). \end{aligned}$$

□

1.4 Immersions et submersions

On introduit maintenant deux catégories particulières de fonctions différentiables, les *immersions* et les *submersions*, qui joueront un rôle très important dans la suite du cours.

Soient $n, m \in \mathbb{N}^*$ des entiers. Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction de classe C^k (pour un certain $k \geq 1$), avec U un ouvert.

Définition 1.13 : immersions et submersions

Pour tout point $x \in U$, on dit que f est une *immersion en x* si $df(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est injective. On dit que f est une *immersion* si c'est une immersion en x pour tout $x \in U$.

Pour tout point $x \in U$, on dit que f est une *submersion en x* si $df(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est surjective. On dit que f est une *submersion* si c'est une submersion en x pour tout $x \in U$.

Remarque

La fonction f ne peut être une immersion que si $n \leq m$ et une submersion que si $n \geq m$.

Si f est une immersion en un point x , elle est injective sur un voisinage de x (c'est une conséquence du théorème 1.14). Toutefois, être une immersion est une propriété significativement plus forte que l'injectivité locale. De même, une submersion est localement surjective mais toutes les applications localement surjectives ne sont pas des submersions.

Lorsque $n \leq m$, l'immersion la plus simple de \mathbb{R}^n vers \mathbb{R}^m est l'application

$$(x_1, \dots, x_n) \in \mathbb{R}^n \quad \rightarrow \quad (x_1, \dots, x_n, 0, \dots, 0) \in \mathbb{R}^m.$$

Le théorème suivant affirme qu'à changement de coordonnées près dans l'espace d'arrivée (c'est-à-dire à transformation près de l'espace d'arrivée par un difféomorphisme), toutes les immersions sont en fait localement égales à celle-là.

Théorème 1.14 : forme normale des immersions

Supposons que $0_{\mathbb{R}^n} \in U$ et $f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$.

Si f est une immersion en $0_{\mathbb{R}^n}$, il existe un voisinage U' de $0_{\mathbb{R}^n}$ et un C^k -difféomorphisme ψ d'un voisinage de $0_{\mathbb{R}^m}$ vers un voisinage de $0_{\mathbb{R}^m}$ tels que

$$\forall (x_1, \dots, x_n) \in U', \quad \psi \circ f(x_1, \dots, x_n) = (x_1, \dots, x_n, 0, \dots, 0).$$

Démonstration. Supposons que f est une immersion en $0_{\mathbb{R}^n}$.

Notons e_1, \dots, e_n les vecteurs de la base canonique de \mathbb{R}^n et $\epsilon_1, \dots, \epsilon_m$ ceux de la base canonique de \mathbb{R}^m . Commençons par supposer que

$$\forall r \in \{1, \dots, n\}, \quad df(0_{\mathbb{R}^n})(e_r) = \epsilon_r.$$

Définissons

$$\begin{aligned} \phi : \quad \mathbb{R}^m &\rightarrow \mathbb{R}^m \\ (x_1, \dots, x_m) &\rightarrow f(x_1, \dots, x_n) + (0, \dots, 0, x_{n+1}, \dots, x_m). \end{aligned}$$

On a $\phi(0) = 0$. De plus, ϕ est une application de classe C^k et, pour tout $h = (h_1, \dots, h_m) \in \mathbb{R}^m$,

$$\phi(0_{\mathbb{R}^m})(h) = df(0_{\mathbb{R}^n})(h_1, \dots, h_n) + (0, \dots, 0, h_{n+1}, \dots, h_m).$$

À partir de cette formule, on peut vérifier que $d\phi(0)(\epsilon_r) = \epsilon_r$ pour tout $r = 1, \dots, m$, c'est-à-dire que $d\phi(0) = \text{Id}_{\mathbb{R}^m}$. En particulier, $d\phi(0)$ est bijective.

D'après le théorème d'inversion locale, il existe des voisinages ouverts V_1, V_2 de $0_{\mathbb{R}^m}$ entre lesquels ϕ réalise un C^k -difféomorphisme. Notons ψ sa réciproque. Pour tout $x = (x_1, \dots, x_n) \in U' \stackrel{\text{déf}}{=} f^{-1}(V_1)$,

$$f(x_1, \dots, x_n) = \phi(x_1, \dots, x_n, 0, \dots, 0),$$

donc

$$\psi \circ f(x_1, \dots, x_n) = (x_1, \dots, x_n, 0, \dots, 0).$$

Cela achève la démonstration du théorème sous l'hypothèse que $df(0)(e_r) = \epsilon_r$ pour tout $r \in \{1, \dots, n\}$.

Cessons maintenant de faire cette hypothèse. Pour tout $r \in \{1, \dots, n\}$, on note $v_r = df(0_{\mathbb{R}^n})(e_r)$. Comme $df(0_{\mathbb{R}^n})$ est injective, la famille (v_1, \dots, v_n) est libre; on peut la compléter en une base de \mathbb{R}^m , (v_1, \dots, v_m) . Soit $L \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ telle que

$$\forall r \in \{1, \dots, m\}, \quad L(v_r) = \epsilon_r.$$

C'est une bijection, puisqu'elle envoie une base sur une base.

Notons $\tilde{f} = L \circ f$. On a $\tilde{f}(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$ et $d\tilde{f}(0_{\mathbb{R}^n}) = L \circ df(0_{\mathbb{R}^n})$; en particulier, $\tilde{f}(0_{\mathbb{R}^n})$ est une immersion en 0. Pour tout $r \in \{1, \dots, n\}$,

$$d\tilde{f}(0_{\mathbb{R}^n})(e_r) = L(df(0_{\mathbb{R}^n})(e_r)) = L(v_r) = \epsilon_r.$$

La fonction \tilde{f} vérifie donc notre hypothèse précédente. En conséquence, il existe U' un voisinage ouvert de $0_{\mathbb{R}^n}$ et $\tilde{\psi}$ un difféomorphisme entre deux voisinages de $0_{\mathbb{R}^m}$ tels que, pour tout $(x_1, \dots, x_n) \in U'$,

$$\begin{aligned} \tilde{\psi} \circ \tilde{f}(x_1, \dots, x_n) &= (x_1, \dots, x_n, 0, \dots, 0), \\ \text{c'est-à-dire } (\tilde{\psi} \circ L) \circ f(x_1, \dots, x_n) &= (x_1, \dots, x_n, 0, \dots, 0). \end{aligned}$$

Il suffit de poser $\psi = \tilde{\psi} \circ L$ pour conclure. □

Un résultat similaire est valable pour les submersions et admet une démonstration similaire. Lorsque $n \geq m$, la submersion la plus simple de \mathbb{R}^n vers \mathbb{R}^m est la projection sur les m premières coordonnées :

$$(x_1, \dots, x_n) \in \mathbb{R}^n \quad \rightarrow \quad (x_1, \dots, x_m) \in \mathbb{R}^m.$$

Quitte à changer de coordonnées dans l'espace de départ, toutes les submersions sont en fait localement égales à celles-là.

Théorème 1.15 : forme normale des submersions

Supposons que $0_{\mathbb{R}^n} \in U$ et $f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$.

Si f est une submersion en $0_{\mathbb{R}^n}$, il existe U_1, U_2 des voisinages ouverts de $0_{\mathbb{R}^n}$ et un C^k -difféomorphisme $\phi : U_1 \rightarrow U_2$ tels que

$$\forall (x_1, \dots, x_n) \in U_1, \quad f \circ \phi(x_1, \dots, x_n) = (x_1, \dots, x_m).$$

1.5 Inégalité des accroissements finis

Terminons ce chapitre avec une inégalité utile, celle des accroissements finis.

Soient $(E, \|\cdot\|_E), (F, \|\cdot\|_F)$ des espaces vectoriels normés. On munit $\mathcal{L}(E, F)$ de la norme uniforme : pour toute $u \in \mathcal{L}(E, F)$,

$$\|u\|_{\mathcal{L}(E,F)} = \sup_{x \in E \setminus \{0\}} \frac{\|u(x)\|_F}{\|x\|_E}.$$

Théorème 1.16 : inégalité des accroissements finis

Soient $U \subset E$ un ouvert convexe et $f : U \rightarrow F$ une application différentiable.

Supposons qu'il existe $M \in \mathbb{R}^+$ tel que

$$\forall x \in U, \quad \|df(x)\|_{\mathcal{L}(E,F)} \leq M.$$

Alors

$$\forall x, y \in U, \quad \|f(x) - f(y)\|_F \leq M\|x - y\|_E.$$

Pour la démonstration de ce résultat, on pourra se référer à [Paulin, 2009, p. 237].

Remarque

Attention à ne pas oublier l'hypothèse de convexité. Le théorème peut être faux si elle n'est pas vérifiée.

Par exemple, la fonction $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ telle que $f(x) = -1$ pour tout $x < 0$ et $f(x) = 1$ pour tout $x > 0$ vérifie

$$|f'(x)| \leq 0 \quad \text{pour tout } x \in \mathbb{R} \setminus \{0\}$$

(car elle est de dérivée nulle).

Pourtant, il n'est pas vrai que $|f(x) - f(y)| = 0$ pour tous $x, y \in \mathbb{R} \setminus \{0\}$.

Chapitre 2

Sous-variétés de \mathbb{R}^n

Dans ce chapitre, on introduit le sujet principal de la première moitié du cours : les sous-variétés de \mathbb{R}^n . L'intuition qu'il faut s'en faire est celle de courbes ou de surfaces « lisses » de \mathbb{R}^n .

Dans toute la suite, n et k sont des entiers fixés strictement positifs, avec k éventuellement égal à $+\infty$.

2.1 Définition

L'exemple le plus simple de sous-variété de \mathbb{R}^n est

$$\mathbb{R}^d \times \{0\}^{n-d} = \{(x_1, \dots, x_d, 0, \dots, 0), x_1, \dots, x_d \in \mathbb{R}\},$$

où d est n'importe quel entier entre 0 et n . La notion de *sous-variété de \mathbb{R}^n* généralise cet exemple : un ensemble est une sous-variété s'il est localement l'image de $\mathbb{R}^d \times \{0\}^{n-d}$ par un difféomorphisme de \mathbb{R}^n . Formalisons cette définition et donnons d'autres définitions qui lui sont équivalentes.

Définition 2.1

Soient $d \in \{0, 1, \dots, n\}$.

Soit $M \subset \mathbb{R}^n$. On dit que M est une *sous-variété de \mathbb{R}^n de dimension d et de classe C^k* si cet ensemble vérifie l'une des propriétés suivantes.

1. (Définition par redressement)

Pour tout $x \in M$, il existe un voisinage $U \subset \mathbb{R}^n$ de x , un voisinage

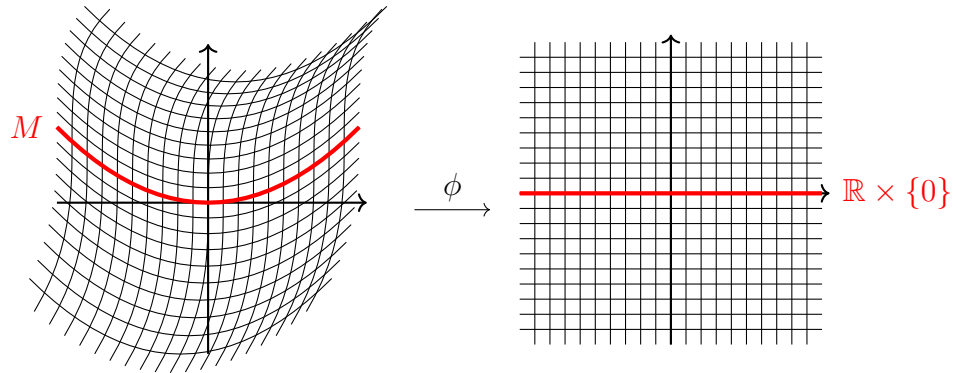


FIGURE 2.1 – Illustration de la propriété 1 de la définition 2.1 : il existe un difféomorphisme local de \mathbb{R}^2 vers \mathbb{R}^2 qui envoie l'ensemble M de la figure de gauche sur $\mathbb{R} \times \{0\}$.

$V \subset \mathbb{R}^n$ de 0 et un C^k -difféomorphisme $\phi : U \rightarrow V$ tel que

$$\phi(M \cap U) = (\mathbb{R}^d \times \{0\}^{n-d}) \cap V.$$

2. (Définition par immersion)

Pour tout $x \in M$, il existe un voisinage $U \subset \mathbb{R}^n$ de x , un ouvert V de \mathbb{R}^d , une application $f : V \rightarrow \mathbb{R}^n$ de classe C^k telle que f réalise un homéomorphisme entre V et $f(V)$,

$$M \cap U = f(V)$$

et, en notant a l'unique antécédant de x par f , f est une immersion en a .

3. (Définition par submersion)

Pour tout $x \in M$, il existe un voisinage $U \subset \mathbb{R}^n$ de x , une application $g : U \rightarrow \mathbb{R}^{n-d}$ de classe C^k qui est une submersion en x telle que

$$M \cap U = g^{-1}(\{0\})$$

4. (Définition par graphe)

Pour tout $x \in M$, il existe un voisinage $U \subset \mathbb{R}^n$ de x , un ouvert V de \mathbb{R}^d , une application $h : V \rightarrow \mathbb{R}^{n-d}$ de classe C^k et un système

de coordonnées ^a dans lequel

$$M \cap U = \text{graphe}(h) \\ \stackrel{\text{d\'ef}}{=} \{(x_1, \dots, x_d, h(x_1, \dots, x_d)), (x_1, \dots, x_d) \in V\}.$$

a. Un *système de coordonnées* est la donnée d'une base (e_1, \dots, e_n) de \mathbb{R}^n . Dans ce système, la notation (x_1, \dots, x_n) désigne le point $x_1e_1 + \dots + x_n e_n$.

Théorème 2.2

Les quatre propriétés de la définition 2.1 sont équivalentes.

Parmi les quatre définitions équivalentes du théorème, la définition par redressement (propriété 1, illustrée sur la figure 2.1) est celle qui rend le plus apparent le lien entre une sous-variété générale et la sous-variété « modèle » $\mathbb{R}^d \times \{0\}^{n-d}$. Toutefois, elle n'est pas la plus facile à appliquer : lorsqu'on veut démontrer qu'un ensemble donné est une sous-variété, les définitions par immersion, submersion ou graphe sont généralement plus commodes, comme on le verra dans la section 2.2.

Remarque

Attention au fait que, dans la définition par submersion (propriété 3), l'application g est à images dans \mathbb{R}^{n-d} et non dans \mathbb{R}^d .

De manière très informelle, dans cette définition, une sous-variété est définie comme l'ensemble des points de \mathbb{R}^n qui vérifient un ensemble d'équations scalaires

$$g(x)_1 = 0, g(x)_2 = 0, \dots$$

On s'attend intuitivement à ce que l'ensemble des solutions ait $n - e$ « degrés de liberté », où e est le nombre d'équations. Pour que la sous-variété ainsi définie soit de dimension d , il faut qu'on ait $e = n - d$, c'est-à-dire que g soit à images dans \mathbb{R}^{n-d} .

On conseille à la lectrice ou au lecteur d'étudier les exemples de la section 2.2 avant de lire la démonstration du théorème 2.2.

Démonstration du théorème 2.2.

1 \Rightarrow 3 : supposons que M vérifie la propriété **1** et montrons qu'elle vérifie la propriété **3**.

Soit $x \in M$. Soient U un voisinage de x dans \mathbb{R}^n , V un voisinage de 0 dans \mathbb{R}^n , $\phi : U \rightarrow V$ un C^k -difféomorphisme tel que

$$\phi(M \cap U) = (\mathbb{R}^d \times \{0\}^{n-d}) \cap V.$$

Notons $\text{pr}_2 : \mathbb{R}^n \rightarrow \mathbb{R}^{n-d}$ la projection sur les $n - d$ dernières coordonnées et définissons

$$g = \text{pr}_2 \circ \phi : U \rightarrow \mathbb{R}^{n-d}.$$

C'est une submersion en x car $dg(x)(\mathbb{R}^n) = \text{pr}_2(d\phi(x)(\mathbb{R}^n)) = \text{pr}_2(\mathbb{R}^n) = \mathbb{R}^{n-d}$ (on a utilisé le fait que, ϕ étant un difféomorphisme, $d\phi(x)$ est bijective).

Vérifions que $M \cap U = g^{-1}(\{0\})$.

Pour tout $x' \in M \cap U$, $\phi(x') \in \phi(M \cap U) = (\mathbb{R}^d \times \{0\}^{n-d}) \cap V \subset \mathbb{R}^d \times \{0\}^{n-d}$ donc $\text{pr}_2 \circ \phi(x') = 0$, c'est-à-dire $g(x') = 0$.

D'autre part, si $x' \in g^{-1}(\{0\})$, alors $\text{pr}_2(\phi(x')) = 0$ donc $\phi(x') \in \mathbb{R}^d \times \{0\}^{n-d}$. Comme $x' \in U$, $\phi(x') \in V$ donc $\phi(x') \in (\mathbb{R}^d \times \{0\}^{n-d}) \cap V = \phi(M \cap U)$, d'où $x' \in M \cap U$.

3 \Rightarrow 4 : supposons que M vérifie la propriété **3** et montrons qu'elle vérifie la propriété **4**.

Soit $x \in M$. Soient U un voisinage de x dans \mathbb{R}^n et $g : U \rightarrow \mathbb{R}^{n-d}$ une application de classe C^k , submersive en x , telle que

$$M \cap U = g^{-1}(\{0\}).$$

Soit (e_1, \dots, e_n) une base orthonormale de \mathbb{R}^n telle que

$$\text{Vect}\{dg(x)(e_{d+1}), \dots, dg(x)(e_n)\} = \mathbb{R}^{n-d}. \quad (2.1)$$

(Une telle base existe grâce au fait que $dg(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{n-d}$ est surjective.) On utilise dorénavant le système de coordonnées défini par cette base. On note, dans ce système,

$$x = (x_1, \dots, x_n).$$

D'après l'équation (2.1), la dérivée de g par rapport à (x_{d+1}, \dots, x_n) est surjective de \mathbb{R}^{n-d} vers \mathbb{R}^{n-d} , donc bijective. Ainsi, d'après le théorème des fonctions implicites (théorème 1.11), il existe $U' \subset U$ un voisinage de x , V un voisinage de (x_1, \dots, x_d) et $h : V \rightarrow \mathbb{R}^{n-d}$ de classe C^k telle que

$$U' \cap g^{-1}(\{0\}) = \{(t, h(t)), t \in V\}.$$

On a alors $M \cap U' = U' \cap g^{-1}(\{0\}) = \text{graphe}(h)$.

$\boxed{4 \Rightarrow 2}$: supposons que M vérifie la propriété 4 et montrons qu'elle vérifie la propriété 2.

Soit $x \in M$. Quitte à translater, on peut supposer $x = 0$, ce qui simplifie les notations. Soient U un voisinage de $x = 0$ dans \mathbb{R}^n , V un ouvert de \mathbb{R}^d et $h : V \rightarrow \mathbb{R}^{n-d}$ de classe C^k telle que, dans un système de coordonnées bien choisi,

$$M \cap U = \text{graphe}(h) = \{(t, h(t)), t \in V\}.$$

Définissons

$$\begin{aligned} f &: V \rightarrow \mathbb{R}^n \\ t &\rightarrow (t, h(t)). \end{aligned}$$

C'est une application de classe C^k . Il s'agit d'une immersion en 0 car, pour tout $t \in \mathbb{R}^d$, $df(0)(t)$ vaut

$$(t_1, \dots, t_d, dh(0)(t)),$$

ce qui ne peut s'annuler que si $t = 0$.

On a bien $f(0) = 0 = x$ (on a $h(0) = 0$ car $x = 0 \in M \cap U = \text{graphe}(h)$) et f réalise un homéomorphisme entre V et $f(V)$ (dont la réciproque est la projection sur les d premières coordonnées). De plus,

$$M \cap U = \text{graphe}(h) = f(V).$$

$\boxed{2 \Rightarrow 1}$: supposons que M vérifie la propriété 2 et montrons qu'elle vérifie la propriété 1.

Soit $x \in M$. Soient U, V des voisinages de x et 0 dans \mathbb{R}^n et \mathbb{R}^d et $f : V \rightarrow \mathbb{R}^n$ une fonction C^k réalisant un homéomorphisme de V vers $f(V)$, telle que

$$M \cap U = f(V)$$

et f est immersive en a , avec a l'unique antécédant de 0 par f . Quitte à translater, on peut supposer, pour simplifier les notations, que $a = 0$, c'est-à-dire $f(0) = x$.

D'après le théorème de forme locale des immersions (théorème 1.14), il existe un voisinage $V' \subset V$ de $0_{\mathbb{R}^d}$ et un C^k -difféomorphisme $\phi : A \rightarrow B$ entre un voisinage A de x et un voisinage B de $0_{\mathbb{R}^n}$ tels que

$$\forall (t_1, \dots, t_d) \in V', \quad \phi \circ f(t_1, \dots, t_d) = (t_1, \dots, t_d, 0, \dots, 0). \quad (2.2)$$

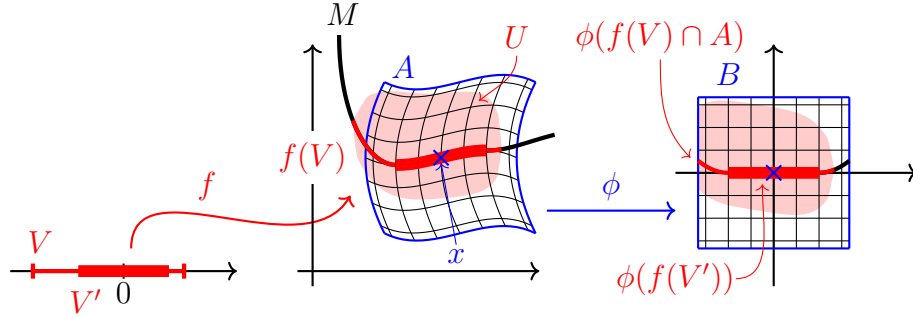


FIGURE 2.2 – Illustrations des différents objets utilisés dans la démonstration de l'implication $\boxed{2 \Rightarrow 1}$ du théorème 2.2

Une illustration des différentes définitions de cette démonstration est donnée dans la figure 2.2.

Soit $E \subset A \cap U$ un voisinage de x tel que

- $f^{-1}(f(V) \cap E) \subset V'$ (un tel voisinage existe car f est un homéomorphisme vers son image, donc f^{-1} est une application bien définie et continue sur $f(V)$);
- $\phi(E) \subset V' \times \mathbb{R}^{n-d}$.

Soit $F = \phi(E)$.

L'application ϕ réalise un C^k -difféomorphisme de E vers F . Montrons que

$$\phi(M \cap E) = (\mathbb{R}^d \times \{0\}^{n-d}) \cap F. \quad (2.3)$$

Pour tout $x' \in M \cap E$, on a $x' \in M \cap U = f(V)$ donc $x' = f(t)$ pour un certain $t \in V$. Comme $x' \in f(V) \cap E$, t est un élément de V' d'après la définition de E . Ainsi, d'après l'équation (2.2), $\phi(x') = \phi(f(t)) \in \mathbb{R}^d \times \{0\}^{n-d}$. De plus, $\phi(x') \in \phi(E) = F$. Donc $\phi(x') \in (\mathbb{R}^d \times \{0\}^{n-d}) \cap F$, ce qui montre

$$\phi(M \cap E) \subset (\mathbb{R}^d \times \{0\}^{n-d}) \cap F.$$

En sens inverse, si $(t_1, \dots, t_d, 0, \dots, 0) \in (\mathbb{R}^d \times \{0\}^{n-d}) \cap F$, alors $t \stackrel{\text{déf}}{=} (t_1, \dots, t_d)$ est un élément de V' (car $F = \phi(E) \subset V' \times \mathbb{R}^{n-d}$). Donc, d'après l'équation (2.2),

$$(t_1, \dots, t_d, 0, \dots, 0) = \phi(f(t)).$$

Comme $f(t) \in f(V) \subset M$ et $f(t) \in \phi^{-1}(F) = E$, cela montre que

$$(t_1, \dots, t_d, 0, \dots, 0) \in \phi(M \cap E).$$

D'où l'inclusion $\phi(M \cap E) \supset (\mathbb{R}^d \times \{0\}^{n-d}) \cap F$ qui achève de démontrer l'équation (2.3). \square

2.2 Exemples et contre-exemples

On a déjà vu, dans la section précédente, que, pour tout $d \in \{0, \dots, n\}$,

$$\mathbb{R}^d \times \{0\}^{n-d}$$

est une sous-variété de \mathbb{R}^n (de classe C^∞ et de dimension d).

Les ouverts constituent un autre exemple simple de sous-variétés : tout ouvert non-vide de \mathbb{R}^n est une sous-variété de dimension n de \mathbb{R}^n .

2.2.1 Sphère

Définition 2.3

La *sphère unité* de \mathbb{R}^n est l'ensemble

$$\mathbb{S}^{n-1} = \{(x_1, \dots, x_n) \in \mathbb{R}^n, x_1^2 + \dots + x_n^2 = 1\}.$$

Proposition 2.4

L'ensemble \mathbb{S}^{n-1} est une sous-variété de \mathbb{R}^n , de classe C^∞ et de dimension $n - 1$ ^a.

^a. C'est parce qu'elle est de dimension $n - 1$ qu'on la note \mathbb{S}^{n-1} et non \mathbb{S}^n .

Démonstration. On va utiliser la définition par submersion (propriété 3 de la définition 2.1).

Soit $x \in \mathbb{S}^{n-1}$. Soit $g : (t_1, \dots, t_n) \in \mathbb{R}^n \rightarrow t_1^2 + \dots + t_n^2 - 1 \in \mathbb{R}$. C'est une application de classe C^∞ . C'est une submersion au point x . En effet, $dg(x)$ est une application linéaire de \mathbb{R}^n vers \mathbb{R} . C'est donc soit l'application nulle, soit une application surjective. Or

$$\forall t = (t_1, \dots, t_n) \in \mathbb{R}^n, \quad dg(x)(t_1, \dots, t_n) = 2(x_1 t_1 + \dots + x_n t_n).$$

Comme $x_1^2 + \dots + x_n^2 = 1$, x n'est pas le vecteur nul, donc $dg(x)$ n'est pas l'application nulle : elle est surjective.

De plus, la définition de g est telle que

$$\mathbb{S}^{n-1} = g^{-1}(\{0\}).$$

La propriété 3 de la définition 2.1 est donc vérifiée (avec $U = \mathbb{R}^n$). \square

2.2.2 Produit de sous-variétés

Proposition 2.5

Soient $n_1, n_2 \in \mathbb{N}^*$, $d_1 \in \{0, \dots, n_1\}$, $d_2 \in \{0, \dots, n_2\}$. Si M_1 est une sous-variété de \mathbb{R}^{n_1} de classe C^k et de dimension d_1 et M_2 est une sous-variété de \mathbb{R}^{n_2} de classe C^k et de dimension d_2 , alors

$$M_1 \times M_2 \stackrel{\text{d\'ef}}{=} \{(x_1, x_2), x_1 \in M_1, x_2 \in M_2\}$$

est une sous-variété de $\mathbb{R}^{n_1+n_2}$ de dimension $d_1 + d_2$.

Démonstration. On utilise la définition par immersion (propriété 2 de la définition 2.1). Soit $x = (x_1, x_2) \in M$.

Comme M_1 est une sous-variété, il existe un voisinage U_1 de x_1 , un ouvert V_1 de \mathbb{R}^{d_1} et une application $f_1 : V_1 \rightarrow \mathbb{R}^{n_1}$ de classe C^k , réalisant un homéomorphisme sur son image, telle que

$$M_1 \cap U_1 = f_1(V_1)$$

et f_1 est immersive en $f_1^{-1}(x_1)$.

Définissons de même U_2, V_2 et $f_2 : V_2 \rightarrow \mathbb{R}^{n_2}$.

L'application $f : (t_1, t_2) \in V_1 \times V_2 \rightarrow (f_1(t_1), f_2(t_2)) \in \mathbb{R}^{n_1+n_2}$ est de classe C^k . Elle réalise un homéomorphisme vers son image (de réciproque $(z_1, z_2) \in f(V_1 \times V_2) \rightarrow (f_1^{-1}(z_1), f_2^{-1}(z_2))$), si on note f_1^{-1} et f_2^{-1} les réciproques respectives de f_1 et f_2). De plus,

$$\begin{aligned} (M_1 \times M_2) \cap (U_1 \times U_2) &= (M_1 \cap U_1) \times (M_2 \cap U_2) \\ &= f_1(V_1) \times f_2(V_2) \\ &= f(V_1 \times V_2). \end{aligned}$$

Enfin, f est immersive en $f^{-1}(x) = (f_1^{-1}(x_1), f_2^{-1}(x_2))$. En effet, pour tout $t = (t_1, t_2) \in \mathbb{R}^{n_1+n_2}$,

$$df(f^{-1}(x_1), f^{-1}(x_2))(t_1, t_2) = (df_1(f_1^{-1}(x_1))(t_1), df_2(f_2^{-1}(x_2))(t_2)),$$

ce qui ne vaut 0 que si $t_1 = 0$ et $t_2 = 0$, puisque $df_1(f_1^{-1}(x_1))$ et $df_2(f_2^{-1}(x_2))$ sont injectives.

L'ensemble $M_1 \times M_2$ vérifie donc bien la propriété 2 de la définition 2.1. \square

Démonstration. On utilise à nouveau la définition par submersion. Soit $x = (x_1, x_2) \in M$.

Comme M_1 est une sous-variété, il existe un voisinage U_1 de x_1 et une application $g_1 : U_1 \rightarrow \mathbb{R}^{n_1-d_1}$ de classe C^k , submersive en x_1 , telle que

$$M_1 \cap U_1 = g_1^{-1}(\{0\}).$$

Définissons de même U_2 et $g_2 : U_2 \rightarrow \mathbb{R}^{n_2-d_2}$.

L'application $g : (t_1, t_2) \in U_1 \times U_2 \rightarrow (g_1(t_1), g_2(t_2)) \in \mathbb{R}^{(n_1+n_2)-(d_1+d_2)}$ est de classe C^k et submersive en x . En effet,

$$\begin{aligned} dg(x)(\mathbb{R}^{n_1+n_2}) &= dg_1(x_1)(\mathbb{R}^{n_1}) \times dg_2(x_2)(\mathbb{R}^{n_2}) \\ &= \mathbb{R}^{n_1-d_1} \times \mathbb{R}^{n_2-d_2} \\ &= \mathbb{R}^{(n_1+n_2)-(d_1+d_2)}. \end{aligned}$$

De plus,

$$\begin{aligned} (M_1 \times M_2) \cap (U_1 \times U_2) &= (M_1 \cap U_1) \times (M_2 \cap U_2) \\ &= g_1^{-1}(\{0\}) \times g_2^{-1}(\{0\}) \\ &= g^{-1}(\{0\}). \end{aligned}$$

L'ensemble $M_1 \times M_2$ vérifie donc bien la propriété 3 de la définition 2.1. \square

Exemple 2.6 : tore

L'ensemble $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$ est une sous-variété de \mathbb{R}^4 , de dimension 2. On l'appelle *tore de dimension 2*.

2.2.3 $O_n(\mathbb{R})$

On note $\mathbb{R}^{n \times n}$ l'ensemble des matrices de taille $n \times n$ à coefficients réels. En réindexant les coordonnées, on peut également voir cet ensemble comme \mathbb{R}^{n^2} . Plusieurs sous-ensembles importants de $\mathbb{R}^{n \times n}$ ont une structure de sous-variété. On se concentre ici sur le groupe orthogonal.

Définition 2.7 : groupe orthogonal

Le *groupe orthogonal* est l'ensemble

$$O_n(\mathbb{R}) = \{A \in \mathbb{R}^{n \times n}, I_n = {}^tAA\}.$$

Proposition 2.8

L'ensemble $O_n(\mathbb{R})$ est une sous-variété de $\mathbb{R}^{n \times n}$, de classe C^∞ et de dimension $\frac{n(n-1)}{2}$.

Démonstration. On va utiliser la définition par submersion. Soit $G \in O_n(\mathbb{R})$. On va écrire $O_n(\mathbb{R})$ sous la forme $g^{-1}(\{0\})$, avec g une application C^∞ submersive en G .

Une première idée est de définir

$$g : A \in \mathbb{R}^{n \times n} \rightarrow {}^tAA - I_n \in \mathbb{R}^{n \times n}.$$

La définition du groupe orthogonal entraîne que $O_n(\mathbb{R}) = g^{-1}(\{0\})$. Toutefois, cette application n'est pas une submersion en G . En effet,

$$\forall A \in \mathbb{R}^{n \times n}, \quad dg(G)(A) = {}^tGA + {}^tAG,$$

de sorte que $dg(G)(\mathbb{R}^{n \times n})$ est inclus dans Sym_n , l'ensemble des matrices symétriques de taille $n \times n$. On a même $dg(G)(\mathbb{R}^n) = \text{Sym}_n$ car, pour tout $S \in \text{Sym}_n$,

$$dg(G) \left(\frac{GS}{2} \right) = \frac{{}^tGGS + {}^tS{}^tGG}{2} = \frac{S + {}^tS}{2} = S.$$

En particulier, $dg(G)(\mathbb{R}^{n \times n}) \neq \mathbb{R}^{n \times n}$.

On définit donc plutôt

$$\tilde{g} = \text{Tri} \circ g : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{\frac{n(n+1)}{2}},$$

où Tri est l'application qui, à toute matrice $n \times n$, associe sa partie triangulaire supérieure :

$$\forall A \in \mathbb{R}^{n \times n}, \quad \text{Tri}(A) = (A_{ij})_{i \leq j} \in \mathbb{R}^{\frac{n(n+1)}{2}}.$$

L'application \tilde{g} est de classe C^∞ . C'est une submersion en G :

$$d\tilde{g}(G)(\mathbb{R}^{n \times n}) = \text{Tri}(dg(G)(\mathbb{R}^{n \times n})) = \text{Tri}(\text{Sym}_n) = \mathbb{R}^{\frac{n(n+1)}{2}}.$$

De plus, pour toute matrice $A \in \mathbb{R}^{n \times n}$, ${}^tAA = I_n$ si et seulement si ${}^tAA - I_n = 0$, ce qui est équivalent à $\text{Tri}({}^tAA - I_n) = 0$, puisque ${}^tAA - I_n$ est une matrice symétrique. Ainsi,

$$O_n(\mathbb{R}) = \tilde{g}^{-1}(\{0\}),$$

de sorte que $O_n(\mathbb{R})$ vérifie bien la propriété 3, avec $U = \mathbb{R}^{n \times n}$ et $d = n - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$. □

2.2.4 Solutions d'équations et images de fonctions

Proposition 2.9

Soit $d \in \{0, \dots, n\}$. Soient U un ouvert de \mathbb{R}^n et

$$g : U \rightarrow \mathbb{R}^{n-d}$$

une application de classe C^k . On suppose que g est une submersion sur $g^{-1}(\{0\})$ (c'est-à-dire que g est une submersion en x , pour tout $x \in g^{-1}(\{0\})$).

Alors $g^{-1}(\{0\})$ est une sous-variété de \mathbb{R}^n , de classe C^k et dimension d .

Démonstration. C'est une application directe de la définition 2.1, version « submersion ». □

On a déjà vu deux exemples de sous-variétés définies comme dans la proposition 2.9 :

- la sphère \mathbb{S}^{n-1} est égale à $g^{-1}(\{0\})$ pour la fonction $g : x \in \mathbb{R}^n \rightarrow \|x\|^2 - 1 \in \mathbb{R}$;
- le groupe orthogonal $O_n(\mathbb{R})$ est égal à $g^{-1}(\{0\})$ pour la fonction $g : A \in \mathbb{R}^{n \times n} \rightarrow \text{Tri}({}^tAA - I_n)$.

Proposition 2.10

Soit $d \in \{0, \dots, n\}$. Soient U un ouvert de \mathbb{R}^d et $f : U \rightarrow \mathbb{R}^n$ une application de classe C^k . On suppose que f est une immersion et réalise un homéomorphisme de U vers $f(U)$.

Alors $f(U)$ est une sous-variété de \mathbb{R}^n , de classe C^k et de dimension d .

Démonstration. C'est une application directe de la définition 2.1, version « immersion ». □

Exemple 2.11 : spirale

Définissons

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R}^2 \\ \theta &\rightarrow (e^\theta \cos(2\pi\theta), e^\theta \sin(2\pi\theta)). \end{aligned}$$

Son image $f(\mathbb{R})$ est une sous-variété. Elle est représentée sur la figure 2.3.

En effet, pour tout $\theta \in \mathbb{R}$,

$$f'(\theta) = e^\theta ((\cos(2\pi\theta), \sin(2\pi\theta)) + 2\pi (-\sin(2\pi\theta), \cos(2\pi\theta))),$$

ce qui ne s'annule jamais (on observe par exemple que $\langle f'(\theta), (\cos(2\pi\theta), \sin(2\pi\theta)) \rangle = e^\theta \neq 0$ pour tout $\theta \in \mathbb{R}$). La fonction f est donc une immersion. De plus, elle réalise un homéomorphisme sur son image car elle est continue et bijective vers $f(\mathbb{R})$ et sa réciproque

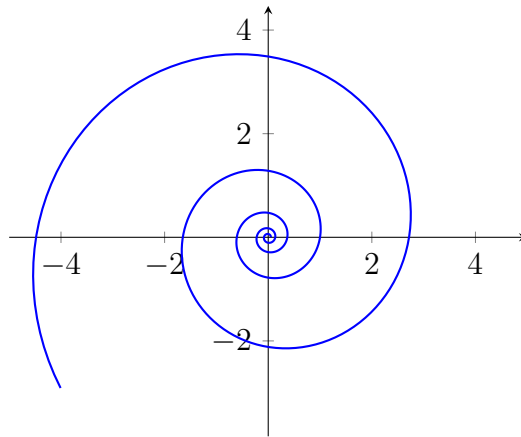
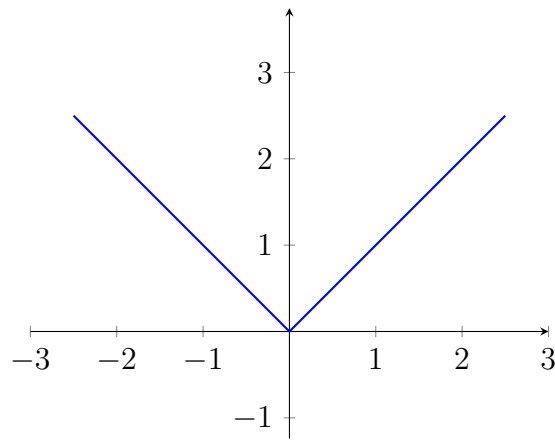
$$\begin{aligned} f^{-1} : f(\mathbb{R}) &\rightarrow \mathbb{R} \\ (x, y) &\rightarrow \frac{1}{2} \log(x^2 + y^2) \end{aligned}$$

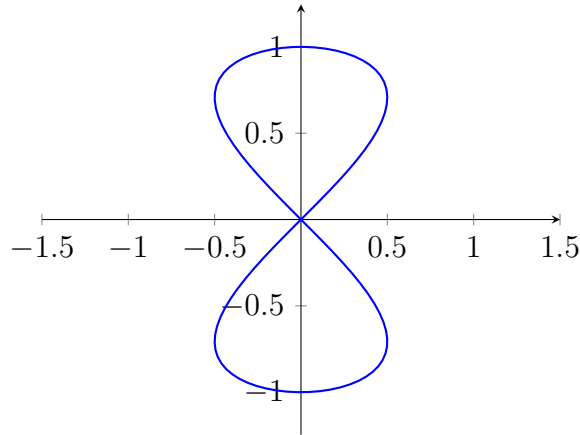
est continue (c'est la restriction à $f(\mathbb{R})$ d'une application continue sur $\mathbb{R}^2 \setminus \{(0, 0)\}$).

2.2.5 Deux contre-exemples

Le graphe de la valeur absolue (figure 2.4) n'est pas une sous-variété de \mathbb{R}^2 . Intuitivement, la raison en est que ce graphe présente un point de « non-régularité » en $(0, 0)$.

Pour le démontrer rigoureusement, le plus simple est de procéder par l'absurde. Supposons donc que c'est une sous-variété et notons d sa dimension. Alors, d'après la définition « submersion » des sous-variétés (propriété 3 de la définition 2.1), il existe $U \subset \mathbb{R}^2$ un voisinage de $(0, 0)$ et $g : U \rightarrow \mathbb{R}^{2-d}$ de

FIGURE 2.3 – Image de la fonction f définie dans l'exemple 2.11FIGURE 2.4 – Le graphe de la valeur absolue n'est pas une sous-variété de \mathbb{R}^2 .

FIGURE 2.5 – Le « huit » n'est pas une sous-variété de \mathbb{R}^2 .

classe au moins C^1 , submersive en $(0, 0)$, telle que

$$\{(t, |t|), t \in \mathbb{R}\} \cap U = g^{-1}(\{0\}). \quad (2.4)$$

Une telle application g doit vérifier, pour tout t assez proche de 0,

$$\begin{aligned} \text{si } t \leq 0, \quad 0 &= g(t, |t|) = g(t, -t), \\ \text{si } t \geq 0, \quad 0 &= g(t, |t|) = g(t, t). \end{aligned}$$

En différentiant ces deux égalités, on obtient :

$$\begin{aligned} \partial_1 g(0, 0) - \partial_2 g(0, 0) &= 0; \\ \partial_1 g(0, 0) + \partial_2 g(0, 0) &= 0. \end{aligned}$$

Cela entraîne que $\partial_1 g(0, 0) = \partial_2 g(0, 0) = 0$, c'est-à-dire $dg(0, 0) = 0$. Comme $dg(0, 0)$ est surjective, c'est impossible, à moins que $\mathbb{R}^{2-d} = \{0\}$, c'est-à-dire $d = 2$. Mais si $d = 2$, alors $g^{-1}(\{0\}) = U$, donc l'égalité (2.4) entraîne que le graphe de la valeur absolue contient un voisinage de $(0, 0)$ dans \mathbb{R}^2 , ce qui n'est pas vrai. On atteint donc une contradiction.

Le « huit » (figure 2.5) n'est pas non plus une sous-variété de \mathbb{R}^2 . Ici, la raison est que le huit est une courbe régulière, mais avec un point d'« auto-intersection » en zéro. On peut le démontrer rigoureusement par la même méthode que précédemment.

Remarque

Cet exemple montre l'importance de la propriété « f réalise un homéomorphisme vers son image » dans la définition « immersion » des sous-variétés (propriété 2 de la définition 2.1), ainsi que dans la proposition 2.10. En effet, le huit est égal à $f(] - \pi; \pi[)$, où f est l'application

$$\begin{aligned} f :] - \pi; \pi[&\rightarrow \mathbb{R}^2 \\ \theta &\rightarrow (\sin(\theta) \cos(\theta), \sin(\theta)), \end{aligned}$$

qui est une immersion et réalise une bijection (mais pas un homéomorphisme !) entre $] - \pi; \pi[$ et $f(] - \pi; \pi[)$.

2.3 Espace tangent

2.3.1 Définition

Intuitivement, l'espace tangent à une sous-variété M en un point x est l'ensemble des directions que pourrait prendre une courbe se déplaçant à la surface de M en partant du point x . Plus formellement, la définition est la suivante.

Définition 2.12 : espace tangent

Soient M une sous-variété de \mathbb{R}^n et x un point de M .

L'espace tangent à M en x , noté $T_x M$, est l'ensemble des vecteurs $v \in \mathbb{R}^n$ tels qu'il existe I un intervalle ouvert de \mathbb{R} contenant 0 et $c : I \rightarrow \mathbb{R}^n$ une application C^1 vérifiant

- $c(t) \in M$ pour tout $t \in I$;
- $c(0) = x$;
- $c'(0) = v$.

Proposition 2.13

On garde les notations de la définition précédente.

L'ensemble $T_x M$ est un sous-espace vectoriel de \mathbb{R}^n , de même dimension que M .

Démonstration. C'est une conséquence du théorème suivant. \square

Les quatre propriétés équivalentes qui permettent de définir la notion de sous-variété (définition 2.1) donnent chacune un moyen de calculer explicitement l'espace tangent.

Théorème 2.14 : calcul de l'espace tangent

Soient à nouveau M une sous-variété de \mathbb{R}^n et x un point de M . On note d la dimension de M .

1. (Calcul par redressement)

Si U et V sont des voisinages, respectivement, de x et de 0 dans \mathbb{R}^n , et si $\phi : U \rightarrow V$ est un C^k -difféomorphisme tel que $\phi(x) = 0$ et $\phi(M \cap U) = (\mathbb{R}^d \times \{0\}^{n-d}) \cap V$, alors

$$T_x M = d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d}).$$

2. (Calcul par immersion)

Si U est un voisinage de x dans \mathbb{R}^n et V un ouvert de \mathbb{R}^d , et si $f : V \rightarrow \mathbb{R}^n$ est une application C^k , réalisant un homéomorphisme entre V et $f(V)$, telle que $M \cap U = f(V)$ et f est une immersion en $z_0 \stackrel{\text{déf}}{=} f^{-1}(x)$, alors

$$T_x M = df(z_0)(\mathbb{R}^d) (= \text{Im}(df(z_0)))$$

3. (Calcul par submersion)

Si U est un voisinage de x et $g : U \rightarrow \mathbb{R}^{n-d}$ est une application C^k , submersive en x , telle que $M \cap U = g^{-1}(\{0\})$, alors

$$T_x M = \text{Ker}(dg(x)).$$

4. (Calcul par graphe)

Si U est un voisinage de x , V un ouvert de \mathbb{R}^d et $h : V \rightarrow \mathbb{R}^{n-d}$ une application C^k telle que, dans un système de coordonnées bien choisi, $M \cap U = \text{graphe}(h)$, alors

$$T_x M = \{(t_1, \dots, t_d, dh(x_1, \dots, x_d)(t_1, \dots, t_d)), t_1, \dots, t_d \in \mathbb{R}\}.$$

Démonstration. Commençons par la propriété 1. Soient U, V, ϕ comme dans l'énoncé de cette propriété.

Montrons tout d'abord l'inclusion $T_x M \subset d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d})$. Soit v un élément de $T_x M$ quelconque; on va montrer qu'il appartient à $d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d})$.

Soit c comme dans la définition de l'espace tangent, c'est-à-dire que c est une application C^1 de I dans \mathbb{R}^n , avec I un intervalle ouvert de \mathbb{R} contenant 0, à images dans M , telle que $c(0) = x$ et $c'(0) = v$.

Pour tout t assez proche de 0, $c(t)$ appartient à U donc $\phi(c(t))$ est bien défini et, puisque $\phi(M \cap U) \subset \mathbb{R}^d \times \{0\}^{n-d}$, on doit avoir

$$0 = \phi(c(t))_{d+1} = \dots = \phi(c(t))_n.$$

Dérivons ces égalités en $t = 0$:

$$\begin{aligned} 0 &= d\phi(c(0))(c'(0))_{d+1} = d\phi(x)(v)_{d+1}, \\ &\dots \\ 0 &= d\phi(x)(v)_n. \end{aligned}$$

On a donc bien $d\phi(x)(v) \in \mathbb{R}^d \times \{0\}^{n-d}$, c'est-à-dire $v \in d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d})$.

Montrons l'inclusion réciproque : $d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d}) \subset T_x M$. Soit $v \in d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d})$; montrons que $v \in T_x M$.

Notons

$$w = d\phi(x)(v) \in \mathbb{R}^d \times \{0\}^{n-d}.$$

Nous devons trouver une application c comme dans la définition de l'espace tangent. Nous allons la définir comme l'antécédant par ϕ d'une application γ à images dans \mathbb{R}^n telle que $\gamma(0) = 0$ et $\gamma'(0) = w$.

Fixons I un intervalle de \mathbb{R} contenant 0 assez petit et définissons

$$\begin{aligned} \gamma &: I \rightarrow \mathbb{R}^n \\ t &\rightarrow tw. \end{aligned}$$

C'est une application C^∞ , vérifiant

$$\gamma(0) = 0 \quad \text{et} \quad \gamma'(0) = w.$$

Si I est assez petit, $\gamma(I) \subset V$. On peut donc définir

$$c = \phi^{-1} \circ \gamma : I \rightarrow \mathbb{R}^n.$$

C'est une application de classe C^k . Elle est à images dans M . En effet, $\gamma(t) \in \mathbb{R}^d \times \{0\}^{n-d}$ pour tout $t \in I$ (car $w \in \mathbb{R}^d \times \{0\}^{n-d}$), donc

$$c(t) \in \phi^{-1}((\mathbb{R}^d \times \{0\}^{n-d}) \cap V) = M \cap U.$$

De plus,

$$c(0) = \phi^{-1}(\gamma(0)) = \phi^{-1}(0) = x$$

et

$$\begin{aligned} c'(0) &= d(\phi^{-1})(\gamma(0))(\gamma'(0)) \\ &= d(\phi^{-1})(0)(w) \\ &= d\phi(\phi^{-1}(0))^{-1}(w) \\ &= d\phi(x)^{-1}(w) \\ &= v. \end{aligned}$$

L'application c vérifie les propriétés requises dans la définition de l'espace tangent. On a donc

$$v \in T_x M.$$

Cela termine la démonstration de l'inclusion $d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d}) \subset T_x M$ et donc de l'égalité

$$T_x M = d\phi(x)^{-1}(\mathbb{R}^d \times \{0\}^{n-d}).$$

Avant de démontrer les trois autres propriétés du théorème, observons que l'égalité que nous venons d'obtenir nous permet déjà de voir que $T_x M$ est un sous-espace vectoriel de \mathbb{R}^n de dimension d . En effet, c'est l'image d'un sous-espace vectoriel de dimension d de \mathbb{R}^n ($\mathbb{R}^d \times \{0\}^{n-d}$) par un isomorphisme linéaire ($d\phi(x)^{-1}$).

Cette observation simplifie la démonstration des propriétés 2, 3 et 4. En effet, les ensembles

$$\begin{aligned} &df(z_0)(\mathbb{R}^d), \text{ Ker}(dg(x)) \\ &\text{et } \{(t_1, \dots, t_d, dh(x_1, \dots, x_d)(t_1, \dots, t_d)), t_1, \dots, t_d \in \mathbb{R}\}, \end{aligned}$$

qui apparaissent dans ces propriétés, sont des sous-espaces vectoriels de \mathbb{R}^n de dimension d (le premier est l'image de \mathbb{R}^d par une application linéaire injective, le deuxième le noyau d'une application linéaire surjective de \mathbb{R}^{n-d} vers \mathbb{R}^n , et le troisième est engendré par la famille libre à d éléments suivante :

$$(1, 0, \dots, 0, dh(x_1, \dots, x_d)(1, 0, \dots, 0)),$$

$$\begin{aligned} & \dots, \\ & (0, \dots, 0, 1, dh(x_1, \dots, x_d)(0, \dots, 0, 1)). \end{aligned}$$

Pour montrer qu'ils sont égaux à $T_x M$, il suffit donc de montrer

- soit qu'ils contiennent $T_x M$,
- soit qu'ils sont inclus dans $T_x M$.

Démontrons la propriété 2. Soient U, V, f comme dans l'énoncé de la propriété. Nous allons montrer que

$$df(z_0)(\mathbb{R}^d) \subset T_x M. \quad (2.5)$$

Soit $v \in df(z_0)(\mathbb{R}^d)$ quelconque; montrons que $v \in T_x M$. Soit $a \in \mathbb{R}^d$ tel que $df(z_0)(a) = v$. Fixons $I \subset \mathbb{R}$ un intervalle contenant 0 assez petit et définissons

$$\begin{aligned} c : I &\rightarrow \mathbb{R}^n \\ t &\rightarrow f(z_0 + ta). \end{aligned}$$

L'application c est bien définie si I est assez petit, car alors $z_0 + ta \in V$ pour tout $t \in I$. C'est une application C^k (donc C^1). Pour tout $t \in I$, $c(t) \in f(V) \subset M$. De plus,

$$c(0) = f(z_0) = x$$

et

$$c'(0) = df(z_0)(a) = v.$$

Cela démontre que $v \in T_x M$. L'équation (2.5) est donc vraie.

Démontrons maintenant la propriété 3. Soient U et g comme dans l'énoncé de la propriété. Nous allons démontrer que

$$T_x M \subset \text{Ker}(dg(x)).$$

Soit $v \in T_x M$ quelconque. On va montrer que c'est aussi un élément de $\text{Ker}(dg(x))$. Soient I un intervalle de \mathbb{R} contenant 0 et $c : I \rightarrow \mathbb{R}^n$ comme dans la définition de l'espace tangent.

Pour tout t assez proche de 0, $c(t)$ est un élément de U ; c'est de plus un élément de M . Comme $M \cap U = g^{-1}(\{0\})$,

$$0 = g(c(t)).$$

Différentions cette égalité en 0 :

$$0 = dg(c(0))(c'(0)) = dg(x)(v).$$

Donc $v \in \text{Ker}(dg(x))$.

Démontrons finalement la propriété 4. Soient U, V, h comme dans l'énoncé de cette propriété. Notons

$$E = \{(t_1, \dots, t_d, dh(x_1, \dots, x_d)(t_1, \dots, t_d)), t_1, \dots, t_d \in \mathbb{R}\}$$

et montrons que

$$E \subset T_x M.$$

Soit $(t, dh(x_1, \dots, x_d)(t)) \in E$, avec $t \in \mathbb{R}^d$. On va montrer que c'est également un élément de $T_x M$.

Fixons I un intervalle de \mathbb{R} contenant 0 assez petit et définissons

$$\begin{aligned} c : I &\rightarrow \mathbb{R}^n \\ s &\rightarrow ((x_1, \dots, x_d) + st, h((x_1, \dots, x_d) + st)). \end{aligned}$$

Cette application est bien définie si I est assez petit, car alors $(x_1, \dots, x_d) + st$ appartient à V pour tout $s \in I$ (puisque V contient (x_1, \dots, x_d) et est ouvert). Elle est de classe C^k (donc C^1). Elle est à images dans $\text{graphe}(h)$, donc dans M . De plus,

$$c(0) = (x_1, \dots, x_d, h(x_1, \dots, x_d)) = x$$

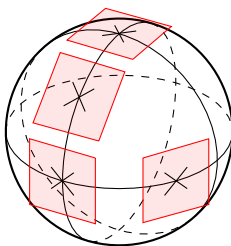
et

$$c'(0) = (t, dh(x_1, \dots, x_d)(t)).$$

Cela démontre que $(t, dh(x_1, \dots, x_d)(t)) \in T_x M$. □

Pour finir ce paragraphe de définitions, introduisons l'espace affine tangent, qui est simplement l'espace tangent, translaté de manière à passer par le point x . Ce n'est pas une notion que nous utiliserons particulièrement dans la suite du cours, sauf dans les figures : il est beaucoup plus naturel de dessiner des espaces tangents touchant¹ réellement la sous-variété à laquelle ils sont associés que des espaces tangents contenant tous 0.

1. Le mot « tangent » vient du verbe latin *tangere*, qui signifie « toucher ».

FIGURE 2.6 – La sphère \mathbb{S}^2 et son espace affine tangent en quelques points.**Définition 2.15**

Si M est une sous-variété de \mathbb{R}^n et $x \in M$, on appelle *espace affine tangent* à M en x l'ensemble

$$x + T_x M.$$

2.3.2 Exemples

Dans ce paragraphe, on reprend les exemples de sous-variétés vus dans la section 2.2 et on donne leurs espaces tangents.

Proposition 2.16 : espace tangent de la sphère

Pour tout $x \in \mathbb{S}^{n-1}$,

$$T_x \mathbb{S}^{n-1} = \{x\}^\perp = \{t \in \mathbb{R}^n, \langle t, x \rangle = 0\}.$$

Démonstration. Définissons, comme dans le paragraphe 2.2.1, l'application

$$g : \begin{array}{ccc} \mathbb{R}^n & \rightarrow & \mathbb{R} \\ (t_1, \dots, t_n) & \rightarrow & t_1^2 + \dots + t_n^2 - 1. \end{array}$$

Elle vérifie $\mathbb{S}^{n-1} = g^{-1}(\{0\})$ et c'est une submersion en x . D'après la propriété 3 du théorème 2.14,

$$T_x \mathbb{S}^{n-1} = \text{Ker}(dg(x)).$$

Or, pour tout $t \in \mathbb{R}^n$, $dg(x)(t) = 2 \langle x, t \rangle$. On a donc bien

$$T_x \mathbb{S}^{n-1} = \{x\}^\perp.$$

□

Proposition 2.17 : espace tangent d'une sous-variété produit

Soient $n_1, n_2 \in \mathbb{N}^*$. Soient M_1 une sous-variété de \mathbb{R}^{n_1} et M_2 une sous-variété de \mathbb{R}^{n_2} . Pour tout $x = (x_1, x_2) \in M_1 \times M_2$,

$$\begin{aligned} T_x(M_1 \times M_2) &= T_{x_1}M_1 \times T_{x_2}M_2 \\ &= \{(t_1, t_2), t_1 \in T_{x_1}M_1, t_2 \in T_{x_2}M_2\}. \end{aligned}$$

Démonstration. Soit $x = (x_1, x_2) \in M_1 \times M_2$.

On va utiliser l'expression pour l'espace tangent associé à la définition « immersion » des sous-variétés (propriété 2 du théorème 2.14).

Notons d_1 la dimension de M_1 . Soient U_1 un voisinage de x_1 dans \mathbb{R}^{n_1} , V_1 un voisinage de 0 dans \mathbb{R}^{d_1} et $f_1 : V_1 \rightarrow \mathbb{R}^{n_1}$ une application réalisant un homéomorphisme sur son image, telle que

$$M_1 \cap U_1 = f_1(V_1)$$

et f_1 est immersive en z_1 , où $z_1 \in V_1$ est le point tel que $f_1(z_1) = x_1$.

Définissons de même $d_2, U_2, V_2, f_2 : V_2 \rightarrow \mathbb{R}^{n_2}$ et z_2 .

D'après la propriété 2 du théorème 2.14, on a

$$T_{x_1}M_1 = df_1(z_1)(\mathbb{R}^{d_1}) \quad \text{et} \quad T_{x_2}M_2 = df_2(z_2)(\mathbb{R}^{d_2}).$$

De plus, comme on l'a vu en démontrant la proposition 2.5, l'application $f : (t_1, t_2) \in V_1 \times V_2 \rightarrow (f_1(t_1), f_2(t_2)) \in \mathbb{R}^{n_1+n_2}$ réalise un homéomorphisme sur son image, vérifie

$$f(V_1 \times V_2) = (M_1 \times M_2) \cap (U_1 \times U_2)$$

et est immersive en $(z_1, z_2) = f^{-1}(x)$. Toujours d'après la propriété 2 du théorème 2.14, on a

$$\begin{aligned} T_x(M_1 \times M_2) &= df(z_1, z_2)(\mathbb{R}^{d_1+d_2}) \\ &= \{df(z_1, z_2)(t_1, t_2), t_1 \in \mathbb{R}^{d_1}, t_2 \in \mathbb{R}^{d_2}\} \\ &= \{(df_1(z_1)(t_1), df_2(z_2)(t_2)), t_1 \in \mathbb{R}^{d_1}, t_2 \in \mathbb{R}^{d_2}\} \\ &= df_1(z_1)(\mathbb{R}^{d_1}) \times df_2(z_2)(\mathbb{R}^{d_2}) \\ &= T_{x_1}M_1 \times T_{x_2}M_2. \end{aligned}$$

□

Exemple 2.18 : espace tangent du tore

Pour tout $(x_1, x_2) \in \mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^2$,

$$T_{(x_1, x_2)}\mathbb{T}^2 = T_{x_1}\mathbb{S}^1 \times T_{x_2}\mathbb{S}^2 = \{x_1\}^\perp \times \{x_2\}^\perp.$$

Si on fixe θ_1, θ_2 tels que $x_1 = (\cos(\theta_1), \sin(\theta_1)), x_2 = (\cos(\theta_2), \sin(\theta_2))$, on a

$$\begin{aligned} \{x_1\}^\perp &= (\sin(\theta_1), -\cos(\theta_1))\mathbb{R} \\ &= \{(t_1 \sin(\theta_1), -t_1 \cos(\theta_1)), t_1 \in \mathbb{R}\} \end{aligned}$$

et de même pour x_2 . Cela permet d'écrire l'expression précédente du tangent au tore de manière légèrement plus explicite :

$$T_{(x_1, x_2)}\mathbb{T}^2 = \{(t_1 \sin(\theta_1), -t_1 \cos(\theta_1), t_2 \sin(\theta_2), -t_2 \cos(\theta_2)), t_1, t_2 \in \mathbb{R}\}.$$

Proposition 2.19 : espace tangent du groupe orthogonal

Pour toute $G \in O_n(\mathbb{R})$,

$$T_G O_n(\mathbb{R}) = \{GR, R \in \mathbb{R}^{n \times n} \text{ est antisymétrique}\}.$$

Démonstration. Soit $G \in O_n(\mathbb{R})$.

Comme on l'a vu en démontrant la proposition 2.8, $O_n(\mathbb{R})$ est égal à $\tilde{g}^{-1}(\{0\})$, pour l'application

$$\begin{aligned} \tilde{g} : \mathbb{R}^{n \times n} &\rightarrow \mathbb{R}^{\frac{n(n+1)}{2}} \\ A &\rightarrow \text{Tri}({}^tAA - I_n), \end{aligned}$$

qui est une submersion en G , de différentielle

$$d\tilde{g}(G) : A \in \mathbb{R}^{n \times n} \rightarrow \text{Tri}({}^tGA + {}^tAG) \in \mathbb{R}^{\frac{n(n+1)}{2}}.$$

D'après la propriété 3 du théorème 2.14,

$$T_G O_n(\mathbb{R}) = \text{Ker}(d\tilde{g}(G)) = \{A \in \mathbb{R}^{n \times n}, \text{Tri}({}^tGA + {}^tAG) = 0\}.$$

Or, pour toute A ,

$$\text{Tri}({}^tGA + {}^tAG) = 0 \iff {}^tGA + {}^tAG = 0$$

$$\begin{aligned}
& (\text{car } {}^tGA + {}^tAG \text{ est symétrique}) \\
\iff & ({}^tGA) + ({}^tGA) = 0 \\
\iff & {}^tGA = R \text{ pour une certaine } R \text{ antisymétrique} \\
\iff & A = GR \text{ pour une certaine } R \text{ antisymétrique} \\
& (\text{car } G^tG = I_n).
\end{aligned}$$

Donc

$$T_G O_n(\mathbb{R}) = \{GR, R \in \mathbb{R}^{n \times n} \text{ est antisymétrique}\}.$$

□

Proposition 2.20

Soit $d \in \{0, \dots, n\}$. Soient U un ouvert de \mathbb{R}^n et $g : U \rightarrow \mathbb{R}^{n-d}$ une application de classe C^k . On suppose que g est une submersion sur $g^{-1}(\{0\})$.

Pour tout $x \in g^{-1}(\{0\})$,

$$T_x(g^{-1}(\{0\})) = \text{Ker}(dg(x)).$$

Démonstration. C'est une application directe de la propriété 3 du théorème 2.14. □

Proposition 2.21

Soit $d \in \{0, \dots, n\}$. Soient U un ouvert de \mathbb{R}^d et $f : U \rightarrow \mathbb{R}^n$ une immersion réalisant un homéomorphisme de U vers $f(U)$.

Pour tout $x \in f(U)$,

$$T_x f(U) = df(z)(\mathbb{R}^d),$$

où z est l'élément de U tel que $x = f(z)$.

Démonstration. C'est une application directe de la propriété 2 du théorème 2.14. □

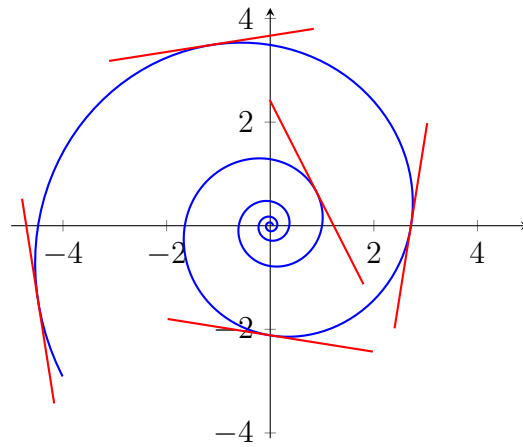


FIGURE 2.7 – La spirale de l'exemple 2.22 et son espace affine tangent en quelques points.

Exemple 2.22 : espace tangent de la spirale

Considérons, comme dans l'exemple 2.11, la fonction

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R}^2 \\ \theta &\rightarrow (e^\theta \cos(2\pi\theta), e^\theta \sin(2\pi\theta)). \end{aligned}$$

Soit $(x, y) \in f(\mathbb{R})$. Notons $\theta \in \mathbb{R}$ le réel tel que $(x, y) = f(\theta)$. D'après la proposition 2.21,

$$\begin{aligned} T_{(x,y)}f(\mathbb{R}) &= f'(\theta)\mathbb{R} \\ &= e^\theta((\cos(2\pi\theta), \sin(2\pi\theta)) + 2\pi(-\sin(2\pi\theta), \cos(2\pi\theta)))\mathbb{R} \\ &= (x - 2\pi y, y + 2\pi x)\mathbb{R} \\ &= \{((x - 2\pi y)t, (y + 2\pi x)t), t \in \mathbb{R}\}. \end{aligned}$$

Une illustration se trouve sur la figure 2.7.

2.4 Applications entre sous-variétés

Dans cette section, on s'intéresse aux fonctions entre deux sous-variétés $M \subset \mathbb{R}^{n_1}$ et $N \subset \mathbb{R}^{n_2}$:

$$f : M \rightarrow N.$$

Si $M = \mathbb{R}^{d_1} \times \{0\}^{n_1-d_1}$ et $N = \mathbb{R}^{d_2} \times \{0\}^{n_2-d_2}$, f est essentiellement une fonction de \mathbb{R}^{d_1} vers \mathbb{R}^{d_2} . Les notions de « différentiabilité » et de « différentielle » sont donc bien définies pour f , conformément au chapitre 1.

En revanche, si M n'est pas un sous-espace vectoriel de \mathbb{R}^{n_1} , ce n'est plus le cas : la définition 1.1 fait intervenir les applications linéaires entre les ensembles de départ et d'arrivée, ce qui n'existe pas si les ensembles ne sont pas des espaces vectoriels.

Pour donner tout de même un sens à la notion de différentiabilité pour f , on peut utiliser le fait que M et N sont identifiables à des ouverts de \mathbb{R}^{d_1} et \mathbb{R}^{d_2} par des difféomorphismes. On dit alors que f est différentiable si, lorsqu'on la compose avec les difféomorphismes en question, on obtient une application différentiable d'un ouvert de \mathbb{R}^{d_1} vers \mathbb{R}^{d_2} . C'est, sous une forme légèrement différente, ce qu'énonce la définition suivante.

Définition 2.23 : fonction C^1 d'une sous-variété vers \mathbb{R}^m

Soit $m \in \mathbb{N}$.

Considérons M une sous-variété de \mathbb{R}^n de classe C^k , ainsi qu'une fonction

$$f : M \rightarrow \mathbb{R}^m.$$

On dit que f est de classe C^1 si, pour tout entier $s \in \mathbb{N}^*$, tout ouvert V de \mathbb{R}^s et toute application $\phi : V \rightarrow \mathbb{R}^n$ de classe C^1 telle que $\phi(V) \subset M$, la fonction

$$f \circ \phi : V \rightarrow \mathbb{R}^m$$

est de classe C^1 .

Remarque

On peut, de même, définir la notion de *fonction de classe C^r* de M vers \mathbb{R}^m , pour tout $r = 1, \dots, k$. Il suffit de remplacer « C^1 » par « C^r » dans la définition ci-dessus.

On peut montrer qu'une fonction de classe C^r est nécessairement de

classe $C^{r'}$ pour tout $r' \leq r$.

Exemple 2.24 : projection sur une coordonnée

Soit $M \subset \mathbb{R}^n$ une sous-variété C^k . Pour tout $r = 1, \dots, n$, on définit la projection sur la r -ème coordonnée

$$\begin{aligned} \pi_r : M &\rightarrow \mathbb{R} \\ (x_1, \dots, x_n) &\rightarrow x_r. \end{aligned}$$

C'est une application C^k .

Démonstration. Soit $r \in \{1, \dots, n\}$. Soient $s \in \mathbb{N}^*$, V un ouvert de \mathbb{R}^s et $\phi : V \rightarrow \mathbb{R}^n$ de classe C^k telle que $\phi(V) \subset M$. Pour tout $x \in \mathbb{R}^s$, on note $\phi(x) = (\phi_1(x), \dots, \phi_n(x))$. Les applications ϕ_1, \dots, ϕ_n ainsi définies sont de classe C^k . Ainsi, $\pi_r \circ \phi = \phi_r$ est de classe C^k . \square

Définition 2.25 : fonction C^1 entre deux sous-variétés

Soient M, N deux sous-variétés de classe C^k , respectivement de \mathbb{R}^{n_1} et \mathbb{R}^{n_2} . Considérons une fonction

$$f : M \rightarrow N.$$

Puisque $N \subset \mathbb{R}^{n_2}$, on peut voir f comme une application de M vers \mathbb{R}^{n_2} , plutôt que de M vers N . On dit que f est *de classe C^1* (plus généralement, C^r , pour $r \in \{1, \dots, k\}$) entre M et N si elle est de classe C^1 (plus généralement, C^r) lorsqu'on la voit comme une application de M vers \mathbb{R}^{n_2} .

Exemple 2.26 : projection dans une sous-variété produit

Soient A, B deux sous-variétés de classe C^k , respectivement de \mathbb{R}^a et \mathbb{R}^b . On rappelle que $A \times B$ est une sous-variété de \mathbb{R}^{a+b} (Proposition 2.5).

On définit la projection sur A par

$$\begin{aligned} \pi_A : A \times B &\rightarrow A \\ (x_A, x_B) &\rightarrow x_A. \end{aligned}$$

C'est une application C^k .

On peut définir de même la projection sur B ; elle est aussi C^k .

Démonstration. Considérons π_A comme une fonction de $A \times B$ vers \mathbb{R}^a et montrons que cette fonction est de classe C^k . Soient $s \in \mathbb{N}^*$, V un ouvert de \mathbb{R}^s et $\phi : V \rightarrow \mathbb{R}^{a+b}$ une application de classe C^k telle que $\phi(V) \subset A \times B$.

Pour tout $x \in \mathbb{R}^s$, on note $\phi(x) = (\phi_1(x), \dots, \phi_{a+b}(x))$. Les applications $\phi_1, \dots, \phi_{a+b}$ sont de classe C^k . L'application $\pi_A \circ \phi$ vaut

$$\begin{aligned} \forall x \in \mathbb{R}^s, \quad \pi_A \circ \phi(x) &= \underbrace{(\phi_1(x), \dots, \phi_a(x))}_{\text{élément de } A}, \underbrace{(\phi_{a+1}(x), \dots, \phi_{a+b}(x))}_{\text{élément de } B} \\ &= (\phi_1(x), \dots, \phi_a(x)). \end{aligned}$$

Donc $\pi_A \circ \phi$ est égale à (ϕ_1, \dots, ϕ_a) , qui est C^k , donc $\pi_A \circ \phi$ est C^k . \square

Les définitions 2.23 et 2.25, plus abstraites que la définition de la différentiabilité pour une fonction de \mathbb{R}^n vers \mathbb{R}^m , ne doivent pas intimider. En pratique, on a rarement besoin de recourir à ces définitions pour montrer qu'une application est de classe C^1 (ou, plus généralement, C^r). En effet, comme c'est le cas pour les fonctions de $\mathbb{R}^n \rightarrow \mathbb{R}^m$, les opérations de base sur les fonctions entre variétés préservent la différentiabilité. Par exemple, si M est une sous-variété et m un entier, la somme de deux fonctions C^r de M vers \mathbb{R}^m est également C^r . De même, le produit de deux fonctions C^r de M vers \mathbb{R} est également C^r . Nous n'énoncerons pas ici chacune de ces propriétés; contentons-nous de celle relative à la composition des fonctions.

Proposition 2.27 : composition de fonctions C^1

Soient M, N, P trois sous-variétés de classe C^k de, respectivement, \mathbb{R}^{n_M} , \mathbb{R}^{n_N} et \mathbb{R}^{n_P} . Considérons deux applications

$$f_1 : M \rightarrow N \quad \text{et} \quad f_2 : N \rightarrow P.$$

Si f_1 et f_2 sont de classe C^1 , alors

$$f_2 \circ f_1 : M \rightarrow P$$

est également de classe C^1 .

Démonstration. Voyons $f_2 \circ f_1$ comme une application de M vers \mathbb{R}^{n_P} et montrons que cette application est C^1 . Soient $s \in \mathbb{N}^*$, V un ouvert de \mathbb{R}^s et $\phi : V \rightarrow \mathbb{R}^{n_M}$ une fonction de classe C^1 telle que $\phi(V) \subset M$. Montrons que $f_2 \circ f_1 \circ \phi$ est de classe C^1 sur V .

Puisque $f_1 : M \rightarrow N$ est de classe C^1 , elle est C^1 lorsqu'on la voit comme une application $f_1 : M \rightarrow \mathbb{R}^{n_N}$, ce qui entraîne que l'application $f_1 \circ \phi : V \rightarrow \mathbb{R}^{n_N}$ est C^1 . De plus, $(f_1 \circ \phi)(V) \subset f_1(M) \subset N$. Comme $f_2 : N \rightarrow P \subset \mathbb{R}^{n_P}$ est C^1 , l'application $f_2 \circ (f_1 \circ \phi)$ est C^1 .

Puisque $f_2 \circ f_1 \circ \phi = f_2 \circ (f_1 \circ \phi)$, cela démontre que $f_2 \circ f_1 \circ \phi$ est C^1 . \square

Remarquons qu'au contraire du cas où les fonctions qu'on considère vont de \mathbb{R}^n vers \mathbb{R}^m , on a défini la notion de *fonction différentiable* entre variétés sans introduire celle de *différentielle*. On peut tout de même définir cette notion de *différentielle*; c'est l'objectif de la définition suivante.

Définition 2.28 : différentielle sur des variétés

Soient M, N deux sous-variétés de classe C^k de, respectivement, \mathbb{R}^{n_1} et \mathbb{R}^{n_2} . Soit

$$f : M \rightarrow N$$

une application de classe C^r , avec $r \in \{1, \dots, k\}$.

Soit $x \in M$. Pour tout $v \in T_x M$, fixons I_v un intervalle ouvert de \mathbb{R} contenant 0 et $c_v : I_v \rightarrow \mathbb{R}^{n_1}$ comme dans la définition de l'espace tangent (2.12), c'est-à-dire une fonction C^1 à images dans M telle que $c_v(0) = x$ et $c'_v(0) = v$.

La *différentielle de f en x* , notée $df(x)$, est l'application suivante :

$$\begin{aligned} df(x) : T_x M &\rightarrow T_{f(x)} N \\ v &\rightarrow (f \circ c_v)'(0). \end{aligned}$$

L'application $df(x)$ est bien définie : $f \circ c_v : I_v \rightarrow \mathbb{R}^{n_2}$ est une application C^1 , à images dans N , telle que $f \circ c_v(0) = f(x)$, donc $(f \circ c_v)'(0)$ est bien un élément de $T_{f(x)} N$.

Remarque

Si M est un ouvert de \mathbb{R}^{n_1} , alors f , vue comme application de cet ouvert de \mathbb{R}^{n_1} vers \mathbb{R}^{n_2} est différentiable au sens usuel, et les différentielles définies aux définitions 1.1 et 2.28 coïncident, car alors, en notant $df(x)$ la différentielle usuelle,

$$(f \circ c_v)'(0) = df(c_v(0))(c'_v(0)) = df(x)(v).$$

Théorème 2.29

On garde les notations de la définition 2.28.
L'application $df(x)$ ne dépend pas du choix des intervalles I_v et des fonctions c_v .
De plus, elle est linéaire.

Démonstration. Soit $v \in T_x M$. Montrons que $df(x)(v) = (f \circ c_v)'(0)$ ne dépend pas du choix de I_v et c_v . Pour cela, nous allons donner une expression alternative de $df(x)(v)$, qui ne fait pas intervenir I_v ou c_v .

Notons d_1 et d_2 les dimensions de M et N . Soient, comme dans la définition par redressement des sous-variétés (propriété 1 de la définition 2.1), $U_M, V_M \subset \mathbb{R}^{n_1}$ des voisinages, respectivement, de x et de 0, et $\phi_M : U_M \rightarrow V_M$ un C^k -difféomorphisme tels que

$$\phi_M(M \cap U_M) = (\mathbb{R}^{d_1} \times \{0\}^{n_1-d_1}) \cap V_M.$$

On note $\phi_{M,0}^{-1}$ la restriction de ϕ_M^{-1} à $(\mathbb{R}^{d_1} \times \{0\}^{n_1-d_1}) \cap V_M$. On a

$$\begin{aligned} df(x)(v) &= (f \circ c_v)'(0) \\ &= (f \circ \phi_{M,0}^{-1} \circ \phi_M \circ c_v)'(0) \\ &= ((f \circ \phi_{M,0}^{-1}) \circ \phi_M \circ c_v)'(0). \end{aligned}$$

L'application $f \circ \phi_{M,0}^{-1}$ est définie (à peu de chose près) sur un ouvert de \mathbb{R}^{d_1} . Elle est de classe C^r sur cet ouvert (d'après la définition de la classe C^r de f , puisque $\phi_{M,0}^{-1}$ est une fonction C^r d'un ouvert de \mathbb{R}^{d_1} vers M). Les fonctions $(f \circ \phi_{M,0}^{-1}, \phi_M$ et c_v sont donc des applications définies sur des ouverts de \mathbb{R}^n (pour certaines valeurs de n) et différentiables au sens habituel. Le théorème usuel de composition des différentielles donne donc

$$df(x)(v) = (d(f \circ \phi_{M,0}^{-1})(\phi_M \circ c_v(0)) \circ d\phi_M(c_v(0)))(c'_v(0))$$

$$= d(f \circ \phi_{M,0}^{-1})(0) \circ d\phi_M(x)(v).$$

Comme annoncé, cette expression ne dépend pas de c_v , ce qui termine la première partie de la démonstration.

La linéarité de $df(x)$ découle du même argument. En effet, le raisonnement qu'on vient de faire a montré que

$$df(x) = d(f \circ \phi_{M,0}^{-1})(0) \circ d\phi_M(x),$$

c'est-à-dire que $df(x)$ est la composée de deux applications linéaires. Elle est donc linéaire. \square

Comme précédemment, la notion de différentielle pour les fonctions entre variétés est régie à peu près par les mêmes règles que pour les fonctions entre \mathbb{R}^m et \mathbb{R}^n . Énonçons à titre d'exemple la règle de composition des différentielles.

Proposition 2.30

Soient M, N, P trois sous-variétés de classe C^k de, respectivement, \mathbb{R}^{n_M} , \mathbb{R}^{n_N} et \mathbb{R}^{n_P} . Considérons deux applications de classe C^r , pour $r \in \{1, \dots, k\}$,

$$f_1 : M \rightarrow N \quad \text{et} \quad f_2 : N \rightarrow P.$$

Pour tout $x \in M$,

$$d(f_2 \circ f_1)(x) = df_2(f_1(x)) \circ df_1(x).$$

Démonstration. Soit $v \in T_x M$. Montrons que

$$d(f_2 \circ f_1)(x)(v) = df_2(f_1(x)) \circ df_1(x)(v).$$

Soient I_v un intervalle ouvert de \mathbb{R} contenant 0 et $c_v : I_v \rightarrow \mathbb{R}^{n_M}$ de classe C^1 telle que $c_v(I_v) \subset M$, $c_v(0) = x$, $c'_v(0) = v$. La définition de la différentielle donne

$$d(f_2 \circ f_1)(x)(v) = (f_2 \circ f_1 \circ c_v)'(0).$$

Notons $w = (f_1 \circ c_v)'(0) = df_1(x)(v) \in \mathbb{R}^{n_N}$. L'application $f_1 \circ c_v : I_v \rightarrow \mathbb{R}^{n_N}$ est de classe C^1 et à images dans N . Elle vérifie $f_1 \circ c_v(0) = f_1(x)$ et, par définition de w , $(f_1 \circ c_v)'(0) = w$. La définition de la différentielle pour f_2 donne donc

$$df_2(f_1(x))(w) = (f_2 \circ f_1 \circ c_v)'(0).$$

Ainsi,

$$\begin{aligned} d(f_2 \circ f_1)(x)(v) &= df_2(f_1(x))(w) \\ &= df_2(f_1(x))(df_1(x)(v)) \\ &= [df_2(f_1(x)) \circ df_1(x)](v). \end{aligned}$$

□

Au-delà des règles portant sur la différentiabilité et les opérations usuelles, beaucoup de notions et résultats de calcul différentiel classique se généralisent de manière naturelle au calcul différentiel sur les variétés. Nous donnons ci-dessous les exemples de la notion de *difféomorphisme* et du théorème d'inversion locale.

Définition 2.31

Soient M, N deux sous-variétés de classe C^k de, respectivement, \mathbb{R}^{n_1} et \mathbb{R}^{n_2} . Considérons une application

$$\phi : M \rightarrow N.$$

Pour tout $r \in \{1, \dots, k\}$, on dit que ϕ est un C^r -difféomorphisme entre M et N si elle vérifie les trois propriétés suivantes :

1. ϕ est une bijection de M vers N ;
2. ϕ est de classe C^r sur M ;
3. ϕ^{-1} est de classe C^r sur N .

Théorème 2.32 : inversion locale sur des sous-variétés

Soient M, N deux sous-variétés de classe C^k de, respectivement, \mathbb{R}^{n_1} et \mathbb{R}^{n_2} . Soit $x_0 \in M$. Pour un $r \in \{1, \dots, k\}$, considérons une application de classe C^r ,

$$f : M \rightarrow N.$$

Si $df(x_0) : T_{x_0}M \rightarrow T_{f(x_0)}N$ est bijective, alors il existe U_{x_0} un voisinage ouvert de x_0 dans M et $V_{f(x_0)}$ un voisinage ouvert de $f(x_0)$ dans N tels que f réalise un C^r -difféomorphisme de U_{x_0} vers $V_{f(x_0)}$.

Démonstration. Soit d la dimension de M . On remarque que N a la même dimension d : $df(x_0)$ est une application linéaire bijective entre $T_{x_0}M$ et $T_{f(x_0)}N$, donc

$$\dim T_{f(x_0)}N = \dim T_{x_0}M = d.$$

Soient, comme dans la définition « redressement » d'une sous-variété (propriété 1 de la définition 2.1), $U_M, V_M \subset \mathbb{R}^{n_1}$ des voisinages ouverts, respectivement, de x_0 et de 0, et $\phi_M : U_M \rightarrow V_M$ un C^k -difféomorphisme tels que

$$\phi_M(M \cap U_M) = (\mathbb{R}^d \times \{0\}^{n_1-d}) \cap V_M.$$

Quitte à composer ϕ_M par une translation, on peut supposer que $\phi_M(x_0) = 0$.

De même, soient $U_N, V_N \subset \mathbb{R}^{n_2}$ des voisinages ouverts de $f(x_0)$ et 0, $\phi_N : U_N \rightarrow V_N$ un C^k -difféomorphisme tels que

$$\phi_N(N \cap U_N) = (\mathbb{R}^d \times \{0\}^{n_2-d}) \cap V_N.$$

On peut supposer que $\phi_N(f(x_0)) = 0$.

Le principe de la démonstration est de se ramener au cas où f est définie sur un ouvert de \mathbb{R}^d puis d'appliquer le théorème d'inversion locale classique. Pour cela, on « transfère » f en une application de $\mathbb{R}^d \times \{0\}^{n_1-d}$ vers $\mathbb{R}^d \times \{0\}^{n_2-d}$ en la composant par les difféomorphismes ϕ_M et ϕ_N .

Plus précisément, notons $\phi_{M,0}^{-1}$ la restriction à $(\mathbb{R}^d \times \{0\}^{n_1-d}) \cap V_M$ de ϕ_M^{-1} . On définit

$$g \stackrel{\text{déf}}{=} \phi_N \circ f \circ \phi_{M,0}^{-1} : (\mathbb{R}^d \times \{0\}^{n_1-d}) \cap V_M \rightarrow (\mathbb{R}^d \times \{0\}^{n_2-d}) \cap V_N.$$

Cette définition est valide, quitte à réduire un peu V_M . La fonction g est de classe C^r et sa différentielle en 0 est injective : c'est la composée de $d\phi_N(f(x_0))$, $df(x_0)$ et $d\phi_{M,0}^{-1}(0)$, qui sont toutes trois injectives. Comme elle va de \mathbb{R}^d vers \mathbb{R}^d ², elle est bijective.

D'après le théorème d'inversion locale (théorème 1.10), il existe E_M, E_N des voisinages ouverts de 0 dans \mathbb{R}^d tels que g réalise un C^r -difféomorphisme de $E_M \times \{0\}^{n_1-d}$ vers $E_N \times \{0\}^{n_2-d}$. Alors f réalise un C^r -difféomorphisme de $U_{x_0} \stackrel{\text{déf}}{=} \phi_M^{-1}(E_M \times \{0\}^{n_1-d})$ vers $V_{f(x_0)} \stackrel{\text{déf}}{=} \phi_N^{-1}(E_N \times \{0\}^{n_2-d})$: sur ces ensembles,

$$f = \phi_N^{-1} \circ g \circ \phi_M.$$

2. On peut voir $\phi_N \circ f \circ \phi_{M,0}^{-1}$ comme une application entre deux ouverts de \mathbb{R}^d .

Puisque ϕ_M réalise un difféomorphisme (de classe C^k donc aussi de classe C^r) de U_{x_0} vers $E_M \times \{0\}^{n_1-d}$, g un C^r -difféomorphisme de $E_M \times \{0\}^{n_1-d}$ vers $E_N \times \{0\}^{n_2-d}$ et ϕ_N^{-1} un difféomorphisme (C^k donc aussi C^r) de $E_N \times \{0\}^{n_2-d}$ vers $V_{f(x_0)}$, l'application f est une composition de C^r -difféomorphismes, donc un C^r -difféomorphisme. □

Chapitre 3

Géométrie riemannienne

Soient $k, n \in \mathbb{N}^*$ fixés.

Dans le chapitre précédent, nous avons introduit la notion de différentiabilité pour une application définie entre deux variétés. Cette notion permet d'étudier les propriétés *topologiques* des sous-variétés : on peut se demander quelles sous-variétés sont difféomorphes entre elles et quelles propriétés caractérisent le fait d'être ou non difféomorphes. Dit de manière informelle, on peut se poser des questions comme : « Une bouée est-elle difféomorphe à un ballon ? ». ¹

Dans ce chapitre, on s'intéresse à des propriétés plus fines des sous-variétés, qui sont des propriétés *métriques*, c'est-à-dire faisant intervenir des notions de longueur, d'angle ... Nous allons en particulier définir une notion d'isométrie, plus restrictive que celle de difféomorphisme (dans le sens que deux variétés qui sont isométriques sont nécessairement difféomorphes, tandis que la réciproque n'est pas vraie).

Comme ces propriétés sont assez difficiles à définir de manière formelle et comme l'objectif, ici, est seulement d'en donner un aperçu et non un panorama complet, on se limitera principalement au cas le plus simple, celui des sous-variétés de dimension 1. Les sous-variétés de dimension générale seront évoquées seulement à la fin du chapitre.

1. Réponse : non.

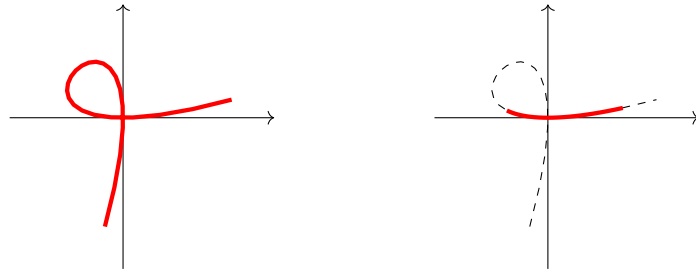


FIGURE 3.1 – L'image de l'arc paramétré $\gamma : t \in \mathbb{R} \rightarrow (t(t+1)^2, t^2(t+1))$ (figure de gauche) n'est pas une sous-variété de \mathbb{R}^2 car $(0,0)$ en est un point multiple. En revanche, $\gamma(] - \epsilon; \epsilon[)$ est une sous-variété de \mathbb{R}^2 pour tout ϵ assez petit (figure de droite).

3.1 Sous-variétés de dimension 1

Définition 3.1 : courbe

Une *courbe* est une sous-variété de \mathbb{R}^n de dimension 1.

3.1.1 Arcs paramétrés

Les courbes ont la particularité, par rapport aux variétés de dimension supérieure, qu'elles admettent nécessairement un paramétrage simple, c'est-à-dire qu'on peut essentiellement les voir comme l'image d'un ouvert de \mathbb{R} par une fonction C^1 . Ce paramétrage permet de définir commodément des quantités métriques, comme nous le verrons dans la suite de la section.

Définition 3.2 : arc paramétré

On appelle *arc paramétré* de classe C^k un couple (I, γ) , où I est un intervalle de \mathbb{R} et $\gamma : I \rightarrow \mathbb{R}^n$ est une fonction de classe C^k .

L'image d'un arc paramétré n'est pas nécessairement une sous-variété de \mathbb{R}^n , en particulier car l'arc peut « se recouper » (on parle de *point multiple*). En revanche, la proposition suivante montre que l'image d'un arc paramétré (I, γ) définit localement une sous-variété, au voisinage des points où γ' ne s'annule pas. Ce résultat est illustré sur la figure 3.1.

Proposition 3.3

Soit (I, γ) un arc paramétré. Soient $t \in \overset{\circ}{I}$ et $x = \gamma(t)$. On dit que x est un *point régulier* si $\gamma'(t) \neq 0$.

Dans ce cas, il existe $\epsilon > 0$ tel que $]t - \epsilon; t + \epsilon[\subset I$ et l'ensemble

$$C \stackrel{\text{déf}}{=} \gamma(]t - \epsilon; t + \epsilon[)$$

est une courbe. De plus,

$$T_x C = \mathbb{R}\gamma'(t).$$

Démonstration. Supposons que x est régulier, c'est-à-dire que γ est une immersion au point t . Si on parvient à montrer que, pour $\epsilon > 0$ assez petit, γ réalise un homéomorphisme de $]t - \epsilon; t + \epsilon[$ vers son image, le théorème est démontré. En effet, on peut alors choisir $\epsilon > 0$ assez petit pour que γ' ne s'annule pas (c'est-à-dire pour que γ soit immersive) sur tout l'intervalle $]t - \epsilon; t + \epsilon[$. La proposition 2.10 garantit alors que

$$C \stackrel{\text{déf}}{=} \gamma(]t - \epsilon; t + \epsilon[)$$

est une sous-variété de \mathbb{R}^n de dimension 1, c'est-à-dire une courbe et la propriété 2 du théorème 2.14 nous dit que

$$T_x C = \text{Im}(d\gamma(t)) = \mathbb{R}\gamma'(t).$$

Pour montrer que γ réalise un homéomorphisme de $]t - \epsilon; t + \epsilon[$ vers son image si $\epsilon > 0$ est assez petit, on utilise le théorème de forme normale des immersions (théorème 1.14). Soient ψ un difféomorphisme d'un voisinage de x vers un voisinage de $0_{\mathbb{R}^n}$ et $\epsilon > 0$ tels que

$$\forall t' \in]t - \epsilon; t + \epsilon[, \quad \psi \circ \gamma(t') = (t', 0, \dots, 0).$$

En notant $\pi_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ la projection sur la première coordonnée, on obtient que

$$\forall t' \in]t - \epsilon; t + \epsilon[, \quad \pi_1 \circ \psi \circ \gamma(t') = t'.$$

En conséquent, γ réalise un homéomorphisme de $]t - \epsilon; t + \epsilon[$ vers son image (de réciproque $\pi_1 \circ \psi$). \square

Réciproquement, toute courbe connexe² est l'image d'un arc paramétré. C'est une conséquence des théorèmes qui suivent.

Théorème 3.4 : courbes compactes

Soit $M \subset \mathbb{R}^n$ une courbe compacte et connexe de classe C^k . Elle est C^k -difféomorphe au cercle \mathbb{S}^1 .

Théorème 3.5 : courbes non-compactes

Soit $M \subset \mathbb{R}^n$ une courbe connexe non compacte de classe C^k . Elle est C^k -difféomorphe à \mathbb{R} .

La démonstration de ces théorèmes est difficile. Nous allons nous limiter à celle du premier ; elle sera donnée dans la sous-section 3.1.2. La démonstration du second utilise en partie la même stratégie mais nécessite des idées supplémentaires.

Comme dit avant les énoncés, ces théorèmes impliquent que toute courbe M est l'image d'un arc paramétré. De manière un peu plus précise, le théorème 3.5 garantit que, si M n'est pas compacte, il existe un arc paramétré (I, γ) de classe C^k , avec I un ouvert de \mathbb{R} , tel que γ réalise un C^k -difféomorphisme de I vers M (dans le théorème, l'intervalle I considéré est \mathbb{R} mais il est utile d'autoriser des paramétrages avec d'autres intervalles). On appelle un tel arc (I, γ) un *paramétrage global* de M .

Si M est compacte, le théorème 3.4 garantit qu'il existe $\phi : \mathbb{S}^1 \rightarrow M$ un C^k -difféomorphisme. Pour tous $a, b \in \mathbb{R}$ avec $a < b$, on peut définir

$$\begin{aligned} \gamma : [a; b[&\rightarrow \mathbb{R}^n \\ t &\rightarrow \phi \left(e^{2\pi i \frac{t-a}{b-a}} \right). \end{aligned}$$

Avec cette définition, γ est un arc paramétré de classe C^k , qui réalise une bijection de $[a; b[$ vers M . On peut observer que les valeurs et dérivées de γ en a « se recollent » avec les valeurs et dérivées en b : pour tout $k' = 0, \dots, k$,

$$\gamma^{(k')}(t) \xrightarrow{t \rightarrow b^-} \gamma^{(k')}(a)$$

(car γ est la restriction à $[a; b[$ d'une application $(b-a)$ -périodique). L'application γ ne réalise pas un C^k -difféomorphisme de $[a; b[$ vers M (sa réciproque

2. Quelques rappels sur la connexité se trouvent dans l'annexe A.

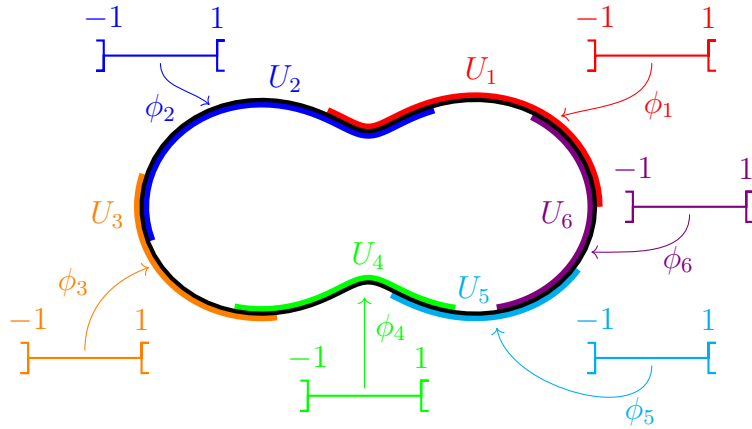


FIGURE 3.2 – Illustration du lemme 3.6 : la courbe M (la ligne noire) et son recouvrement par les ouverts U_s .

n'est pas continue en $\gamma(a) = \phi(1)$) mais elle réalise un C^k -difféomorphisme de $]a; b[$ vers $M - \{\phi(1)\}$.

On appelle un tel arc $(]a; b[, \gamma)$ un *paramétrage global* de M .

3.1.2 Démonstration du théorème 3.4

La démonstration est délicate. Elle utilise un certain nombre de lemmes intermédiaires dont la démonstration sera donnée ensuite.

Le premier lemme, dont la démonstration repose seulement sur la définition de sous-variété et la compacité de M , affirme que M peut être recouverte par un nombre fini d'ouverts difféomorphes à $] - 1; 1[$.

Lemme 3.6

Il existe un nombre fini d'ouverts de M , qu'on note U_1, \dots, U_S , tels que

1. $M = U_1 \cup \dots \cup U_S$;
2. pour tout $s \leq S$, U_s est C^k -difféomorphe à $] - 1; 1[$.

Le principe de la démonstration est de considérer un recouvrement fini comme dans le lemme précédent et de construire étape par étape un recouvrement de plus en plus petit, en fusionnant progressivement les ouverts du

recouvrement. Soit donc (U_1, \dots, U_S) un recouvrement comme dans le lemme 3.6. Pour tout s , notons

$$\phi_s :]-1; 1[\rightarrow U_s$$

un C^k -difféomorphisme.

Nous allons maintenant choisir judicieusement deux ouverts U_{s_1}, U_{s_2} et les fusionner pour obtenir, selon les propriétés de $U_{s_1} \cap U_{s_2}$

- soit directement que M est C^k -difféomorphe à \mathbb{S}^1 ;
- soit qu'il existe un recouvrement comme dans le lemme 3.6, de taille $S - 1$ plutôt que S .

Dans le premier cas, la démonstration sera terminée. Dans le deuxième cas, on réappliquera itérativement la procédure pour obtenir un recouvrement avec un nombre d'éléments de plus en plus petit.

Le lemme qui suit indique ce à quoi peut ressembler $U_{s_1} \cap U_{s_2}$.

Lemme 3.7

Pour tous $s_1, s_2 \leq S$ distincts, l'intersection $U_{s_1} \cap U_{s_2}$ vérifie l'une des propriétés suivantes :

1. $U_{s_1} \cap U_{s_2}$ est vide.
2. $U_{s_1} \cap U_{s_2}$ a une seule composante connexe. Dans ce cas, on est dans l'une des situations suivantes :
 - (a) $U_{s_1} \subset U_{s_2}$ ou $U_{s_2} \subset U_{s_1}$;
 - (b) $\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2})$ et $\phi_{s_2}^{-1}(U_{s_1} \cap U_{s_2})$ sont des intervalles de la forme $] - 1; \alpha[$ ou $]\alpha; 1[$, avec $\alpha \in] - 1; 1[$.
3. $U_{s_1} \cap U_{s_2}$ a deux composantes connexes. Dans ce cas, $\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2})$ et $\phi_{s_2}^{-1}(U_{s_1} \cap U_{s_2})$ sont de la forme $] - 1; \alpha[\cup]\beta; 1[$, avec $\alpha, \beta \in] - 1; 1[$, $\alpha < \beta$.

Choisissons $s_1, s_2 \in \{1, \dots, S\}$ distincts tels que $U_{s_1} \cap U_{s_2} \neq \emptyset$. C'est possible : de tels s_1, s_2 existent. Pour le démontrer, procédons par l'absurde. Supposons qu'il n'existe pas $s_1 \neq s_2$ tels que $U_{s_1} \cap U_{s_2} \neq \emptyset$. Alors on est dans l'une des situations suivantes :

1. $S = 1$;
2. $S > 1$ et $U_{s_1} \cap U_{s_2} = \emptyset$ pour tous $s_1 \neq s_2$.

Dans le premier cas, on doit avoir $M = U_{s_1}$. Puisque U_{s_1} est C^k -difféomorphe à $] - 1; 1[$, M l'est aussi. C'est impossible : un ensemble compact (ici, M) ne peut pas être difféomorphe à un ensemble non-compact (ici, $] - 1; 1[$). Dans le deuxième cas,

$$U_{s_1} \text{ et } U_{s_2} \cup \dots \cup U_S$$

sont des ouverts non-vides³ et disjoints, dont l'union est M . Donc M n'est pas connexe : on aboutit de nouveau à une impossibilité, ce qui conclut sur l'existence de s_1, s_2 .

Puisque l'intersection $U_{s_1} \cap U_{s_2}$ est non-vide, on est dans la situation 2 ou 3 du lemme 3.7. Si on est dans la situation 3, le lemme suivant conclut directement la démonstration du théorème.

Lemme 3.8 : deux composantes connexes

Si U_{s_1}, U_{s_2} vérifient la propriété 3 du lemme 3.7, alors M est C^k -difféomorphe à \mathbb{S}^1 .

Si on est au contraire dans la situation 2, c'est un autre lemme qu'il faut utiliser.

Lemme 3.9 : une composante connexe

Si U_{s_1}, U_{s_2} vérifient la propriété 2 du lemme 3.7, alors $U_{s_1} \cup U_{s_2}$ est C^k -difféomorphe à $] - 1; 1[$.

Dans ce cas, on obtient que $\{U_s, s \neq s_1, s_2\} \cup \{U_{s_1} \cup U_{s_2}\}$ est un ensemble d'ouverts C^k -difféomorphes à $] - 1; 1[$ dont l'union est M tout entier. On a donc trouvé un ensemble $\tilde{U}_1, \dots, \tilde{U}_{S-1}$ d'ouverts vérifiant les propriétés du lemme 3.6 mais de cardinal strictement inférieur à S .

On peut alors réappliquer le même raisonnement : il existe $\tilde{s}_1 \neq \tilde{s}_2$ tels que $\tilde{U}_{\tilde{s}_1} \cap \tilde{U}_{\tilde{s}_2} \neq \emptyset$. Si l'intersection a deux composantes connexes, alors M est C^k -difféomorphe à \mathbb{S}^1 , ce qui conclut la démonstration, et si elle a une seule composante connexe, alors on peut trouver un ensemble de $S - 2$ ouverts vérifiant les propriétés du lemme 3.6. Et ainsi de suite.

Le raisonnement ne peut pas être appliqué plus de S fois (sinon on trouverait un recouvrement de M par un nombre négatif d'ouverts). Nécessairement, il arrive donc un moment où l'intersection a deux composantes

3. Les U_s ne peuvent pas être vides, sinon ils ne seraient pas difféomorphes à $] - 1; 1[$.

connexes, ce qui entraîne que M est C^k -difféomorphe à \mathbb{S}^1 et conclut.

Démonstration du lemme 3.6. Soit d'abord $x \in M$ quelconque, fixé. Soient V un voisinage ouvert de x dans \mathbb{R}^n , I un voisinage ouvert de 0 dans \mathbb{R} et $f : I \rightarrow V$ une application C^k réalisant un homéomorphisme sur son image, telle que

$$f(I) = V \cap M$$

et immersive en $z_0 = f^{-1}(x)$. (C'est la définition « immersion » d'une sous-variété de dimension 1 - propriété 2 de la définition 2.1.)

Quitte à réduire un peu I et V , on peut supposer que I est un intervalle ouvert borné et que f est immersive sur tout I . Notons

$$U(x) = f(I) = V \cap M.$$

C'est un ouvert de M . De plus, il est C^k -difféomorphe à I (en effet, il est homéomorphe à I , par hypothèse sur f ; pour tout x' , $df(x')$ est injective, donc bijective, de $T_{x'}I$ vers $T_{f(x')}M$; d'après le théorème d'inversion locale 2.32, f est donc un C^k -difféomorphisme local, ce qui entraîne que f^{-1} est C^k). Comme tout intervalle ouvert non-vidé de \mathbb{R} est C^k -difféomorphe à $] - 1; 1[$, $U(x)$ est C^k -difféomorphe à $] - 1; 1[$.

On ne considère maintenant plus un x fixé.

Pour tout $x \in M$, $x \in U(x) \subset \cup_{x' \in M} U(x')$. Donc

$$M \subset \bigcup_{x' \in M} U(x'),$$

c'est-à-dire que les $U(x')$ forment un recouvrement de M par des ouverts. Puisque M est compacte, on peut extraire de ce recouvrement un sous-recouvrement fini : il existe x_1, \dots, x_S tels que

$$M = U(x_1) \cup \dots \cup U(x_S).$$

Comme on a vu que $U(x_s)$ était difféomorphe à $] - 1; 1[$ pour tout s , on a démontré le résultat. \square

Démonstration du lemme 3.7. L'ensemble $\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2})$ est un ouvert de $] - 1; 1[$. On peut donc l'écrire comme une union d'intervalles ouverts disjoints de $] - 1; 1[$ (voir l'exemple A.5) :

$$\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2}) = \bigcup_{l \in E}]a_l; b_l[$$

où E est un ensemble d'indices (qui peut être fini ou infini).

Commençons par supposer qu'il existe $k \in E$ tel que $-1 < a_k < b_k < 1$. Nous allons montrer qu'alors $U_{s_2} \subset U_{s_1}$.

L'application $\phi_{s_2}^{-1} \circ \phi_{s_1} : \phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2}) \rightarrow]-1; 1[$ est continue et injective (c'est la composée de deux applications continues et injectives). Elle est donc monotone sur chaque intervalle inclus dans $\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2})$. Supposons par exemple qu'elle est croissante sur $]a_k; b_k[$ (on pourrait faire un raisonnement identique si elle était décroissante).

Notons

$$B_k = \lim_{t \rightarrow b_k^-} \phi_{s_2}^{-1} \circ \phi_{s_1}(t).$$

(On prend soin d'observer que la limite existe : $\phi_{s_2}^{-1} \circ \phi_{s_1}$ est une fonction croissante et majorée, par -1 , sur $]a_k; b_k[$; elle admet donc une limite, qui est inférieure ou égale à -1 .)

Il est impossible que $B_k < 1$. En effet, si $B_k < 1$, alors $\phi_{s_2}(B_k)$ est bien défini et, par continuité de ϕ_{s_2} ,

$$\begin{aligned} \phi_{s_2}(B_k) &= \phi_{s_2}\left(\lim_{t \rightarrow b_k^-} \phi_{s_2}^{-1} \circ \phi_{s_1}(t)\right) \\ &= \lim_{t \rightarrow b_k^-} \phi_{s_1}(t) \\ &= \phi_{s_1}(b_k). \end{aligned}$$

Donc $\phi_{s_1}(b_k) \in \phi_{s_1}(]-1; 1[) \cap \phi_{s_2}(]-1; 1[) = U_{s_1} \cap U_{s_2}$, de sorte que

$$b_k \in \phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2}) = \bigcup_{l \in E}]a_l; b_l[.$$

Donc $b_k \in]a_l; b_l[$ pour un certain $l \in E$ tel que $l \neq k$, et, pour ce l , on doit avoir $]a_k; b_k[\cap]a_l; b_l[\neq \emptyset$, ce qui contredit le fait que les intervalles $]a_l; b_l[$ sont disjoints. Donc $B_k = 1$.

De même, on définit

$$A_k = \lim_{t \rightarrow a_k^+} \phi_{s_2}^{-1} \circ \phi_{s_1}(t)$$

et on montre que $A_k = -1$.

L'image de $]a_k; b_k[$ par $\phi_{s_2}^{-1} \circ \phi_{s_1}$ est un intervalle (c'est l'image d'un intervalle par une application continue); elle est incluse dans $]-1; 1[$ et on vient de voir que

$$\phi_{s_2}^{-1} \circ \phi_{s_1}(t) \xrightarrow{t \rightarrow b_k^-} 1 \quad \text{et} \quad \phi_{s_2}^{-1} \circ \phi_{s_1}(t) \xrightarrow{t \rightarrow a_k^+} -1.$$

On a donc

$$\begin{aligned} & \phi_{s_2}^{-1} \circ \phi_{s_1} (]a_k; b_k[) =]-1; 1[\\ \Rightarrow U_{s_2} = \phi_{s_2} (]-1; 1]) &= \phi_{s_2} (\phi_{s_2}^{-1} \circ \phi_{s_1} (]a_k; b_k[)) = \phi_{s_1} (]a_k; b_k[) \subset U_{s_1}. \end{aligned}$$

Nous avons ainsi montré que, s'il existait $k \in E$ tel que $-1 < a_k < b_k < 1$, alors $U_{s_2} \subset U_{s_1}$: on est dans le cas **2a** de l'énoncé du lemme. Supposons maintenant qu'il n'existe pas $k \in E$ tel que $-1 < a_k < b_k < 1$. Cela signifie que, pour tout $l \in E$, $a_l = -1$ ou $b_l = 1$ (ou les deux). Compte tenu du fait que les intervalles $]a_l; b_l[$ sont disjoints, on a cinq possibilités :

- (i) $\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2}) = \emptyset$;
- (ii) $\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2}) =]-1; 1[$;
- (iii) $\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2}) =]-1; \alpha[$ pour un certain $\alpha \in]-1; 1[$;
- (iv) $\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2}) =]\alpha; 1[$ pour un certain $\alpha \in]-1; 1[$;
- (v) $\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2}) =]-1; \alpha \cup]\beta; 1[$, avec $\alpha, \beta \in]-1; 1[$, $\alpha < \beta$.

Dans le cas (i), on doit avoir $U_{s_1} \cap U_{s_2} = \emptyset$ (puisque ϕ_{s_1} est surjective vers U_{s_1}); on est alors dans le cas **1** de l'énoncé.

Dans le cas (ii), on a que

$$U_{s_1} = \phi_{s_1} (]-1; 1]) = \phi_{s_1} (\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2})) = U_{s_1} \cap U_{s_2}$$

donc $U_{s_1} \subset U_{s_2}$: on est dans le cas **2a** de l'énoncé.

Dans le cas (iii) ou (iv), $U_{s_1} \cap U_{s_2}$ a exactement une composante connexe (voir Proposition A.7); dans le cas (v), $U_{s_1} \cap U_{s_2}$ a deux composantes connexes. On est donc, respectivement, dans le cas **2b** ou **3** de l'énoncé. (On observe que le raisonnement que nous avons fait pour $\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2})$ est aussi valable pour $\phi_{s_2}^{-1}(U_{s_1} \cap U_{s_2})$: cet ensemble est aussi de la forme $]-1; \alpha[$ ou $]\alpha; 1[$ si $U_{s_1} \cap U_{s_2}$ a une seule composante connexe et $U_{s_1} \not\subset U_{s_2}, U_{s_2} \not\subset U_{s_1}$, et de la forme $]-1; \alpha \cup]\alpha; \beta[$ si $U_{s_1} \cap U_{s_2}$ a deux composantes connexes.)

□

Démonstration du lemme 3.8.

Première étape : commençons par admettre que $U_{s_1} \cup U_{s_2}$ est C^k -difféomorphe à \mathbb{S}^1 . Alors $U_{s_1} \cup U_{s_2}$ est un sous-ensemble ouvert et fermé de M (ouvert car c'est une union d'ouverts, fermé car il est homéomorphe à un ensemble compact, donc compact). Comme M est connexe et $U_{s_1} \cup U_{s_2}$ est non-vide, on doit avoir (d'après la proposition A.2)

$$M = U_{s_1} \cup U_{s_2}.$$

Donc M est C^k -difféomorphe à \mathbb{S}^1 .

Deuxième étape : montrons que $U_{s_1} \cup U_{s_2}$ est C^k -difféomorphe à \mathbb{S}^1 .

Soient C_1, C_2 les deux composantes connexes de $U_{s_1} \cap U_{s_2}$. Puisqu'on est dans le cas 3 du lemme 3.7, il existe α_1, β_1 tels que

$$\begin{aligned} \phi_{s_1}^{-1}(C_1) &=]-1; \alpha_1[\quad \text{et} \quad \phi_{s_1}^{-1}(C_2) =]\beta_1; 1[\\ \text{ou} \quad \phi_{s_1}^{-1}(C_1) &=]\beta_1; 1[\quad \text{et} \quad \phi_{s_1}^{-1}(C_2) =]-1; \alpha_1[. \end{aligned} \quad (3.1)$$

Quitte à échanger C_1 et C_2 , on peut supposer que c'est l'équation (3.1) qui est vraie. De même, il existe α_2, β_2 tels que

$$\begin{aligned} \phi_{s_2}^{-1}(C_1) &=]-1; \alpha_2[\quad \text{et} \quad \phi_{s_2}^{-1}(C_2) =]\beta_2; 1[\\ \text{ou} \quad \phi_{s_2}^{-1}(C_1) &=]\beta_2; 1[\quad \text{et} \quad \phi_{s_2}^{-1}(C_2) =]-1; \alpha_2[. \end{aligned} \quad (3.2)$$

Quitte à remplacer ϕ_{s_2} par $\tilde{\phi}_{s_2} : t \in]-1; 1[\rightarrow \phi_{s_2}(-t)$ (qui est aussi un C^k -difféomorphisme de $] - 1; 1[$ vers U_{s_2}), on peut supposer que c'est l'équation (3.2) qui est vraie.

Proposition 3.10

L'application $\phi_{s_2}^{-1} \circ \phi_{s_1}$ réalise un C^k -difféomorphisme décroissant de $] - 1; \alpha_1[$ vers $] - 1; \alpha_2[$, ainsi que de $] \beta_1; 1[$ vers $] \beta_2; 1[$.

Démonstration. Faisons la démonstration pour les intervalles $] - 1; \alpha_1[$ et $] - 1; \alpha_2[$; elle est identique pour $] \beta_1; 1[$ et $] \beta_2; 1[$.

Puisque ϕ_{s_1} réalise un C^k -difféomorphisme de $] - 1; \alpha_1[$ vers C_1 et $\phi_{s_2}^{-1}$ un C^k -difféomorphisme de C_1 vers $] - 1; \alpha_2[$, l'application $\phi_{s_2}^{-1} \circ \phi_{s_1}$ réalise un C^k -difféomorphisme de $] - 1; \alpha_1[$ vers $] - 1; \alpha_2[$. Montrons qu'il est décroissant.

Comme un difféomorphisme entre deux intervalles est toujours strictement monotone, il suffit de montrer qu'il n'est pas croissant. Supposons par l'absurde qu'il est croissant. Alors

$$\phi_{s_2}^{-1} \circ \phi_{s_1}(t) \xrightarrow{t \rightarrow \alpha_1} \alpha_2,$$

d'où

$$\phi_{s_2}(\alpha_2) = \phi_{s_2} \left(\lim_{t \rightarrow \alpha_1} \phi_{s_2}^{-1} \circ \phi_{s_1}(t) \right) = \lim_{t \rightarrow \alpha_1} \phi_{s_1}(t) = \phi_{s_1}(\alpha_1)$$

et donc

$$\phi_{s_1}(\alpha_1) \in U_{s_1} \cap U_{s_2},$$

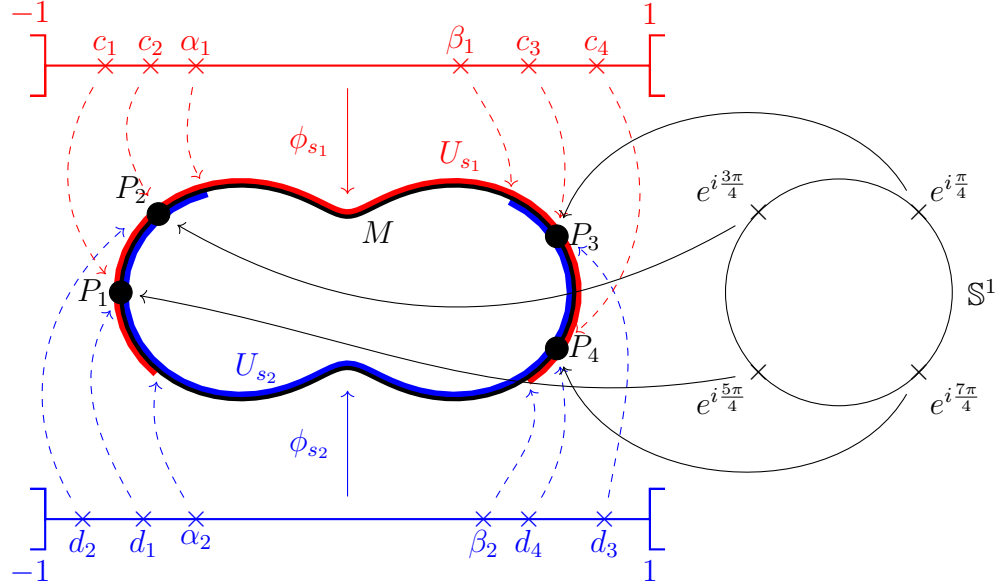


FIGURE 3.3 – Illustration des notations du lemme 3.8 et représentation schématique du difféomorphisme de \mathbb{S}^1 vers M ($e^{i\frac{\pi}{4}}$ est envoyé sur P_3 etc.).

ce qui est en contradiction avec le fait que $\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2}) =]-1; \alpha_1[\cup]\beta_1; 1[$ et ne contient donc pas α_1 . Donc il est impossible que $\phi_{s_2}^{-1} \circ \phi_{s_1}$ soit croissante. \square

Fixons quatre réels c_1, c_2, c_3, c_4 tels que $-1 < c_1 < c_2 < \alpha_1$ et $\beta_1 < c_3 < c_4 < 1$ (voir la figure 3.3 pour une illustration des notations). On note, pour tout $k = 1, 2, 3, 4$,

$$P_k = \phi_{s_1}(c_k) \quad \text{et} \quad d_k = \phi_{s_2}^{-1}(P_k) = \phi_{s_2}^{-1}(\phi_{s_1}(c_k))$$

Comme c_1, c_2 appartiennent à $] - 1; \alpha_1[$ et $c_1 < c_2$, la proposition 3.10 implique que d_1, d_2 appartiennent à $] - 1; \alpha_2[$ et $d_2 < d_1$. De même, d_3, d_4 appartiennent à $] \beta_2; 1[$ et $d_4 < d_3$. On remarque (cela servira à la fin) qu'en vertu à nouveau de la proposition 3.10 :

$$\begin{aligned} \phi_{s_1}(]-1; c_1]) &= \phi_{s_2}(\phi_{s_2}^{-1} \circ \phi_{s_1}(]-1; c_1])) = \phi_{s_2}([d_1; \alpha_2]), \\ \phi_{s_1}([c_1; c_2]) &= \phi_{s_2}([d_2; d_1]), \quad \phi_{s_1}([c_2; \alpha_1]) = \phi_{s_2}(]-1; d_2]), \\ \phi_{s_1}(] \beta_1; c_3]) &= \phi_{s_2}([d_3; 1]), \\ \phi_{s_1}([c_3; c_4]) &= \phi_{s_2}([d_4; d_3]), \quad \phi_{s_1}([c_4; 1]) = \phi_{s_2}(] \beta_2; d_4]). \end{aligned} \tag{3.3}$$

Construisons maintenant un C^k -difféomorphisme $\psi : \mathbb{S}^1 \rightarrow M$. Nous allons imposer, comme indiqué sur la figure 3.3,

$$\psi\left(e^{i\frac{\pi}{4}}\right) = P_3, \quad \psi\left(e^{i\frac{3\pi}{4}}\right) = P_2, \quad \psi\left(e^{i\frac{5\pi}{4}}\right) = P_1, \quad \psi\left(e^{i\frac{7\pi}{4}}\right) = P_4. \quad (3.4)$$

Nous allons donner pour ψ une définition par morceaux :

$$\psi(e^{i\theta}) = \phi_{s_1}(\delta_{s_1}(\theta)) \text{ pour tout } \theta \in \left[-\frac{\pi}{4}; \frac{5\pi}{4}\right]; \quad (3.5a)$$

$$\psi(e^{i\theta}) = \phi_{s_2}(\delta_{s_2}(\theta)) \text{ pour tout } \theta \in \left[\frac{3\pi}{4}; \frac{9\pi}{4}\right], \quad (3.5b)$$

avec $\delta_{s_1} : \left[-\frac{\pi}{4}; \frac{5\pi}{4}\right] \rightarrow]-1; 1[$ et $\delta_{s_2} : \left[\frac{3\pi}{4}; \frac{9\pi}{4}\right] \rightarrow]-1; 1[$ des applications bien choisies.

On commence par choisir δ_{s_1} . Soit δ_{s_1} un C^∞ -difféomorphisme de $\left[-\frac{\pi}{4}; \frac{5\pi}{4}\right]$ vers $[c_1; c_4]$ tel que

$$\delta_{s_1}\left(-\frac{\pi}{4}\right) = c_4, \quad \delta_{s_1}\left(\frac{\pi}{4}\right) = c_3, \quad \delta_{s_1}\left(\frac{3\pi}{4}\right) = c_2, \quad \delta_{s_1}\left(\frac{5\pi}{4}\right) = c_1. \quad (3.6)$$

(Ces égalités permettent de garantir que l'équation (3.4) est vérifiée.) Un tel difféomorphisme existe (voir la proposition B.3 en annexe).

Définissons maintenant δ_{s_2} . Les définitions des lignes (3.5a) et (3.5b) doivent coïncider sur les points où elles donnent toutes les deux une valeur à ψ . On doit donc avoir, pour tout $\theta \in \left[\frac{3\pi}{4}; \frac{5\pi}{4}\right]$,

$$\phi_{s_1}(\delta_{s_1}(\theta)) = \phi_{s_2}(\delta_{s_2}(\theta))$$

et, pour tout $\theta \in \left[\frac{7\pi}{4}; \frac{9\pi}{4}\right]$,

$$\phi_{s_1}(\delta_{s_1}(\theta - 2\pi)) = \phi_{s_2}(\delta_{s_2}(\theta)).$$

Définissons donc

$$\delta_{s_2}(\theta) = \phi_{s_2}^{-1}(\phi_{s_1}(\delta_{s_1}(\theta))) \text{ pour tout } \theta \in \left[\frac{3\pi}{4}; \frac{5\pi}{4}\right], \quad (3.7a)$$

$$\delta_{s_2}(\theta) = \phi_{s_2}^{-1}(\phi_{s_1}(\delta_{s_1}(\theta - 2\pi))) \text{ pour tout } \theta \in \left[\frac{7\pi}{4}; \frac{9\pi}{4}\right]. \quad (3.7b)$$

On peut vérifier que les quantités ci-dessus sont bien définies, grâce aux égalités de l'équation (3.6), qui entraînent que $\delta_{s_1}(\theta)$ et $\delta_{s_1}(\theta - 2\pi)$ appartiennent à

$] - 1; \alpha_1[\cup]\beta_1; 1[$. Avec ces définitions, δ_{s_2} réalise déjà un C^∞ -difféomorphisme entre $[\frac{3\pi}{4}; \frac{5\pi}{4}]$ et

$$\left[\phi_{s_2}^{-1} \left(\phi_{s_1} \left(\delta_{s_1} \left(\frac{3\pi}{4} \right) \right) \right); \phi_{s_2}^{-1} \left(\phi_{s_1} \left(\delta_{s_1} \left(\frac{5\pi}{4} \right) \right) \right) \right] = [d_2; d_1]$$

et entre $[\frac{7\pi}{4}; \frac{9\pi}{4}]$ et

$$\left[\phi_{s_2}^{-1} \left(\phi_{s_1} \left(\delta_{s_1} \left(-\frac{\pi}{4} \right) \right) \right); \phi_{s_2}^{-1} \left(\phi_{s_1} \left(\delta_{s_1} \left(\frac{\pi}{4} \right) \right) \right) \right] = [d_4; d_3].$$

Sur $[\frac{5\pi}{4}; \frac{7\pi}{4}]$, on définit δ_{s_2} comme n'importe quel C^∞ -difféomorphisme croissant de $[\frac{5\pi}{4}; \frac{7\pi}{4}]$ vers $[d_1; d_4]$ dont les dérivées jusqu'à l'ordre k aux extrémités du segment sont compatibles avec celles des définitions (3.7a) et (3.7b) : pour tout $k' = 1, \dots, k$,

$$\begin{aligned} \delta_{s_2}^{(k')} \left(\frac{5\pi}{4} \right) &= (\phi_{s_2}^{-1} \circ \phi_{s_1} \circ \delta_{s_1})^{(k')} \left(\frac{5\pi}{4} \right), \\ \delta_{s_2}^{(k')} \left(\frac{7\pi}{4} \right) &= (\phi_{s_2}^{-1} \circ \phi_{s_1} \circ \delta_{s_1})^{(k')} \left(-\frac{\pi}{4} \right). \end{aligned}$$

Un tel difféomorphisme existe (voir la proposition B.4 en annexe). Avec ces définitions, δ_{s_2} est un C^k -difféomorphisme de $[\frac{3\pi}{4}; \frac{9\pi}{4}]$ vers $[d_2; d_3]$.

Nous avons maintenant fini de définir ψ , conformément aux équations (3.5a) et (3.5b). Vérifions que cette définition en fait bien un C^k -difféomorphisme de \mathbb{S}^1 vers $U_{s_1} \cup U_{s_2}$. Tout d'abord, c'est bien une application de classe C^k : elle est C^k sur $\{e^{i\theta}, \theta \in]-\frac{\pi}{4}; \frac{5\pi}{4}[\}$ car $\phi_{s_1} \circ \delta_{s_1}$ l'est, et elle est C^k sur $\{e^{i\theta}, \theta \in]\frac{3\pi}{4}; \frac{9\pi}{4}[\}$ car $\phi_{s_2} \circ \delta_{s_2}$ l'est. Elle est donc C^k sur la réunion de ces deux ensembles, qui est \mathbb{S}^1 tout entier.

La proposition qui suit garantit qu'elle est bijective de \mathbb{S}^1 vers $U_{s_1} \cup U_{s_2}$.

Proposition 3.11

L'application ψ réalise une bijection de \mathbb{S}^1 vers $U_{s_1} \cup U_{s_2}$, dont la réciproque vaut :

$$\begin{aligned} \zeta(x) &= e^{i\delta_{s_1}^{-1}(\phi_{s_1}^{-1}(x))} \quad \text{pour tout } x \in \phi_{s_1}([c_1; c_4]), \\ &= e^{i\delta_{s_2}^{-1}(\phi_{s_2}^{-1}(x))} \quad \text{pour tout } x \in \phi_{s_2}([d_2; d_3]). \end{aligned}$$

Démonstration. L'application ψ est surjective vers $U_{s_1} \cup U_{s_2}$. En effet, d'après sa définition (équations (3.5a) et (3.5b)),

$$\begin{aligned}\psi(\mathbb{S}^1) &= \phi_{s_1} \left(\delta_{s_1} \left(\left[-\frac{\pi}{4}; \frac{5\pi}{4} \right] \right) \right) \cup \phi_{s_2} \left(\delta_{s_2} \left(\left[\frac{3\pi}{4}; \frac{9\pi}{4} \right] \right) \right) \\ &= \phi_{s_1}([c_1; c_4]) \cup \phi_{s_2}([d_2; d_3])\end{aligned}$$

Or

$$\begin{aligned}U_{s_1} \cup U_{s_2} &= \phi_{s_1}(] - 1; 1[) \cup \phi_{s_2}(] - 1; 1[) \\ &= \phi_{s_1}(] - 1; c_1[) \cup \phi_{s_1}(]c_1; c_4[) \cup \phi_{s_1}([c_4; 1[) \\ &\quad \cup \phi_{s_2}(] - 1; d_2[) \cup \phi_{s_2}(]d_2; d_3[) \cup \phi_{s_2}([d_3; 1[) \\ &= \phi_{s_2}([d_1; \alpha_2[) \cup \phi_{s_1}(]c_1; c_4[) \cup \phi_{s_2}(] \beta_2; d_4[) \\ &\quad \cup \phi_{s_1}([c_2; \alpha_2[) \cup \phi_{s_2}(]d_2; d_3[) \cup \phi_{s_1}(] \beta_1; c_3[) \\ &\quad (\text{par l'équation (3.3)}) \\ &\subset \phi_{s_1}(]c_1; c_4[) \cup \phi_{s_2}(]d_2; d_3[) \\ &\subset U_{s_1} \cup U_{s_2},\end{aligned}$$

d'où $\phi_{s_1}([c_1; c_4]) \cup \phi_{s_2}([d_2; d_3]) = U_{s_1} \cup U_{s_2}$.

D'autre part, elle est injective. Pour le montrer, supposons donnés $\theta, \theta' \in \mathbb{R}$ tels que

$$\psi(e^{i\theta}) = \psi(e^{i\theta'})$$

et montrons que $e^{i\theta} = e^{i\theta'}$. Tout d'abord, si θ et θ' appartiennent tous deux à $[-\frac{\pi}{4}; \frac{5\pi}{4}]$ (modulo 2π), alors, d'après la définition (3.5a) et l'injectivité de ϕ_{s_1} et δ_{s_1} ,

$$\theta \equiv \theta' [2\pi] \quad \Rightarrow \quad e^{i\theta} = e^{i\theta'}.$$

De même, si θ et θ' appartiennent tous deux à $[\frac{3\pi}{4}; \frac{9\pi}{4}]$ modulo 2π , alors $e^{i\theta} = e^{i\theta'}$. Supposons maintenant qu'on n'est dans aucune des deux situations précédentes, par exemple que θ appartient à $[-\frac{\pi}{4}; \frac{5\pi}{4}]$ mais pas à $[\frac{3\pi}{4}; \frac{9\pi}{4}]$ (c'est-à-dire que θ appartient à $] \frac{\pi}{4}; \frac{3\pi}{4} [$) et θ' appartient à $[\frac{3\pi}{4}; \frac{9\pi}{4}]$ mais pas à $[-\frac{\pi}{4}; \frac{5\pi}{4}]$ (c'est-à-dire que θ' appartient à $] \frac{5\pi}{4}; \frac{7\pi}{4} [$). Alors

$$\begin{aligned}\psi(e^{i\theta}) &\in \phi_{s_1} \left(\delta_{s_1} \left(\left[\frac{\pi}{4}; \frac{3\pi}{4} \right] \right) \right) = \phi_{s_1}(]c_2; c_3[) \\ \psi(e^{i\theta'}) &\in \phi_{s_2} \left(\delta_{s_2} \left(\left[\frac{5\pi}{4}; \frac{7\pi}{4} \right] \right) \right) = \phi_{s_2}(]d_1; d_4[).\end{aligned}$$

Mais $\phi_{s_1}(]c_2; c_3[)$ et $\phi_{s_2}(]d_1; d_4[)$ sont d'intersection vide (voir la figure 3.3 ; cela se vérifie avec l'équation (3.3)). On ne peut donc pas avoir $\psi(e^{i\theta}) = \psi(e^{i\theta'})$: ce cas est impossible. Cela finit de démontrer l'injectivité.

Nous avons donc montré que ψ est une bijection. La formule de la réciproque est une conséquence de la définition de ψ , aux équations (3.5a) et (3.5b). \square

Enfin, comme $\psi^{-1} = \zeta$ est de classe C^k (les fonctions $\delta_{s_1}, \delta_{s_2}, \phi_{s_1}, \phi_{s_2}$ le sont), ψ est un C^k -difféomorphisme. \square

Démonstration du lemme 3.9. La démonstration étant assez similaire à celle du lemme 3.8, on n'en donnera que les grandes lignes ici.

On suppose que U_{s_1}, U_{s_2} vérifient la propriété 2 du lemme 3.7. Si $U_{s_1} \subset U_{s_2}$, alors $U_{s_1} \cup U_{s_2} = U_{s_2}$ est C^k -difféomorphe à $] - 1; 1[$, d'après nos hypothèses sur U_{s_2} . De même si $U_{s_2} \subset U_{s_1}$.

On peut donc supposer que c'est plutôt la sous-propriété 2b qui est vérifiée : $\phi_{s_1}^{-1}(U_{s_1} \cap U_{s_2})$ et $\phi_{s_2}^{-1}(U_{s_1} \cap U_{s_2})$ sont de la forme $] - 1; \alpha[$ ou $]\alpha; 1[$. On peut supposer qu'ils valent respectivement $]\alpha_1; 1[$ et $]\alpha_2; 1[$ pour des réels $\alpha_1, \alpha_2 \in] - 1; 1[$ (voir la figure 3.4 pour une illustration des notations).

Soient $c_1, c_2 \in]\alpha_1; 1[$ tels que $c_1 < c_2$. On note

$$\begin{aligned} P_1 &= \phi_{s_1}(c_1), & P_2 &= \phi_{s_1}(c_2), \\ d_1 &= \phi_{s_2}^{-1}(P_1), & d_2 &= \phi_{s_2}^{-1}(P_2). \end{aligned}$$

Comme $\phi_{s_2}^{-1} \circ \phi_{s_1}$ réalise un C^k -difféomorphisme décroissant de $]\alpha_1; 1[$ vers $]\alpha_2; 1[$ (pour les mêmes raisons qu'à la proposition 3.10), on a $\alpha_2 < d_2 < d_1 < 1$.

On définit $\psi :] - 1; 1[\rightarrow U_{s_1} \cup U_{s_2}$ par

$$\psi(x) = \phi_{s_1}(\delta_{s_1}(x)) \text{ pour tout } x \in \left] -1; \frac{1}{2} \right] \quad (3.8a)$$

$$= \phi_{s_2}(\delta_{s_2}(x)) \text{ pour tout } x \in \left[-\frac{1}{2}; 1 \right[, \quad (3.8b)$$

où δ_{s_1} est un C^∞ difféomorphisme de $] -1; \frac{1}{2}]$ vers $] -1; c_2]$ tel que

$$\delta_{s_1} \left(-\frac{1}{2} \right) = c_1, \quad \delta_{s_1} \left(\frac{1}{2} \right) = c_2,$$

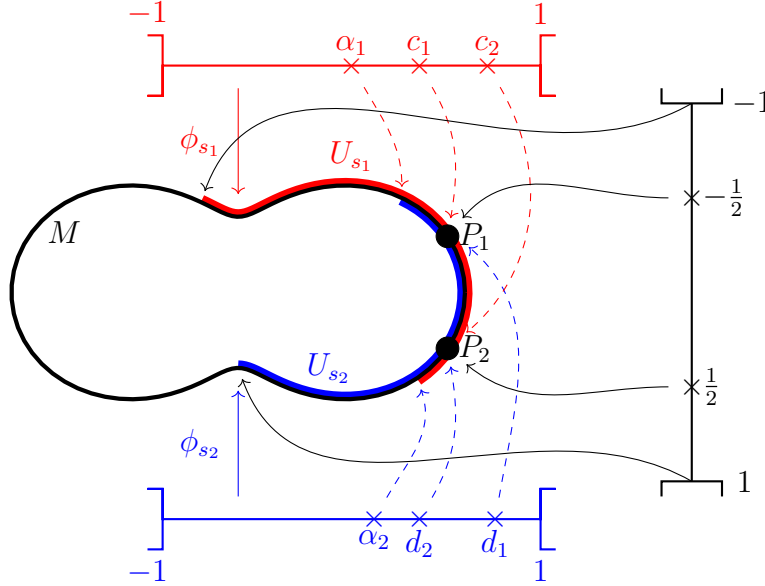


FIGURE 3.4 – Illustration des notations du lemme 3.9 et représentation schématique du difféomorphisme de $] - 1; 1[$ vers $U_{s_1} \cup U_{s_2}$.

et δ_{s_2} est un C^k -difféomorphisme décroissant de $[-\frac{1}{2}; 1[$ vers $] - 1; d_1]$ tel que, sur $[-\frac{1}{2}; \frac{1}{2}]$,

$$\delta_{s_2} = \phi_{s_2}^{-1} \circ \phi_{s_1} \circ \delta_{s_1}$$

et, sur $[\frac{1}{2}; 1[$, δ_{s_2} est n'importe quel C^k -difféomorphisme décroissant de $[\frac{1}{2}; 1[$ vers $] - 1; d_2]$ tel que, pour tout $k' = 1, \dots, k$.

$$\delta_{s_2}^{(k')} \left(\frac{1}{2} \right) = (\phi_{s_2}^{-1} \circ \phi_{s_1} \circ \delta_{s_1})^{(k')} \left(\frac{1}{2} \right).$$

L'existence de $\delta_{s_1}, \delta_{s_2}$ est assurée par les propositions B.3 et B.4. Avec ces définitions pour $\delta_{s_1}, \delta_{s_2}$, la définition de ψ aux équations (3.8a) et (3.8b) est valide. De plus, la fonction ψ est de classe C^k .

Le même raisonnement qu'à la proposition 3.11 permet de montrer que ψ est une bijection entre $] - 1; 1[$ et $U_{s_1} \cup U_{s_2}$. Sa réciproque vaut

$$\begin{aligned} \zeta(x) &= \delta_{s_1}^{-1}(\phi_{s_1}^{-1}(x)) \quad \text{pour tout } x \in \phi_{s_1}^{-1}(] - 1; c_2]), \\ &= \delta_{s_2}^{-1}(\phi_{s_2}^{-1}(x)) \quad \text{pour tout } x \in \phi_{s_2}^{-1}(] - 1; d_1]). \end{aligned}$$

Comme cette réciproque est C^k , ψ est donc un C^k -difféomorphisme entre $] - 1; 1[$ et $U_{s_1} \cup U_{s_2}$.

□

3.1.3 Longueur et abscisse curviligne

Nous allons maintenant définir la *longueur* d'une courbe. Intuitivement, de quoi s'agit-il? Notons (I, γ) un paramétrage global de la courbe et imaginons une fourmi qui parcourt l'arc : à l'instant t , la fourmi se trouve au point $\gamma(t)$. La longueur de l'arc est la distance totale parcourue par la fourmi au cours du temps. Comme, à un instant t , sa vitesse absolue est $\|\gamma'(t)\|_2$, la longueur doit être définie comme l'intégrale sur I de $\|\gamma'\|_2$.

Définition 3.12 : longueur d'une courbe

Soit M une courbe connexe. Soit (I, γ) un paramétrage global de M . On définit la *longueur* de M comme

$$\ell(M) = \int_I \|\gamma'(t)\|_2 dt.$$

Proposition 3.13

La définition précédente « a un sens » : si (I, γ) et (J, δ) sont deux paramétrages globaux de M , alors

$$\int_I \|\gamma'(t)\|_2 dt = \int_J \|\delta'(t)\|_2 dt.$$

Démonstration. Traitons le cas où M est non-compacte. Alors γ et δ sont des difféomorphismes de (respectivement) I et J vers M . Notons

$$\theta = \gamma^{-1} \circ \delta : J \rightarrow I.$$

C'est un difféomorphisme de J vers I et on a $\delta = \gamma \circ \theta$. Alors

$$\begin{aligned} \int_J \|\delta'(t)\|_2 dt &= \int_J \|(\gamma \circ \theta)'(t)\|_2 dt \\ &= \int_J |\theta'(t)| \|\gamma' \circ \theta(t)\|_2 dt \\ &= \int_I \|\gamma'(t)\|_2 dt. \end{aligned}$$

La dernière égalité provient de la formule du changement de variable appliquée à la fonction $(s \in I \rightarrow \|\gamma'(s)\|_2 \in \mathbb{R})$, avec le changement de variable donné par θ .

On omet le cas où M est compacte. Le principe est identique, avec une subtilité liée au fait que γ et δ ne sont pas exactement des difféomorphismes de leur intervalle de définition vers M .⁴ \square

Définition 3.14 : abscisse curviligne

Un paramétrage global (I, γ) d'une courbe connexe M est appelé *abscisse curviligne* si

$$\|\gamma'(t)\|_2 = 1, \quad \forall t \in I.$$

On remarque que, si (I, γ) est une abscisse curviligne de M , alors la longueur de M est égale à la longueur de I :

$$\ell(M) = \int_I 1 dt = \sup I - \inf I.$$

Théorème 3.15 : existence d'une abscisse curviligne

Toute courbe connexe M admet une abscisse curviligne.

Démonstration. Traitons le cas où M n'est pas compacte (à part que les notations y sont légèrement différentes, le cas compact est identique). Soit $\phi : \mathbb{R} \rightarrow M$ un C^k -difféomorphisme. On va chercher une abscisse curviligne sous la forme $(I, \phi \circ \theta)$ avec I un intervalle ouvert contenant 0 et $\theta : I \rightarrow \mathbb{R}$ un C^k -difféomorphisme croissant tel que $\theta(0) = 0$.

Pour que $(I, \phi \circ \theta)$ soit une abscisse curviligne, il faut que, pour tout $t \in I \cap \mathbb{R}_+^*$,

$$t = \text{longueur}(]0; t]) = \ell(\phi \circ \theta(]0; t]))$$

4. Pour les lecteurs/trices particulièrement curieux/ses, voici comment résoudre cette difficulté. Notons a, b, c, d les réels tels que $I = [a; b[$ et $J = [c; d[$. Notons $\alpha \in [0; d - c[$ tel que $\gamma(a) = \delta(c + \alpha)$. En remplaçant (ce qui ne change pas l'intégrale de $\|\delta'\|$) (J, δ) par $(\tilde{J}, \tilde{\delta})$, où $\tilde{J} = [c + \alpha; d + \alpha[$ et $\tilde{\delta} = \delta$ sur $[c + \alpha; d[$ et $\tilde{\delta} = \delta(\cdot - (d - c))$ ailleurs, on peut supposer que $\gamma(a) = \delta(c)$. Alors γ et δ réalisent des difféomorphismes de $]a; b[$ vers $M - \{\gamma(a)\}$. On peut définir, comme dans le cas non-compact,

$$\theta = \gamma^{-1} \circ \delta :]c; d[\rightarrow]a; b[$$

et procéder de la même manière que précédemment.

$$\begin{aligned}
&= \ell(\phi([\theta(0); \theta(t)]) \\
&= \int_0^{\theta(t)} \|\phi'(s)\|_2 ds. \tag{3.9}
\end{aligned}$$

On vérifie que la même équation doit être valable pour $t \in I \cap \mathbb{R}^-$.

Définissons alors

$$\begin{aligned}
L : \mathbb{R} &\rightarrow \mathbb{R} \\
T &\rightarrow \int_0^T \|\phi'(s)\|_2 ds.
\end{aligned}$$

Il s'agit d'une application de classe C^k dont la dérivée ne s'annule pas : elle réalise un C^k -difféomorphisme entre \mathbb{R} et son image, qui est un intervalle ouvert. Soit I cette image. Définissons, comme l'impose l'équation (3.9),

$$\theta = L^{-1} : I \rightarrow \mathbb{R}.$$

Avec cette définition, $(I, \phi \circ \theta)$ est un paramétrage global de M . Pour tout $t \in I$,

$$\begin{aligned}
(\phi \circ \theta)'(t) &= \theta'(t)\phi'(\theta(t)) \\
&= (L^{-1})'(t)\phi'(\theta(t)) \\
&= \frac{\phi'(\theta(t))}{L'(L^{-1}(t))} \\
&= \frac{\phi'(\theta(t))}{L'(\theta(t))} \\
&= \frac{\phi'(\theta(t))}{\|\phi'(\theta(t))\|_2}.
\end{aligned}$$

Ce vecteur est toujours de norme 1 : $(I, \phi \circ \theta)$ est une abscisse curviligne. \square

La notion d'abscisse curviligne permet de définir assez simplement un certain nombre de quantités qui décrivent la « forme locale » des courbes. Nous n'avons pas le temps de les présenter en détail dans ce cours mais, pour la culture, citons-en quelques unes. Si (I, γ) est une abscisse curviligne, on appelle, pour tout $t \in I$,

$$\gamma'(t)$$

le *vecteur tangent unitaire* au point $\gamma(t)$. Si γ est de classe C^2 , on appelle

$$\frac{\gamma''(t)}{\|\gamma''(t)\|_2}$$

le *vecteur normal principal unitaire* en $\gamma(t)$ (qui n'est bien défini que si $\gamma''(t) \neq 0$) et

$$\|\gamma''(t)\|_2$$

la *courbure* en $\gamma(t)$ (qui peut être munie d'un signe, positif ou négatif, lorsque la courbe est une sous-variété de \mathbb{R}^2). Informellement, la courbure caractérise à quel point la courbe « tourne » vite au voisinage du point $\gamma(t)$.

3.2 Sous-variétés de dimension générale

Dans cette section, un certain nombre de démonstrations sont reportées à l'annexe, afin de faciliter la lecture. Le lecteur ou la lectrice est bien sûr encouragé-e à aller les lire dans l'annexe une fois qu'il ou elle aura bien compris toutes les notions et propriétés énoncées dans cette section.

3.2.1 Distance et géodésiques

On va maintenant utiliser la notion de longueur vue à la définition 3.12 pour définir une distance sur toute sous-variété connexe M de \mathbb{R}^n : la distance entre deux points x_1, x_2 est l'infimum des longueurs des chemins qui relient ces points.

Dans cette section, on appelle *chemin* reliant deux points x_1 et x_2 toute application $\gamma : [0; A] \rightarrow M$, pour un certain $A \in \mathbb{R}^+$, telle que

- γ est continue ;
- γ est C^1 par morceaux ;
- $\gamma(0) = x_1$ et $\gamma(A) = x_2$.

On peut étendre la définition 3.12 des courbes aux chemins : la *longueur* d'un chemin γ est

$$\ell(\gamma) = \int_0^A \|\gamma'(t)\|_2 dt.$$

Définition 3.16 : distance sur une sous-variété

Soit M une sous-variété connexe de \mathbb{R}^n . On définit une distance sur M comme suit : pour tous $x_1, x_2 \in M$,

$$\text{dist}_M(x_1, x_2) = \inf\{\ell(\gamma), \gamma \text{ est un chemin reliant } x_1 \text{ et } x_2\}.$$

Proposition 3.17

La fonction dist_M est bien définie : pour tous x_1, x_2 , il existe un chemin reliant x_1 et x_2 .

Démonstration. Voir la section C.1. □

Proposition 3.18

La fonction dist_M est bien une distance.

Démonstration.

- Symétrie : soient $x_1, x_2 \in M$. Soit $(\gamma_n)_{n \in \mathbb{N}}$ une suite de chemins reliant x_1 à x_2 telle que

$$\ell(\gamma_n) \xrightarrow{n \rightarrow +\infty} \text{dist}_M(x_1, x_2).$$

Pour tout n , notons $[0; A_n]$ l'intervalle de définition de γ_n et définissons

$$\begin{aligned} \delta_n : [0; A_n] &\rightarrow M \\ t &\rightarrow \gamma_n(A_n - t). \end{aligned}$$

C'est un chemin reliant x_2 à x_1 . De plus, pour tout n ,

$$\ell(\delta_n) = \int_0^{A_n} \|\dot{\gamma}_n(A_n - t)\|_2 dt = \int_0^{A_n} \|\dot{\gamma}_n(t)\|_2 dt = \ell(\gamma_n),$$

donc $\text{dist}_M(x_2, x_1) \leq \ell(\delta_n) = \ell(\gamma_n)$. Par passage à la limite quand $n \rightarrow +\infty$, on en déduit

$$\text{dist}_M(x_2, x_1) \leq \text{dist}_M(x_1, x_2).$$

Le raisonnement qu'on vient de faire reste vrai si on échange x_1 et x_2 . On a donc aussi

$$\text{dist}_M(x_1, x_2) \leq \text{dist}_M(x_2, x_1).$$

D'où $\text{dist}_M(x_1, x_2) = \text{dist}_M(x_2, x_1)$.

- Inégalité triangulaire : soient $x_1, x_2, x_3 \in M$. Montrons l'inégalité

$$\text{dist}_M(x_1, x_3) \leq \text{dist}_M(x_1, x_2) + \text{dist}_M(x_2, x_3).$$

Soient $(\gamma_n : [0; A_n] \rightarrow M)_{n \in \mathbb{N}}$ et $(\delta_n : [0; B_n] \rightarrow M)_{n \in \mathbb{N}}$ des suites de chemins reliant, respectivement, x_1 à x_2 et x_2 à x_3 , telles que

$$\begin{aligned} \ell(\gamma_n) &\xrightarrow{n \rightarrow +\infty} \text{dist}_M(x_1, x_2); \\ \ell(\delta_n) &\xrightarrow{n \rightarrow +\infty} \text{dist}_M(x_2, x_3). \end{aligned}$$

Pour tout n , on définit

$$\begin{aligned} \zeta_n : [0; A_n + B_n] &\rightarrow M \\ t &\rightarrow \begin{cases} \gamma_n(t) & \text{si } t \leq A_n \\ \delta_n(t - A_n) & \text{si } A_n < t. \end{cases} \end{aligned}$$

Pour tout n , on a $\zeta_n(0) = x_1$ et $\zeta_n(A_n + B_n) = x_3$. Comme γ_n et δ_n sont continues, ζ_n est continue sur $[0; A_n[$ et sur $]A_n; A_n + B_n]$. Elle est également continue en A_n car elle admet en ce point des limites à gauche et à droite, qui sont identiques :

$$\zeta_n(t) \xrightarrow{t \rightarrow A_n^-} \gamma_n(A_n) = x_2 = \delta_n(0) \xleftarrow{t \rightarrow A_n^+} \zeta_n(t).$$

La fonction ζ_n est donc continue. Elle est de plus C^1 par morceaux car γ_n et δ_n le sont : c'est un chemin. Sa longueur est

$$\begin{aligned} \ell(\zeta_n) &= \int_0^{A_n+B_n} \|\zeta_n'(t)\|_2 dt \\ &= \int_0^{A_n} \|\gamma_n'(t)\|_2 dt + \int_{A_n}^{A_n+B_n} \|\delta_n'(t - A_n)\|_2 dt \\ &= \int_0^{A_n} \|\gamma_n'(t)\|_2 dt + \int_0^{B_n} \|\delta_n'(t)\|_2 dt \\ &= \ell(\gamma_n) + \ell(\delta_n). \end{aligned}$$

Ainsi, pour tout n , $\text{dist}_M(x_1, x_3) \leq \ell(\gamma_n) + \ell(\delta_n)$, ce qui implique par passage à la limite

$$\text{dist}_M(x_1, x_3) \leq \text{dist}_M(x_1, x_2) + \text{dist}_M(x_2, x_3).$$

- Séparation : pour tout $x \in M$, $\text{dist}_M(x, x) = 0$: en choisissant γ constante, de valeur x , on a $\text{dist}_M(x, x) \leq \ell(\gamma) = 0$.

Montrons la réciproque. Pour tous $x_1, x_2 \in M$ et tout chemin γ reliant x_1 à x_2 ,

$$\begin{aligned} \ell(\gamma) &= \int_0^A \|\gamma'(t)\|_2 dt \\ &\geq \left\| \int_0^A \gamma'(t) dt \right\|_2 \quad (\text{par inégalité triangulaire}) \\ &= \left\| [\gamma(t)]_0^A \right\|_2 \\ &= \|x_2 - x_1\|_2. \end{aligned}$$

En conséquent,

$$\text{dist}_M(x_1, x_2) \geq \|x_2 - x_1\|_2.$$

En particulier, si $\text{dist}_M(x_1, x_2) = 0$, on doit aussi avoir $\|x_2 - x_1\|_2 = 0$ et donc $x_1 = x_2$.

□

Théorème 3.19 : existence de chemins minimisants

Soit toujours M une sous-variété de \mathbb{R}^n de classe C^k connexe. On suppose de plus que

- $k \geq 2$;
- M est fermée dans \mathbb{R}^n .

Alors, pour tous $x_1, x_2 \in M$, l'infimum de la définition 3.16 est un minimum : il existe un chemin γ reliant x_1 à x_2 tel que

$$\ell(\gamma) = \text{dist}_M(x_1, x_2).$$

Si γ est un chemin minimisant, comme dans le théorème précédent, il en existe une reparamétrisation $\tilde{\gamma} \stackrel{\text{déf}}{=} \gamma \circ \phi$ qui soit de vitesse constante : pour un certain c ,

$$\|\tilde{\gamma}'(t)\|_2 = c \text{ pour tout } t.$$

(L'argument est le même que pour le théorème 3.15 ; on peut même imposer $c = 1$ si on le souhaite.)

Ces chemins minimisants parcourus à vitesse constante sont caractérisés par une équation différentielle simple, donnée dans un nouveau théorème.

Théorème 3.20 : équation géodésique

On garde les mêmes notations et hypothèses que dans le théorème précédent. Soit $\gamma : [0; A] \rightarrow M$ un chemin reliant x_1 à x_2 , parcouru à vitesse constante, tel que $\ell(\gamma) = \text{dist}_M(x_1, x_2)$. Alors γ est de classe C^2 et

$$\gamma''(t) \in (T_{\gamma(t)}M)^\perp, \quad \forall t \in [0; A]. \quad (3.10)$$

Démonstration simultanée des théorèmes 3.19 et 3.20. Fixons x_1, x_2 et, pour simplifier les notations, définissons $D = \text{dist}_M(x_1, x_2)$.

Soit $(\gamma_N)_{N \in \mathbb{N}}$ une suite quelconque de chemins reliant x_1 à x_2 telle que

$$\ell(\gamma_N) \xrightarrow{n \rightarrow +\infty} D.$$

Quitte à reparamétriser, on peut supposer que les γ_n sont parcourus à vitesse constante $c > 0$ (en choisissant c arbitrairement) :

$$\|\gamma'_N(t)\|_2 = c, \quad \forall N \in \mathbb{N}, \forall t.$$

Après cette reparamétrisation, l'ensemble de définition de l'application γ_N est $[0; \ell(\gamma_N)/c]$.

Un argument de compacité permet de supposer que $(\gamma_N)_{N \in \mathbb{N}}$ converge uniformément vers une certaine limite. C'est ce que dit la proposition qui suit, démontrée dans la section C.2.

Proposition 3.21

Il existe une fonction $\delta : [0; D/c] \rightarrow M$ et une extraction $\rho : \mathbb{N} \rightarrow \mathbb{N}$ telles que

$$\|\gamma_{\rho(N)} - \delta\|_\infty \xrightarrow{n \rightarrow +\infty} 0.$$

De plus,

- δ est c -lipschitzienne ;
- $\delta(0) = x_1$ et $\delta(D/c) = x_2$.

Définissons δ et ρ comme dans la proposition. Quitte à remplacer $(\gamma_N)_{N \in \mathbb{N}}$ par la sous-suite $(\gamma_{\rho(N)})_{n \in \mathbb{N}}$, on peut supposer $\rho = \text{Id}$ (pour simplifier les notations). On va montrer que δ est un chemin reliant x_1 à x_2 (il ne reste

qu'à montrer qu'il est C^1 par morceaux), qu'il est même C^2 , vérifie l'équation (3.10) et

$$\ell(\delta) = D.$$

Cela démontrera directement le théorème 3.19 et impliquera également le théorème 3.20 (car pour tout chemin γ tel que $\ell(\gamma) = D$, on peut appliquer le raisonnement qu'on vient de faire à la suite constante $\gamma_N = \gamma, \forall N \in \mathbb{N}$; le seul point d'adhérence de cette suite est $\delta \stackrel{\text{déf}}{=} \gamma$ donc, si δ vérifie l'équation (3.10), γ la vérifie).

La fin de cette démonstration reste à écrire; elle sera disponible dans une prochaine version de ces notes.

□

Remarque

Le théorème 3.19, qui garantit l'existence d'un chemin de longueur minimale entre deux points arbitraires, n'est plus nécessairement vrai si la sous-variété considérée n'est pas fermée. Par exemple, dans la sous-variété $M \stackrel{\text{déf}}{=} \mathbb{R}^2 \setminus \{(0, 0)\}$, il n'existe pas de chemin minimisant entre $(-1, 0)$ et $(1, 0)$.

Toutefois, lorsque la sous-variété M n'est plus fermée, on peut démontrer (et la démonstration est très similaire à celle qui précède) que tout point $x_1 \in M$ admet un voisinage V tel que, pour tout $x_2 \in V$, il existe un chemin minimisant la longueur entre x_1 et x_2 .

Le théorème 3.20 reste vrai si la sous-variété considérée n'est pas fermée.

Les courbes vérifiant l'équation (3.10), qu'elles soient ou non des chemins de longueur minimale entre deux points, s'appellent les *géodésiques*.

Définition 3.22 : géodésiques

Soit M une sous-variété de \mathbb{R}^n de classe C^k avec $k \geq 2$. On appelle *géodésique* toute application $\gamma : I \rightarrow M$ (pour I un certain intervalle non-vide de \mathbb{R}) de classe C^2 telle que, pour tout $t \in I$,

$$\gamma''(t) \in (T_{\gamma(t)}M)^\perp.$$

Proposition 3.23

Une géodésique γ est toujours de vitesse constante : $\|\gamma'(t)\|_2$ est indépendante de t .

Démonstration. Soit $\gamma : I \rightarrow M$ une géodésique dans une certaine sous-variété M . Définissons

$$N : t \in I \rightarrow \|\gamma'(t)\|_2^2.$$

Cette fonction est dérivable et, pour tout t ,

$$N'(t) = 2 \langle \gamma'(t), \gamma''(t) \rangle.$$

Or, pour tout t , $\gamma'(t) \in T_{\gamma(t)}M$ et, puisque γ est une géodésique, $\gamma''(t) \in (T_{\gamma(t)}M)^\perp$. Donc, pour tout t ,

$$N'(t) = 0,$$

ce qui veut dire que N , et donc aussi $\|\gamma'\|_2$, est une fonction constante. \square

Exemple 3.24

Soient $n \in \mathbb{N}^*$ et $d \in \{1, \dots, n\}$. Les géodésiques de la sous-variété « modèle » $M = \mathbb{R}^d \times \{0\}^{n-d}$, sont les applications $\gamma : I \rightarrow \mathbb{R}^n$ de classe C^2 telles que

1. $\gamma_{d+1}(t) = \dots = \gamma_n(t) = 0$ pour tout $t \in I$ (puisque $\gamma(t) \in M$);
2. $\gamma_1''(t) = \dots = \gamma_d''(t) = 0$ pour tout $t \in I$ (puisque $\gamma''(t) \in (T_{\gamma(t)}M)^\perp = \{0\}^d \times \mathbb{R}^{n-d}$).

Il s'agit donc des applications dont les $n - d$ dernières composantes sont nulles et les d premières sont affines. Les géodésiques sont donc exactement les applications de la forme

$$\gamma : t \in I \rightarrow x_0 + tv,$$

pour n'importe quels $x_0, v \in \mathbb{R}^d \times \{0\}^{n-d}$.

De manière plus géométrique, on peut dire que les géodésiques sont les applications qui parcourent une droite de M à vitesse constante.

3.2.2 Distance et géodésiques sur la sphère

Proposition 3.25 : géodésiques de \mathbb{S}^{n-1}

Soit $n \geq 2$.

Les géodésiques de \mathbb{S}^{n-1} sont toutes les fonctions de la forme

$$\begin{aligned} \gamma : I &\rightarrow \mathbb{S}^{n-1} \\ t &\rightarrow \cos(ct)e_1 + \sin(ct)e_2, \end{aligned}$$

pour n'importe quel intervalle $I \neq \emptyset$, n'importe quel réel $c > 0$ et n'importe quels vecteurs $e_1, e_2 \in \mathbb{R}^n$ tels que

$$\|e_1\|_2 = \|e_2\|_2 = 1 \quad \text{and} \quad \langle e_1, e_2 \rangle = 0.$$

Remarque

Cela revient à dire que les géodésiques de la sphère sont les applications qui parcourent un « grand cercle »

$$\{\cos(s)e_1 + \sin(s)e_2, s \in \mathbb{R}\}$$

à vitesse constante, ou bien un arc de celui-ci.

Démonstration de la proposition 3.25. Soit d'abord γ une application de la forme indiquée. Vérifions qu'il s'agit d'une géodésique. Pour tout t ,

$$(T_{\gamma(t)}\mathbb{S}^{n-1})^\perp = (\{\gamma(t)\}^\perp)^\perp = \text{Vect}\{\gamma(t)\}.$$

Or, pour tout $t \in I$,

$$\begin{aligned} \gamma'(t) &= c(-\sin(ct)e_1 + \cos(ct)e_2); \\ \gamma''(t) &= -c^2(\cos(ct)e_1 + \sin(ct)e_2) = -c^2\gamma(t) \in \text{Vect}\{\gamma(t)\}. \end{aligned}$$

L'équation des géodésiques est donc bien vérifiée.

Réciproquement, soit γ une géodésique, définie sur un intervalle I . Montrons qu'elle est de la forme indiquée. Notons c sa vitesse (c'est-à-dire le réel positif tel que $\|\gamma'(t)\|_2 = c$ pour tout t ; on rappelle que γ est de vitesse constante d'après la proposition 3.23). Si $c = 0$, γ est constante et elle est

bien de la forme voulue (pour $e_1 = \gamma(t_0)$ et e_2 quelconque). Supposons donc $c > 0$.

Pour tout $t \in I$, $\gamma'(t) \in T_{\gamma(t)}\mathbb{S}^{n-1} = \{\gamma(t)\}^\perp$ donc

$$0 = \langle \gamma(t), \gamma'(t) \rangle.$$

Différencions cette égalité : pour tout t ,

$$\begin{aligned} 0 &= \langle \gamma(t), \gamma''(t) \rangle + \langle \gamma'(t), \gamma'(t) \rangle \\ &= \langle \gamma(t), \gamma''(t) \rangle + c^2. \end{aligned}$$

Donc $\langle \gamma(t), \gamma''(t) \rangle = -c^2$. Comme $\gamma''(t) \in (T_{\gamma(t)}\mathbb{S}^{n-1})^\perp = \text{Vect}\{\gamma(t)\}$ et $\gamma(t)$ est un vecteur unitaire, on doit avoir

$$\gamma''(t) = -c^2\gamma(t).$$

On sait qu'une solution de cette équation différentielle est nécessairement de la forme

$$\gamma : t \in I \rightarrow \cos(ct)e_1 + \sin(ct)e_2.$$

Fixons e_1, e_2 tels que γ a cette expression. Il reste à vérifier que $\|e_1\|_2 = \|e_2\|_2 = 1$ et $\langle e_1, e_2 \rangle = 0$.

Pour cela, fixons $t_0 \in I$ quelconque. Notons

$$v_1 = \gamma(t_0) \text{ et } v_2 = \frac{\gamma'(t_0)}{c}.$$

Ce sont deux vecteurs de norme 1 et orthogonaux entre eux. Exprimons e_1, e_2 en fonction de v_1, v_2 :

$$\begin{aligned} v_1 &= \gamma(t_0) = \cos(ct_0)e_1 + \sin(ct_0)e_2; \\ v_2 &= \frac{\gamma'(t_0)}{c} = -\sin(ct_0)e_1 + \cos(ct_0)e_2. \end{aligned}$$

On en déduit

$$e_1 = \cos(ct_0)v_1 - \sin(ct_0)v_2 \text{ et } e_2 = \sin(ct_0)v_1 + \cos(ct_0)v_2.$$

On a donc $\|e_1\|_2^2 = \cos^2(ct_0)\|v_1\|_2^2 - 2\cos(ct_0)\sin(ct_0)\langle v_1, v_2 \rangle + \sin^2(ct_0)\|v_2\|_2^2 = 1$ et, de même, $\|e_2\|_2^2 = 1$, $\langle e_1, e_2 \rangle = 0$. \square

Remarque

On voit à travers cet exemple que la longueur d'une géodésique n'est pas nécessairement égale à la distance entre ses points extrémaux : pour tous e_1, e_2 , la géodésique

$$\gamma : t \in [0; 2\pi] \rightarrow \cos(t)e_1 + \sin(t)e_2$$

est de longueur 2π . Pourtant, ses extrémités sont confondues : $\gamma(0) = e_1 = \gamma(2\pi)$; elles sont donc à distance nulle.

Toutefois, on peut montrer que, si $\gamma : I \rightarrow M$ est une géodésique de vitesse c , alors, pour tout $t \in I$, il existe $\epsilon > 0$ tel que, pour tout $t' \in I \cap]t - \epsilon; t + \epsilon[$,

$$\text{dist}_M(\gamma(t), \gamma(t')) = \int_{\min(t, t')}^{\max(t, t')} \|\gamma'(s)\|_2 ds = c|t - t'|.$$

On dit que les géodésiques sont *localement minimisantes*.

Remarque

On voit également à travers l'exemple de la sphère qu'il peut exister plusieurs chemins γ entre deux points x_1 et x_2 tels que

$$\ell(\gamma) = \text{dist}_M(x_1, x_2)$$

qui sont différents même à reparamétrisation près.

Par exemple, pour tous vecteurs e_1, e_2 de norme 1 et orthogonaux entre eux, les géodésiques

$$\gamma_1 : t \in [0; \pi] \rightarrow \cos(t)e_1 + \sin(t)e_2,$$

$$\gamma_2 : t \in [0; \pi] \rightarrow \cos(t)e_1 - \sin(t)e_2$$

sont des chemins de longueur minimale entre e_1 et $-e_1$, mais elles ne sont pas égales, même si on les reparamètre.

On peut en revanche montrer qu'on a une unicité locale du chemin de longueur minimale.

Corollaire 3.26 : distance sur \mathbb{S}^{n-1}

Soit $n \geq 2$. Soient $x_1, x_2 \in \mathbb{S}^{n-1}$. Alors

$$\text{dist}_{\mathbb{S}^{n-1}} = \arccos(\langle x_1, x_2 \rangle).$$

Démonstration. D'après les théorèmes 3.19 et 3.20, il existe au moins un chemin γ reliant x_1 à x_2 tel que

$$\ell(\gamma) = \text{dist}_{\mathbb{S}^{n-1}}(x_1, x_2)$$

et un tel chemin, si on le reparamètre à vitesse constante, est une géodésique. On a donc

$$\text{dist}_{\mathbb{S}^{n-1}}(x_1, x_2) = \min\{\ell(\gamma), \gamma \text{ géodésique reliant } x_1 \text{ et } x_2\}.$$

Calculons ce minimum.

Soit γ une géodésique quelconque reliant x_1 à x_2 . Déterminons les valeurs possibles pour sa longueur. Quitte à la composer par une translation, on peut supposer qu'elle est définie sur un intervalle de la forme $[0; A]$. Soient c, e_1, e_2 tels que, pour tout $t \in [0; A]$,

$$\gamma(t) = \cos(ct)e_1 + \sin(ct)e_2.$$

On doit avoir $x_1 = \gamma(0) = e_1$ et

$$x_2 = \gamma(A) = \cos(cA)e_1 + \sin(cA)e_2.$$

En particulier, $\langle x_1, x_2 \rangle = \langle e_1, x_2 \rangle = \cos(cA)$, de sorte que

$$\begin{aligned} cA &= \arccos(\langle x_1, x_2 \rangle) + 2k\pi \\ \text{ou } cA &= (2\pi - \arccos(\langle x_1, x_2 \rangle)) + 2k\pi, \end{aligned}$$

pour un certain $k \in \mathbb{Z}$ (en fait, $k \in \mathbb{N}$ car $cA \geq 0$). Comme $\ell(\gamma) = cA$, on en déduit que la longueur de γ est au moins

$$\min(\arccos(\langle x_1, x_2 \rangle), 2\pi - \arccos(\langle x_1, x_2 \rangle)) = \arccos(\langle x_1, x_2 \rangle).$$

Ainsi,

$$\text{dist}_{\mathbb{S}^{n-1}}(x_1, x_2) \geq \arccos(\langle x_1, x_2 \rangle).$$

Pour montrer que l'inégalité est une égalité, on remarque que, si on pose $e_2 = \frac{x_2 - \langle x_1, x_2 \rangle x_1}{\sqrt{1 - \langle x_1, x_2 \rangle^2}}$, la géodésique

$$\begin{aligned} \gamma : [0; \arccos(\langle x_1, x_2 \rangle)] &\rightarrow \mathbb{S}^{n-1} \\ t &\rightarrow \cos(t)x_1 + \sin(t)e_2 \end{aligned}$$

relie x_1 à x_2 et est de longueur $\arccos(\langle x_1, x_2 \rangle)$. □

3.2.3 Isométries

Définition 3.27 : isométries

Soient M_1, M_2 deux sous-variétés de (respectivement) \mathbb{R}^{n_1} et \mathbb{R}^{n_2} , de classe C^k avec $k \geq 1$. Une application $\phi : M_1 \rightarrow M_2$ de classe C^1 est une *isométrie* si

1. elle est bijective ;
2. elle préserve les distances : pour tous $x, x' \in M_1$,

$$\text{dist}_{M_1}(x, x') = \text{dist}_{M_2}(\phi(x), \phi(x')).$$

Théorème 3.28

On garde les mêmes notations que dans la définition précédente. Une application ϕ est une isométrie si et seulement si

1. elle est bijective ;
2. pour tout $x \in M_1$, $d\phi(x)$ réalise une isométrie^a entre $T_x M_1$ et $T_{\phi(x)} M_2$ (munis du produit scalaire euclidien usuel).

^a. Une application linéaire est une *isométrie* si elle est bijective et préserve la norme.

Démonstration. Supposons que ϕ est une isométrie. Elle est donc bijective. Montrons que, pour tout x , $d\phi(x)$ réalise une isométrie entre $T_x M_1$ et $T_{\phi(x)} M_2$.

On montre que $d\phi(x)$ préserve la norme grâce à la proposition suivante, démontrée dans l'annexe C.3.

Proposition 3.29

Soit M une sous-variété de \mathbb{R}^n . Soient $A > 0$ et $f : [0; A] \rightarrow M$ une fonction de classe C^1 . Alors

$$\text{dist}_M(f(0), f(t)) = t\|f'(0)\|_2 + o(t), \quad \text{quand } t \rightarrow 0^+.$$

Soit $v \in T_x M_1$ quelconque. Soient I un intervalle ouvert de \mathbb{R} contenant 0 et $f : I \rightarrow M_1$ une application C^1 telle que $f(0) = x$ et $f'(0) = v$. Pour t tendant vers 0 par valeurs positives,

$$\begin{aligned} t\|v\|_2 + o(t) &= t\|f'(0)\|_2 + o(t) \\ &= \text{dist}_{M_1}(f(0), f(t)) \\ &= \text{dist}_{M_2}(\phi \circ f(0), \phi \circ f(t)) \\ &\quad (\text{car } \phi \text{ est une isométrie}) \\ &= t\|(\phi \circ f)'(0)\|_2 + o(t) \\ &\quad (\text{d'après la proposition 3.29}) \\ &= t\|d\phi(f(0))(f'(0))\|_2 + o(t) \\ &= t\|d\phi(x)(v)\|_2 + o(t). \end{aligned}$$

Par identification des développements limités, $\|v\|_2 = \|d\phi(x)(v)\|_2$.

Maintenant que nous avons montré que $d\phi(x)$ préserve la norme, il reste à montrer qu'il s'agit d'une bijection de $T_x M_1$ vers $T_{\phi(x)} M_2$. Comme $d\phi(x)$ est injective (une application qui préserve la norme a nécessairement un noyau trivial), il faut seulement montrer qu'elle est surjective.

Soit $w \in T_{\phi(x)} M_2$ non nul. Montrons que $w \in \text{Im}(d\phi(x))$. Le principe de la démonstration va essentiellement être de choisir une suite de points de M_2 de la forme

$$\phi(x) + \frac{w}{n} + o\left(\frac{1}{n}\right),$$

de considérer leurs antécédants par ϕ et de montrer que ceux-ci (à extraction de sous-suite près) sont de la forme

$$x + \frac{v}{n} + o\left(\frac{1}{n}\right),$$

avec v tel que $d\phi(x)(v) = w$.

Plus formellement, soient I un intervalle ouvert contenant 0 et $g : I \rightarrow M_2$ telle que

$$g(0) = \phi(x), \quad g'(0) = w.$$

Remarquons que $\text{dist}_{M_2}(\phi(x), g(1/n)) \xrightarrow{n \rightarrow +\infty} 0$ (cela se déduit de la proposition 3.29), ce qui entraîne que $\text{dist}_{M_1}(x, \phi^{-1} \circ g(1/n)) \xrightarrow{n \rightarrow +\infty} 0$. Pour tout n assez grand, $x \neq \phi^{-1} \circ g(1/n)$ (car $g(1/n) = \phi(x) + \frac{w}{n} + o\left(\frac{1}{n}\right) \neq \phi(x)$ pour n assez grand). La suite

$$\left(\frac{\phi^{-1} \circ g(1/n) - x}{\|\phi^{-1} \circ g(1/n) - x\|_2} \right)_{n \in \mathbb{N}^*}$$

est donc bien définie à partir d'un certain rang. Elle est bornée. Soit $\alpha : \mathbb{N} \rightarrow \mathbb{N}$ une extraction définissant une sous-suite convergente. Notons v la limite. C'est un élément de $T_x M_1$.⁵

On a

$$\begin{aligned} \phi(x) + \frac{1}{\alpha(n)}w + o\left(\frac{1}{\alpha(n)}\right) &= g\left(\frac{1}{\alpha(n)}\right) \\ &= \phi(\phi^{-1} \circ g(1/\alpha(n))) \\ &= \phi(x + v\|\phi^{-1} \circ g(1/n) - x\|_2 + o(\|\phi^{-1} \circ g(1/n) - x\|_2)) \\ &= \phi(x) + \|\phi^{-1} \circ g(1/n) - x\|_2 d\phi(x)(v) + o(\|\phi^{-1} \circ g(1/n) - x\|_2). \end{aligned}$$

On en déduit que

$$\frac{1}{\alpha(n)}w \xrightarrow{n \rightarrow +\infty} \|\phi^{-1} \circ g(1/n) - x\|_2 d\phi(x)(v)$$

5. Pour tout élément x d'une sous-variété M et toute suite $(x_n)_{n \in \mathbb{N}}$ convergeant vers x telle que $\left(\frac{x_n - x}{\|x_n - x\|_2}\right)_{n \in \mathbb{N}}$ converge vers une limite v , cette limite appartient à $T_x M$. En effet, si V est un voisinage de x et $h : V \rightarrow \mathbb{R}^{n-d}$ est une submersion telle que $M \cap V = h^{-1}(0)$, alors

$$\begin{aligned} 0 &= h(x_n) \\ &= h(x + v\|x_n - x\|_2 + o(\|x_n - x\|_2)) \\ &= h(x) + \|x_n - x\|_2 dh(x)(v) + o(\|x_n - x\|_2) \\ &= \|x_n - x\|_2 dh(x)(v) + o(\|x_n - x\|_2), \end{aligned}$$

ce qui entraîne, par unicité du développement limité, que $dh(x)(v) = 0$, c'est-à-dire que $v \in \text{Ker}(dh(x)) = T_x M$.

et donc que $d\phi(x)(v)$ est colinéaire à w (et non-nul, puisque $d\phi(x)$ est injective). Donc $w \in \text{Im}(d\phi(x))$. Cela achève de montrer que $d\phi(x)$ est une isométrie et conclut donc la partie « seulement si » de la démonstration.

Montrons maintenant l'autre direction. Supposons que ϕ est bijective et $d\phi(x)$ est une isométrie de $T_x M_1$ vers $T_{\phi(x)} M_2$ pour tout x . Montrons qu'il s'agit d'une isométrie, c'est-à-dire montrons qu'elle préserve la distance.

Remarquons d'abord que ϕ est un difféomorphisme. En effet, c'est une bijection continue ; il suffit de montrer que ϕ^{-1} est C^1 . Comme ϕ est un difféomorphisme sur un voisinage de chaque point (car $d\phi(x)$ est bijective pour tout x), sa réciproque est C^1 au voisinage de chaque point. Elle est donc globalement C^1 .

Soient $x, x' \in M_1$. Soit $(\gamma_n : [0; A_n] \rightarrow M_1)_{n \in \mathbb{N}}$ une suite de chemins de M_1 reliant x à x' telle que

$$\ell(\gamma_n) \xrightarrow{n \rightarrow +\infty} \text{dist}_{M_1}(x, x').$$

Pour tout n , $\phi \circ \gamma_n$ est un chemin de M_2 reliant $\phi(x)$ à $\phi(x')$ et

$$\begin{aligned} \ell(\gamma_n) &= \int_0^{A_n} \|\gamma_n'(t)\|_2 dt \\ &= \int_0^{A_n} \|d\phi(\gamma_n(t))(\gamma_n'(t))\|_2 dt \\ &\quad \text{(car } d\phi(\gamma_n(t)) \text{ est une isométrie pour tout } t) \\ &= \int_0^{A_n} \|(\phi \circ \gamma_n)'(t)\|_2 dt \\ &= \ell(\phi \circ \gamma_n). \end{aligned}$$

On en déduit que

$$\text{dist}_{M_2}(\phi(x), \phi(x')) \leq \lim_{n \rightarrow +\infty} \ell(\gamma_n) = \text{dist}_{M_1}(x, x').$$

On peut appliquer le même raisonnement en échangeant les rôles de M_1 et M_2 et en remplaçant ϕ par ϕ^{-1} (c'est là qu'il est utile d'avoir montré que ϕ est un difféomorphisme). On a donc aussi

$$\text{dist}_{M_2}(\phi(x), \phi(x')) \geq \text{dist}_{M_1}(x, x').$$

Et donc $\text{dist}_{M_2}(\phi(x), \phi(x')) = \text{dist}_{M_1}(x, x')$.

□

Proposition 3.30

On garde toujours les mêmes notations mais on suppose cette fois que $k \geq 2$ et que ϕ est de classe C^2 . Une application ϕ est une isométrie si et seulement si

1. c'est un difféomorphisme de M_1 vers M_2 ;
2. pour tout géodésique $\gamma : I \rightarrow M_1$, $\phi \circ \gamma$ est une géodésique de M_2 parcourue à la même vitesse.

Démonstration. Supposons d'abord que ϕ est un difféomorphisme qui envoie une géodésique sur une géodésique de même vitesse. Montrons que c'est une isométrie. Le fait que ϕ soit un difféomorphisme entraîne que ϕ est une bijection et que, pour tout $x \in M_1$, $d\phi(x) : T_x M_1 \rightarrow T_{\phi(x)} M_2$ est aussi une bijection. D'après le théorème 3.28, il suffit de montrer que, pour tout $x \in M_1$, $d\phi(x)$ préserve la norme.

Pour tout $v \in T_x M_1$, il existe une géodésique $\gamma : I \rightarrow M_1$, avec I un intervalle ouvert contenant 0, telle que

$$\gamma(0) = x \text{ et } \gamma'(0) = v.$$

(Nous admettons cette propriété ici ; elle se démontre à l'aide du théorème de Cauchy-Lipschitz qui sera présenté dans le chapitre suivant.)

Puisque γ est une géodésique de vitesse $\|v\|_2$ de M_1 , l'application $\phi \circ \gamma$ est une géodésique de M_2 de vitesse $\|v\|_2$. En particulier,

$$\|d\phi(x)(v)\|_2 = \|d\phi(\gamma(0))(\gamma'(0))\|_2 = \|(\phi \circ \gamma)'(0)\|_2 = \|v\|_2.$$

Donc $d\phi(x)$ préserve la norme pour tout x . Cela conclut la première partie de la démonstration.

Montrons maintenant la deuxième partie. Supposons que ϕ est une isométrie. D'après le théorème 3.28, c'est donc une bijection, dont la différentielle en chaque point est une isométrie. Ainsi, par le même argument que dans la démonstration du théorème 3.28, ϕ est un difféomorphisme.

Montrons que ϕ envoie une géodésique sur une géodésique de même vitesse. Soit $\gamma : I \rightarrow M_1$ une géodésique. Notons c sa vitesse. Nous allons montrer que $\phi \circ \gamma$ est une géodésique de M_2 .

Soit $t \in I$ quelconque. D'après la remarque suivant la démonstration de la proposition 3.25, il existe $\epsilon > 0$ tel que, pour tout $t' \in I \cap]t - \epsilon; t + \epsilon[$,

$$\text{dist}_{M_1}(\gamma(t), \gamma(t')) = c|t - t'|.$$

Puisque ϕ est une isométrie, on a pour tout t' ,

$$\text{dist}_{M_1}(\gamma(t), \gamma(t')) = \text{dist}_{M_2}(\phi \circ \gamma(t), \phi \circ \gamma(t')) = c|t - t'|.$$

De plus, $\|d\phi(\gamma(s))(\gamma'(s))\|_2 = \|\gamma'(s)\|_2 = c$ pour tout $s \in I$ (car $d\phi(\gamma(s))$ est une isométrie de $T_{\gamma(s)}M_1$ vers $T_{\phi \circ \gamma(s)}M_2$, d'après le théorème précédent). Donc, pour tout $t' \in I \cap]t - \epsilon; t + \epsilon[$,

$$\text{dist}_{M_2}(\phi \circ \gamma(t), \phi \circ \gamma(t')) = c|t - t'| = \int_t^{t'} \|(\phi \circ \gamma)'(s)\|_2 ds.$$

L'application $\phi \circ \gamma$ restreinte à $[t; t']$ est donc un chemin reliant $\phi \circ \gamma(t)$ à $\phi \circ \gamma(t')$, parcouru à vitesse constante, tel que $\ell(\phi \circ \gamma|_{[t; t']}) = \text{dist}_{M_2}(\phi \circ \gamma(t), \phi \circ \gamma(t'))$. D'après le théorème 3.20, c'est une géodésique.

Ainsi, $\phi \circ \gamma$ vérifie l'équation (3.10) sur $I \cap]t - \epsilon; t + \epsilon[$. En particulier,

$$(\phi \circ \gamma)''(t) \in (T_{\phi \circ \gamma(t)}M_2)^\perp.$$

Comme on avait choisi t quelconque, c'est vrai pour tout $t \in I$, c'est-à-dire que $\phi \circ \gamma$ est une géodésique. Elle est de même vitesse que γ car $d\phi$ est une isométrie en tout point. □

3.2.4 Exemples d'isométries

Nous donnons ici deux exemples simples et un contre-exemple de sous-variétés isométriques.

Le premier exemple est celui des courbes. Grâce aux propriétés que nous avons montrées dans la section 3.1 de ce chapitre, on peut exactement caractériser quelles courbes connexes sont isométriques.

Proposition 3.31

Soient C_1, C_2 deux courbes connexes de longueur finie. Elles sont isométriques si et seulement si elles vérifient les deux propriétés suivantes :

1. elles sont toutes deux compactes ou toutes deux non-compactes ;
2. elles ont la même longueur.

Démonstration. Si C_1 et C_2 sont isométriques, elles sont en particulier homéomorphes et sont donc toutes deux compactes ou toutes deux non-compactes. Vérifions qu'elles sont la même longueur. Soit (I, γ) un paramétrage global de C_1 . En notant $\phi : C_1 \rightarrow C_2$ une isométrie, on a que $(I, \phi \circ \gamma)$ est un paramétrage global de C_2 . Puisque $d\phi(x)$ est une isométrie linéaire pour tout x ,

$$\begin{aligned} \ell(C_1) &= \int_I \|\gamma'(t)\|_2 dt \\ &= \int_I \|(\phi \circ \gamma)'(t)\|_2 dt \\ &= \ell(C_2). \end{aligned}$$

Montrons maintenant que si C_1 et C_2 ont la même longueur et sont toutes deux compactes ou toutes deux non-compactes, alors elles sont isométriques. Supposons par exemple qu'elles sont toutes deux non-compactes ; la démonstration est essentiellement identique si elles sont compactes.

Soient (I_1, γ_1) et (I_2, γ_2) des paramétrages par abscisse curviligne de C_1 et C_2 . Comme C_1 et C_2 sont de mêmes longueurs, I_1 et I_2 sont des intervalles (ouverts) de même longueur. Quitte à composer γ_1 et γ_2 par des translations, on peut supposer que

$$I_1 = I_2 =]0; L[,$$

avec L la longueur commune de C_1 et C_2 . Alors γ_1 et γ_2 sont des isométries de $]0; L[$ vers, respectivement, C_1 et C_2 . Si on pose $\phi = \gamma_2 \circ \gamma_1^{-1}$, c'est une isométrie de C_1 vers C_2 . □

Remarque

La proposition précédente n'est pas toujours vraie pour des courbes de longueur infinie : \mathbb{R} et $]0; +\infty[$ ne sont pas isométriques.

Le deuxième exemple est celui du plan \mathbb{R}^2 et du cylindre

$$\text{Cyl} \stackrel{\text{d\'ef}}{=} \mathbb{S}^1 \times \mathbb{R} \subset \mathbb{R}^3.$$

Ces deux sous-variétés ne sont pas globalement isométriques mais elles le sont localement.

Exemple 3.32 : isométrie du plan vers le cylindre

Soit $c = (c_1, c_2, c_3) \in \text{Cyl}$. Soit $\theta \in \mathbb{R}$ tel que $(c_1, c_2) = (\cos(\theta), \sin(\theta))$. Définissons

$$\begin{aligned} \phi :]-\pi; \pi[\times \mathbb{R} &\rightarrow \text{Cyl} \\ (x, y) &\rightarrow (\cos(\theta + x), \sin(\theta + x), c_3 + y). \end{aligned}$$

Cette application est une isométrie entre $]-\pi; \pi[\times \mathbb{R}$, qui est un voisinage de $(0, 0)$ dans \mathbb{R}^2 et $\text{Cyl} \setminus \{(-\cos(\theta), -\sin(\theta), h), h \in \mathbb{R}\}$, qui est un voisinage de c dans Cyl .

En effet, elle est C^1 (et même C^∞). Il s'agit d'une bijection. Pour tout (x, y) , $d\phi(x, y)$ est l'application

$$(h_1, h_2) \in \mathbb{R}^2 \rightarrow (-\sin(\theta + x)h_1, \cos(\theta + x)h_1, h_2) \in T_{\phi(x, y)}\text{Cyl}.$$

Elle préserve la norme. En particulier, elle est injective. Comme ses espaces de départ et d'arrivée ont la même dimension, elle est en fait bijective : c'est une isométrie.

Terminons avec le contre-exemple.

Proposition 3.33 : la sphère et le plan

Un ouvert de la sphère \mathbb{S}^2 et un ouvert du plan \mathbb{R}^2 ne sont jamais isométriques.

Démonstration. Raisonnons par l'absurde et supposons qu'il existe une isométrie $\phi : \Omega_1 \rightarrow \Omega_2$, avec Ω_1 un ouvert de \mathbb{R}^2 et Ω_2 un ouvert de \mathbb{S}^2 . Quitte à composer ϕ par une translation, on peut supposer que Ω_1 contient $(0, 0)$.

Pour tout $\epsilon > 0$ assez petit, définissons

$$x_0 = (0, 0), \quad x_{1, \epsilon} = (\epsilon, 0), \quad x_{2, \epsilon} = (-\epsilon, 0), \quad x_{3, \epsilon} = (0, \epsilon).$$

Nous allons considérer les images de ces points par ϕ et montrer qu'il est impossible que les distances entre les points images soient identiques aux distances entre les $x_{k, \epsilon}$.

Notons $e_0 = \phi(x_0)$. Pour tout $\epsilon > 0$ assez petit,⁶

$$\text{dist}_{\Omega_1}(x_0, x_{k,\epsilon}) = \epsilon, \quad \text{pour } k = 1, 2, 3.$$

Comme, pour tout k , $\text{dist}_{\Omega_2}(e_0, \phi(x_{k,\epsilon})) = \text{dist}_{\Omega_1}(x_0, x_{k,\epsilon})$, le corollaire 3.26 entraîne que

$$\langle e_0, \phi(x_{k,\epsilon}) \rangle = \cos(\epsilon).$$

Cela permet de montrer qu'il existe des vecteurs $e_{1,\epsilon}, e_{2,\epsilon}, e_{3,\epsilon}$ de norme 1, orthogonaux à e_0 , tels que, pour tout $k = 1, 2, 3$,

$$x_{k,\epsilon} = \cos(\epsilon)e_0 + \sin(\epsilon)e_{k,\epsilon}.$$

Puisque $\text{dist}_{\Omega_1}(x_{1,\epsilon}, x_{2,\epsilon}) = 2\epsilon$, on doit avoir

$$\begin{aligned} 2\epsilon &= \text{dist}_{\Omega_2}(\phi(x_{1,\epsilon}), \phi(x_{2,\epsilon})) \\ &= \arccos(\langle \cos(\epsilon)e_0 + \sin(\epsilon)e_{1,\epsilon}, \cos(\epsilon)e_0 + \sin(\epsilon)e_{2,\epsilon} \rangle) \\ &= \arccos(\cos^2(\epsilon) + \sin^2(\epsilon) \langle e_{1,\epsilon}, e_{2,\epsilon} \rangle). \end{aligned}$$

Donc

$$\begin{aligned} \cos^2(\epsilon) + \sin^2(\epsilon) \langle e_{1,\epsilon}, e_{2,\epsilon} \rangle &= \cos(2\epsilon) \\ &= \cos^2(\epsilon) - \sin^2(\epsilon), \end{aligned}$$

ce qui donne $\langle e_{1,\epsilon}, e_{2,\epsilon} \rangle = -1$, donc (puisque les deux vecteurs sont de norme 1) $e_{2,\epsilon} = -e_{1,\epsilon}$.

Finalement, comme $\text{dist}_{\Omega_1}(x_{1,\epsilon}, x_{3,\epsilon}) = \text{dist}_{\Omega_1}(x_{2,\epsilon}, x_{3,\epsilon}) = \sqrt{2}\epsilon$, on doit avoir

$$\begin{aligned} \sqrt{2}\epsilon &= \text{dist}_{\Omega_2}(\phi(x_{1,\epsilon}), \phi(x_{3,\epsilon})) \\ &= \arccos(\cos^2(\epsilon) + \sin^2(\epsilon) \langle e_{1,\epsilon}, e_{3,\epsilon} \rangle), \end{aligned}$$

donc $\cos^2(\epsilon) + \sin^2(\epsilon) \langle e_{1,\epsilon}, e_{3,\epsilon} \rangle = \cos(\sqrt{2}\epsilon)$. De même, $\cos^2(\epsilon) + \sin^2(\epsilon) \langle e_{2,\epsilon}, e_{3,\epsilon} \rangle = \cos(\sqrt{2}\epsilon)$.

En sommant ces inégalités et en utilisant le fait que $e_{1,\epsilon} = -e_{2,\epsilon}$, on obtient

$$\cos^2(\epsilon) = \cos(\sqrt{2}\epsilon).$$

6. On n'a pas a priori $\text{dist}_{\Omega_1} = \text{dist}_{\mathbb{R}^2}$ car les chemins reliant deux points dans Ω_1 ne sont pas les mêmes que les chemins reliant deux points dans \mathbb{R}^2 . Toutefois, on peut montrer que ces deux distances coïncident au voisinage de chaque point de Ω_1 .

Mais cette égalité est fausse pour tout $\epsilon > 0$ assez petit (une façon de le voir est que le DL à l'ordre 4 du membre de gauche est $1 - \epsilon^2 + \frac{\epsilon^4}{3} + o(\epsilon^4)$, tandis que celui du membre de droite est $1 - \epsilon^2 + \frac{\epsilon^4}{6} + o(\epsilon^4)$).

□

Chapitre 4

Équations différentielles : existence et unicité

4.1 Théorème de Cauchy-Lipschitz

On appelle *problème de Cauchy* une équation différentielle où l'inconnue est une fonction d'une variable (la variable est souvent notée t), assortie d'une condition initiale. C'est donc un problème de la forme suivante :

$$\begin{array}{l} u' = f(t, u), \\ u(t_0) = u_0. \end{array} \quad \text{(Cauchy)}$$

Ici,

- $f : I \times U \rightarrow \mathbb{R}^n$ est une fonction fixée, avec I un intervalle ouvert de \mathbb{R} et U un ouvert de \mathbb{R}^n (pour un certain $n \in \mathbb{N}^*$);
- t_0 est un élément de I et u_0 un élément de U ;
- u est la fonction inconnue, qui doit être définie sur un intervalle J tel que $t_0 \in J \subset I$, à valeurs dans U et dérivable.

(Précisons que l'égalité « $u' = f(t, u)$ » est une écriture raccourcie pour « $u'(t) = f(t, u(t))$ » : u est bien une *fonction* qui dépend d'une variable notée ici t .)

Remarque

Dans le problème (**Cauchy**), nous imposons à l'équation différentielle d'être d'ordre 1 (c'est-à-dire de ne contenir qu'une seule dérivation). Ce n'est pas une restriction. En effet, un problème de Cauchy contenant une équation différentielle d'ordre $N \geq 1$ quelconque peut être reformulé comme un problème de Cauchy d'ordre 1. Précisément, considérons un problème de la forme

$$\begin{aligned} u^{(N)} &= g(t, u, u', \dots, u^{(N-1)}) \\ u(t_0) &= u_{0,0}, \quad u'(t_0) = u_{0,1}, \quad \dots, \quad u^{(N-1)}(t_0) = u_{0,N-1}. \end{aligned}$$

Si on note $v_0 = u, v_1 = u', \dots, v_{N-1} = u^{(N-1)}$, il est équivalent à

$$\begin{aligned} v_0' &= v_1 \\ &\dots \\ v_{N-2}' &= v_{N-1} \\ v_{N-1}' &= g(t, v_0, v_1, \dots, v_{N-1}) \\ v_0(t_0) &= u_{0,0}, \quad v_1(t_0) = u_{0,1}, \quad \dots, \quad v_{N-1}(t_0) = u_{0,N-1}, \end{aligned}$$

ce qui est un problème d'ordre 1 sur la fonction inconnue $\begin{pmatrix} v_0 \\ \vdots \\ v_{N-1} \end{pmatrix}$.

Le point de départ de la théorie des équations différentielles est le théorème de Cauchy-Lipschitz, qui, sous des hypothèses de régularité sur f , garantit essentiellement que le problème (**Cauchy**) admet une unique solution au voisinage de t_0 .

Théorème 4.1 : Cauchy-Lipschitz

Supposons que f est continue et qu'il existe un voisinage $H \subset I \times U$ de (t_0, u_0) sur lequel elle est lipschitzienne en sa deuxième variable :

$$\forall t, u, v \text{ tq } (t, u), (t, v) \in H, \quad \|f(t, u) - f(t, v)\|_2 \leq C \|u - v\|_2, \quad (4.1)$$

pour une certaine constante $C > 0$ (qui ne doit pas dépendre de t).

On a alors les conclusions suivantes :

- (Existence)

Il existe un intervalle $J \subset I$ dont l'intérieur contient t_0 et une fonction $u : J \rightarrow U$ de classe C^1 qui est solution du problème (Cauchy).

— (Unicité locale)

Si u_1, u_2 sont deux fonctions C^1 solutions du problème (Cauchy), définies sur des intervalles J_1, J_2 contenant t_0 (à l'intérieur ou non), alors

$$u_1 = u_2 \quad \text{sur } J_1 \cap J_2 \cap [t_0 - \epsilon; t_0 + \epsilon]$$

pour tout $\epsilon > 0$ assez petit.

La démonstration la plus classique de ce théorème utilise (implicitement ou explicitement) le *théorème du point fixe de Picard*. Le lecteur ou la lectrice intéressée la trouvera par exemple dans [Benzoni-Gavage, 2010, p. 142].

La condition de lipschitziannité au voisinage de (t_0, u_0) (équation (4.1)) est automatiquement vérifiée dès que f est de classe C^1 . En effet, dans ce cas, on peut prendre $H = \overline{B}((t_0, u_0), \epsilon)$, pour tout $\epsilon > 0$ assez petit. L'équation (4.1) est alors une conséquence de l'inégalité des accroissements finis (Théorème 1.16), avec

$$C = \max_{(t,u) \in \overline{B}((t_0,u_0),\epsilon)} \|df(t,u)\|_{\mathcal{L}(\mathbb{R}^{n+1}, \mathbb{R}^n)}.$$

La partie « existence » du théorème est vraie même sans la condition de lipschitziannité (il suffit que f soit continue ; c'est le *théorème de Peano*). En revanche, la partie « unicité » peut être fautive sans cette condition. Pour donner un exemple de possible non-unicité, considérons le problème de Cauchy

$$\begin{aligned} u' &= \sqrt{u}, \\ u(0) &= 0. \end{aligned}$$

On peut vérifier que les applications

$$u_1 : \mathbb{R} \rightarrow \mathbb{R} \\ t \rightarrow \begin{cases} \frac{t^2}{4} & \text{si } t \geq 0, \\ 0 & \text{si } t < 0, \end{cases}$$

$$\begin{aligned} u_2 : \mathbb{R} &\rightarrow \mathbb{R} \\ t &\rightarrow 0, \end{aligned}$$

sont toutes deux solutions de ce problème. Pourtant, elles ne sont pas identiques.

Terminons cette section par une propriété simple mais bien utile sur la régularité des solutions d'un problème de Cauchy.

Proposition 4.2

Si f est de classe C^r pour un certain $r \in \mathbb{N}$, toute solution u du problème (Cauchy) est de classe C^{r+1} .

En particulier, si f est C^∞ , toute solution est C^∞ .

Démonstration. On démontre le résultat par récurrence sur r . Pour $r = 0$, il est vrai : si u est une solution, elle est dérivable par définition. En particulier, elle est continue. Sa dérivée vaut

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

Comme f et u sont continues, u' l'est aussi, c'est-à-dire que u est C^1 .

Supposons le résultat démontré pour un certain r et démontrons-le pour $r + 1$. Supposons f de classe C^{r+1} et soit u une solution. Comme f est aussi de classe C^r , l'hypothèse de récurrence nous dit que u est C^{r+1} . Donc

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

est une composée de fonctions de classe C^{r+1} . C'est donc une fonction de classe C^{r+1} , c'est-à-dire que u est C^{r+2} . □

Remarque : extension aux espaces de Banach

Nous nous limitons ici aux équations différentielles en dimension finie, c'est-à-dire que la fonction u du problème (Cauchy) est à valeurs dans \mathbb{R}^n . On peut considérer, plus généralement, des équations dont l'inconnue est à valeurs dans un espace de Banach ^a et tout ce qui a été dit dans cette section reste vrai, à l'exception du théorème de Peano.

^a. c'est-à-dire un espace vectoriel normé complet,

4.2 Solutions maximales

Définition 4.3 : solutions maximales

Soit $u : J \rightarrow U$ une solution d'un problème de la forme (Cauchy). On dit que c'est une *solution maximale* du problème si on ne peut pas la prolonger à un intervalle plus grand : pour toute autre solution $\tilde{u} : \tilde{J} \rightarrow U$ telle que $J \subset \tilde{J}$ et $\tilde{u}|_J = u$, on a

$$\tilde{J} = J \quad \text{et} \quad \tilde{u} = u.$$

Proposition 4.4 : existence d'une unique solution maximale

Si la fonction f du problème (Cauchy) est continue et lipschitzienne en sa deuxième variable au voisinage de tout point, alors le problème admet une unique solution maximale.

De plus, si on note $u : J \rightarrow U$ cette solution maximale, l'ensemble des solutions du problème (Cauchy) est

$$\left\{ u|_{\tilde{J}} : \tilde{J} \rightarrow U \text{ avec } \tilde{J} \text{ intervalle tel que } t_0 \in \tilde{J} \subset J \right\}. \quad (4.2)$$

Démonstration. On commence par une proposition (démontrée juste après cette démonstration-ci) qui établit un résultat d'unicité des solutions du problème (Cauchy). Ce résultat est très similaire à celui qui est contenu dans le théorème de Cauchy-Lipschitz mais il est global, tandis que le théorème de Cauchy-Lipschitz donne une garantie locale (l'unicité est vraie seulement sur un voisinage de t_0). Cette globalité provient du fait qu'ici, f est lipschitzienne en sa deuxième variable *au voisinage de tout point* et pas seulement au voisinage de (t_0, u_0) .

Proposition 4.5

Si $u_1 : J_1 \rightarrow U$ et $u_2 : J_2 \rightarrow U$ sont deux solutions du problème (Cauchy), alors

$$u_1 = u_2 \quad \text{sur } J_1 \cap J_2.$$

De plus, la fonction $u : J_1 \cup J_2 \rightarrow U$ qui coïncide avec u_1 sur J_1 et u_2 sur J_2 est solution du problème (Cauchy).

De cette proposition, on peut déjà déduire que la solution maximale, si elle existe, est unique et que l'ensemble des solutions du problème (Cauchy) est bien celui donné dans l'équation (4.2).

En effet, supposons qu'il existe une solution maximale u , définie sur un intervalle J . Pour tout intervalle \tilde{J} tel que $t_0 \in \tilde{J} \subset J$, $u|_{\tilde{J}}$ est une solution du problème (Cauchy). Réciproquement, si $v : \tilde{J} \rightarrow U$ est une solution du problème, la proposition précédente nous dit qu'il existe une solution définie sur $J \cup \tilde{J}$ qui coïncide avec u sur J . Comme u est maximale, on doit avoir $J \cup \tilde{J} = J$, c'est-à-dire $\tilde{J} \subset J$. De plus, la proposition nous dit que $v = u$ sur $\tilde{J} \cap J = \tilde{J}$. On a donc

$$v = u|_{\tilde{J}}.$$

Cela démontre l'équation (4.2).

Pour ce qui est de l'unicité de la solution maximale, imaginons qu'il en existe une autre, qu'on note $v : \tilde{J} \rightarrow U$. D'après l'équation (4.2) que nous venons de démontrer,

$$\tilde{J} \subset J \quad \text{et} \quad v = u|_{\tilde{J}}.$$

Cela entraîne que u est un prolongement de v à l'intervalle J . Puisque v est maximale, on doit avoir $\tilde{J} = J$ et

$$v = u.$$

La solution maximale est donc bien unique.

Pour finir, montrons l'existence. Soit

$$J = \{t \in \mathbb{R}, (\text{Cauchy}) \text{ a une solution définie sur } [t_0; t]\}.$$

Pour tout $t \in J$, on note v_t une solution du problème (Cauchy) définie sur $[t_0; t]$ ¹ et on pose

$$u(t) = v_t(t).$$

Cela définit une application $u : J \rightarrow U$.

Montrons d'abord que u est solution du problème (Cauchy). Son ensemble de définition J est un intervalle : pour tous $t, t' \in J$ et tout $t'' \in [t; t']$, on a que $[t_0; t]$ ou $[t_0; t']$ (ou les deux) contient $[t_0; t'']$. La restriction à $[t_0; t'']$ de v_t ou $v_{t'}$ est donc bien définie et c'est une solution de (Cauchy). On a donc $t'' \in J$.

1. On note l'intervalle « $[t_0; t]$ » par simplicité mais, bien sûr, si $t_0 < t$, on considère en fait l'intervalle « $[t; t_0]$ ».

L'application u vérifie la condition initiale : $u(t_0) = v_{t_0}(t_0)$ et, comme v_{t_0} est solution du problème, on a $v_{t_0}(t_0) = u_0$, d'où

$$u(t_0) = u_0.$$

On montre ensuite que, pour tout $t \in J$, u est dérivable en t et vérifie l'égalité

$$u'(t) = f(t, u(t)). \quad (4.3)$$

Fixons $t \in J$ quelconque. Supposons pour simplifier les notations que $t > t_0$ (on peut faire exactement le même raisonnement si $t < t_0$ et un raisonnement très proche si $t = t_0$) et distinguons deux cas.

- Premier cas : $t < \sup J$. Dans ce cas, soit $t' \in]t; \sup J[$. La fonction u coïncide avec $v_{t'}$ sur $[t_0; t']$. En effet, pour tout $t'' \in [t_0; t']$, d'après la proposition 4.5,

$$v_{t'} = v_{t''} \quad \text{sur } [t; t'] \cap [t; t''] = [t; t''].$$

Donc $u(t'') = v_{t''}(t'') = v_{t'}(t'')$.

Comme $v_{t'}$ est dérivable et solution du problème de Cauchy, l'égalité $u = v_{t'}$ sur $[t_0; t']$ entraîne que u est également dérivable sur $]t_0; t'[$, en particulier dérivable en t , et vérifie l'égalité (4.3).

- Deuxième cas : $t = \sup J$. Dans ce cas, J est de la forme $[\alpha; t]$ ou $] \alpha; t]$, pour un certain $\alpha \in [-\infty; t_0]$.

Suivant le même raisonnement que dans le premier cas, on voit que u coïncide avec v_t sur $[t_0; t]$. Cela implique que u est dérivable sur $]t_0; t]$, qui est un voisinage de t dans J , et que l'égalité (4.3) est vérifiée.

Cela achève de montrer que u est une solution du problème (Cauchy).

Montrons pour conclure que cette solution est maximale. Soit $\tilde{u} : \tilde{J} \rightarrow U$ une solution prolongeant u (c'est-à-dire que $J \subset \tilde{J}$ et $\tilde{u}|_J = u$). Pour tout $t \in \tilde{J}$, $\tilde{u}|_{[t_0; t]}$ est une solution du problème (Cauchy) donc t appartient à J . Ainsi, $\tilde{J} \subset J$. Donc $\tilde{J} = J$ et $\tilde{u} = u$. \square

Démonstration de la proposition 4.5. Soient $u_1 : J_1 \rightarrow U$ et $u_2 : J_2 \rightarrow U$ deux solutions du problème (Cauchy). Notons

$$H = \{t \in J_1 \cap J_2 \text{ tq } u_1(t) = u_2(t)\}.$$

L'ensemble H est non-vide (il contient t_0) et fermé dans $J_1 \cap J_2$ (car u_1 et u_2 sont continues). Si on parvient à montrer qu'il est ouvert dans $J_1 \cap J_2$, cela

entraîne que $H = J_1 \cap J_2$ (car $J_1 \cap J_2$ est une intersection d'intervalles, donc un intervalle, donc un ensemble connexe) et donc que

$$u_1 = u_2 \quad \text{sur } H = J_1 \cap J_2.$$

Montrons donc qu'il est ouvert. Soit $t_1 \in H$ quelconque. Considérons le problème de Cauchy modifié suivant.

$$\begin{aligned} u' &= f(t, u), & (\text{Cauchy } t_1) \\ u(t_1) &= u_1(t_1). \end{aligned}$$

Les fonctions u_1 et u_2 sont toutes deux solutions de ce problème car elles sont solutions de (Cauchy) et $u_1(t_1) = u_2(t_1)$ d'après la définition de H .

On peut appliquer le théorème de Cauchy-Lipschitz à (Cauchy t_1) : f est continue et lipschitzienne par rapport à sa deuxième variable sur un voisinage de $(t_1, u_1(t_1))$. D'après le résultat d'unicité locale de ce théorème, il existe $\epsilon > 0$ tel que

$$u_1 = u_2 \quad \text{sur } J_1 \cap J_2 \cap [t_1 - \epsilon; t_1 + \epsilon].$$

Cela entraîne que $J_1 \cap J_2 \cap [t_1 - \epsilon; t_1 + \epsilon] \subset H$ et donc que H contient un voisinage de t_1 dans $J_1 \cap J_2$. Cela termine de montrer que H est ouvert dans $J_1 \cap J_2$.

Pour conclure, notons $u : J_1 \cup J_2 \rightarrow U$ la fonction qui coïncide avec u_1 sur J_1 et u_2 sur J_2 . Vérifions qu'elle est solution du problème (Cauchy).

Elle vérifie la condition $u(t_0) = u_0$ (car u_1 et u_2) la satisfont. Montrons qu'elle est dérivable et vérifie l'égalité

$$u' = f(t, u). \tag{4.4}$$

On vérifie (en utilisant les propriétés basiques des intervalles) que $(J_1 \cup J_2) \cap [t_0; +\infty[$ est inclus dans J_1 ou J_2 . La fonction u est donc dérivable sur cet intervalle (elle coïncide avec u_1 ou u_2 , qui l'est) et vérifie l'égalité (4.4) (car u_1 et u_2 la vérifient). De même sur $(J_1 \cup J_2) \cap]-\infty; t_0]$. Cela entraîne que u est dérivable et vérifie (4.4) sur $(J_1 \cup J_2) \setminus \{t_0\}$. De plus, elle admet des dérivées à gauche et à droite en t_0 , qui satisfont aussi (4.4). À cause de cette égalité, les dérivées à gauche et à droite coïncident (elles valent $f(t_0, u_0)$) donc u est dérivable en t_0 et vérifie (4.4) en ce point également. □

4.3 Théorème des bouts

Dans cette section, on considère un problème de Cauchy et on suppose que f est continue et lipschitzienne par rapport à sa deuxième variable au voisinage de tout point. Cela permet d'appliquer les résultats de la section précédente : il existe une unique solution maximale $u : J \rightarrow U$.

Proposition 4.6

L'ensemble de définition maximal J est un intervalle ouvert de \mathbb{R} .

Démonstration. On sait que J est un intervalle. Il faut montrer qu'il est ouvert.

Soit $T \in J$ quelconque. D'après le théorème de Cauchy-Lipschitz, le problème de Cauchy

$$\begin{aligned} v' &= f(t, v), \\ v(T) &= u(T) \end{aligned}$$

admet une solution v définie sur un intervalle dont l'intérieur contient T . Notons H cet intervalle.

D'après la proposition 4.5, comme v et u sont toutes deux solutions de ce problème de Cauchy, elles coïncident sur l'intersection de leurs ensembles de définition et la fonction $w : J \cup H \rightarrow U$ qui coïncide avec u sur J et v sur H est aussi une solution. Cette fonction w est également solution du problème (Cauchy) de départ (car $w(t_0) = u(t_0) = u_0$).

Puisque u est une solution maximale, on doit avoir $J \cup H \subset J$, c'est-à-dire $H \subset J$. Donc J contient un voisinage de T .

C'est vrai pour tout $T \in J$ donc J est ouvert. \square

Une question importante à propos de la solution maximale est de déterminer son ensemble de définition. En particulier, cette solution maximale est-elle globale, c'est-à-dire définie sur le même intervalle I que la fonction f ? Le *théorème des bouts* donne un critère qui, dans certains cas, permet de répondre à cette question.

Théorème 4.7 : théorème des bouts

On suppose toujours que $f : I \times U \rightarrow \mathbb{R}^n$ est continue et lipschitzienne par rapport à sa deuxième variable au voisinage de tout point. On note

toujours $u : J \rightarrow U$ la solution maximale du problème (Cauchy).
L'une des deux propriétés suivantes est nécessairement vraie.

1. $\sup J = \sup I$;
2. u « sort de tout compact de U » au voisinage de $\sup J$: pour tout compact $K \subset U$, il existe $\eta < \sup J$ tel que, pour tout $t \in]\eta; \sup J[$,

$$u(t) \in U \setminus K.$$

Un résultat analogue est vrai pour $\inf J$.

Démonstration. Raisonnons par l'absurde et supposons que les deux propriétés sont fausses. En particulier, $\sup J < \sup I$ donc $\sup J \in I$. Soit $K \subset U$ un compact dont u ne sort pas : pour tout $\eta < \sup J$, il existe $t \in]\eta; \sup J[$ tel que $u(t) \in K$.

Il existe alors (et on fixe pour le reste de la démonstration) $(t_n)_{n \in \mathbb{N}}$ une suite d'éléments de J telle que

$$\begin{aligned} t_n &\xrightarrow{n \rightarrow +\infty} \sup J; \\ u(t_n) &\in K, \quad \forall n \in \mathbb{N}. \end{aligned}$$

Puisque K est compact, on peut supposer, quitte à remplacer $(t_n)_{n \in \mathbb{N}}$ par une sous-suite, que $(u(t_n))_{n \in \mathbb{N}}$ converge vers un certain $u_{\text{lim}} \in K$.

La démonstration va être en deux étapes :

1. on montre que $u(t) \rightarrow u_{\text{lim}}$ quand $t \rightarrow \sup J$;
2. on en déduit que u peut se prolonger en une solution du problème (Cauchy) définie sur $J \cup \{\sup J\}$, ce qui contredit la maximalité de u .

Première étape : comme f est continue, elle est bornée sur un voisinage de $(u_{\text{lim}}, \sup J)$. Soient donc $M \in \mathbb{R}$ et $\epsilon > 0$ tels que

$$\forall (t, v) \in]\sup J - \epsilon; \sup J + \epsilon[\times B(u_{\text{lim}}, \epsilon), \quad \|f(t, v)\|_2 \leq M.$$

Intuitivement, cette inégalité entraîne que si, pour un certain n , t_n est proche de $\sup J$ et $u(t_n)$ est proche de u_{lim} , alors $u' = f(t, u)$ est bornée par M au voisinage de t_n ; en particulier, $\|u(t) - u(t_n)\|_2 \leq M|t - t_n|$ pour tout t dans un voisinage de t_n dont on peut estimer la taille. Ceci est formalisé par la proposition suivante (donc la démonstration est donnée à la fin de celle du théorème des bouts).

Proposition 4.8

Soit n n'importe quel entier tel que

$$|t_n - \sup J| < \frac{\epsilon}{2} \quad \text{et} \quad \|u(t_n) - u_{\text{lim}}\|_2 < \frac{\epsilon}{2}. \quad (4.5)$$

Pour tout $t \in]t_n - \frac{\epsilon}{2 \max(M, 1)}; t_n + \frac{\epsilon}{2 \max(M, 1)}[\cap J$,

$$\|u(t) - u(t_n)\|_2 \leq M|t - t_n|.$$

Puisque $(t_n, u(t_n)) \xrightarrow{n \rightarrow +\infty} (\sup J, u_{\text{lim}})$, on a, pour tout n assez grand,

$$|t_n - \sup J| < \frac{\epsilon}{2 \max(M, 1)} \quad \text{et} \quad \|u(t_n) - u_{\text{lim}}\|_2 < \frac{\epsilon}{2}.$$

Pour de telles valeurs de n , l'hypothèse (4.5) est vérifiée donc

$$\|u(t) - u(t_n)\|_2 \leq M|t - t_n|, \quad \forall t \in]t_n - \frac{\epsilon}{2 \max(M, 1)}; t_n + \frac{\epsilon}{2 \max(M, 1)}[\cap J.$$

Comme $t_n + \frac{\epsilon}{2 \max(M, 1)} > \sup J$, cela implique que, pour tout $t \in [t_n; \sup J[$,

$$\begin{aligned} \|u(t) - u_{\text{lim}}\|_2 &\leq \|u(t) - u(t_n)\|_2 + \|u(t_n) - u_{\text{lim}}\|_2 \\ &\leq M|t - t_n| + \|u(t_n) - u_{\text{lim}}\|_2 \\ &\leq M|t_n - \sup J| + \|u(t_n) - u_{\text{lim}}\|_2 \\ &\rightarrow 0 \text{ quand } n \rightarrow +\infty. \end{aligned}$$

Donc $u(t) \rightarrow u_{\text{lim}}$ quand $t \rightarrow \sup J$.

Deuxième étape : prolongeons u par continuité à $J \cup \{\sup J\}$, c'est-à-dire définissons

$$\begin{aligned} \bar{u} : J \cup \{\sup J\} &\rightarrow U \\ t &\rightarrow u(t) \quad \text{si } t < \sup J \\ &u_{\text{lim}} \quad \text{sinon.} \end{aligned}$$

C'est une application continue. Elle est dérivable sur J et

$$u'(t) = f(t, u(t)) \xrightarrow{t \rightarrow \sup J} f(\sup J, u_{\text{lim}}),$$

ce qui permet de montrer que u est également dérivable en $\sup J$, de dérivée $f(\sup J, u_{\text{lim}})$.

L'application \bar{u} est donc une solution du problème (**Cauchy**), qui prolonge u mais n'est pas égale à u . Cela contredit la maximalité de u . \square

Démonstration de la proposition 4.8. On va d'abord montrer que, pour tout $t \in \left[t_n; t_n + \frac{\epsilon}{2 \max(M, 1)} \right[\cap J$, $\|u(t) - u_{\text{lim}}\|_2 < \epsilon$. On peut supposer que l'ensemble $\{t \in J, t \geq t_n, \|u(t) - u_{\text{lim}}\|_2 \geq \epsilon\}$ est non-vidé, sinon cette propriété est automatiquement vraie. Définissons

$$T = \inf\{t \in J, t \geq t_n, \|u(t) - u_{\text{lim}}\|_2 \geq \epsilon\}$$

et montrons que $T \geq t_n + \frac{\epsilon}{2 \max(M, 1)}$. Supposons par l'absurde que ce n'est pas le cas.

Par continuité de u , on doit avoir $\|u(T) - u_{\text{lim}}\|_2 \geq \epsilon$. Pour tout $t \in [t_n; T[$, on a

$$\|u(t) - u_{\text{lim}}\|_2 < \epsilon.$$

Comme, de plus, $|t_n - \sup J| < \frac{\epsilon}{2}$ et $|t - t_n| \leq |T - t_n| < \frac{\epsilon}{2}$, on a $|t - \sup J| < \epsilon$, d'où

$$\|u'(t)\|_2 = \|f(t, u(t))\|_2 \leq M.$$

C'est aussi vrai en $t = T$ par continuité de u' . Cela montre que u est M -lipschitzienne sur $[t_n; T]$. En particulier,

$$\begin{aligned} \|u(T) - u_{\text{lim}}\|_2 &\leq \|u(T) - u(t_n)\|_2 + \|u(t_n) - u_{\text{lim}}\|_2 \\ &\leq M|T - t_n| + \|u(t_n) - u_{\text{lim}}\|_2 \\ &< M \frac{\epsilon}{2 \max(M, 1)} + \frac{\epsilon}{2} \\ &\leq \epsilon. \end{aligned}$$

Cela contredit l'inégalité $\|u(T) - u_{\text{lim}}\|_2 \geq \epsilon$.

On a donc montré que, pour tout $t \in \left[t_n; t_n + \frac{\epsilon}{2 \max(M, 1)} \right[\cap J$, $\|u(t) - u_{\text{lim}}\|_2 < \epsilon$. De même, on peut montrer que, pour tout $t \in \left] t_n - \frac{\epsilon}{2 \max(M, 1)}; t_n \right[$, $\|u(t_n) - u_{\text{lim}}\|_2 < \epsilon$.

Un raisonnement identique à celui qu'on vient de faire montre alors que, pour tout $t \in \left] t_n - \frac{\epsilon}{2 \max(M, 1)}; t_n + \frac{\epsilon}{2 \max(M, 1)} \right[\cap J$,

$$\|u'(t)\|_2 \leq M.$$

En conséquent, u est M -lipschitzienne sur l'intervalle considéré, ce qui entraîne le résultat annoncé. \square

L'exemple suivant montre comment le théorème des bouts permet de démontrer qu'une solution maximale d'une équation différentielle est globale.

Exemple 4.9

Considérons le problème (Cauchy), pour une fonction $f : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$. Supposons que f est continue, lipschitzienne par rapport à sa deuxième variable sur un voisinage de tout point, et vérifie l'inégalité

$$\|f(t, u)\|_2 \leq \|u\|_2, \quad \forall (t, u) \in \mathbb{R} \times \mathbb{R}^n. \quad (4.6)$$

Sa solution maximale est globale, c'est-à-dire définie sur \mathbb{R} tout entier.

Démonstration. Notons $u : J \rightarrow \mathbb{R}^n$ cette solution maximale et montrons que $J = \mathbb{R}$. Nous allons montrer que $\sup J = +\infty$; un raisonnement analogue montre que $\inf J = -\infty$.

Raisonnons par l'absurde et supposons que $\sup J < +\infty$. D'après le théorème des bouts, u sort de tout compact au voisinage de $\sup J$.

Considérons l'application $N : t \in J \rightarrow \|u(t)\|_2^2 \in \mathbb{R}$. C'est une application dérivable, dont la dérivée vérifie, en tout $t \in J$:

$$\begin{aligned} |N'(t)| &= |2 \langle u(t), u'(t) \rangle| \\ &= 2 |\langle u(t), f(t, u(t)) \rangle| \\ &\leq 2 \|u(t)\|_2 \|f(t, u(t))\|_2 \\ &\leq 2 \|u(t)\|_2^2 \\ &= 2N(t). \end{aligned}$$

Si on définit $N_2 : t \in J \rightarrow N(t)e^{-2t}$, on obtient alors que, pour tout t ,

$$N_2'(t) = (N'(t) - 2N(t))e^{-2t} \leq 0,$$

donc N_2 est décroissante et, pour tout $t \in]t_0; \sup J[$, $N_2(t) \leq N_2(t_0) = \|u_0\|_2^2 e^{-2t_0}$, ce qui entraîne

$$N(t) \leq (\|u_0\|_2 e^{t-t_0})^2.$$

(On aurait pu obtenir ce résultat via le lemme de Gronwall, sans définir explicitement la fonction N_2 .)

On en déduit que, pour tout $t \in]t_0; \sup J[$,

$$\|u(t)\|_2 \leq \|u_0\|_2 e^{t-t_0} \leq \|u_0\|_2 e^{\sup J - t_0}.$$

Si on pose $M = \|u_0\|_2 e^{\sup J - t_0}$, la fonction u ne sort pas du compact $\bar{B}(0, M)$. On a obtenu une contradiction. □

Le résultat énoncé dans l'exemple reste valable si on remplace la borne (4.6) par une majoration affine plus générale

$$\|f(t, u)\|_2 \leq C_1 \|u\|_2 + C_2, \quad \forall (t, u) \in \mathbb{R} \times \mathbb{R}^n,$$

pour des constantes $C_1, C_2 > 0$.

En revanche, il n'est plus valable si on remplace la borne « $\|u\|_2$ » par « $\|u\|_2^\alpha$ » pour une puissance $\alpha > 1$. Pour s'en convaincre, on peut considérer le problème de Cauchy suivant, portant sur une fonction u à valeurs dans \mathbb{R} :

$$\begin{aligned} u' &= |u|^\alpha, \\ u(0) &= 1. \end{aligned}$$

On peut vérifier que sa solution maximale est

$$u : \begin{array}{l}]-\infty; \frac{1}{\alpha-1}[\\ t \end{array} \rightarrow \begin{array}{l} \mathbb{R} \\ \frac{1}{(1-(\alpha-1)t)^{\frac{1}{\alpha-1}}} \end{array}$$

et n'est donc pas définie sur \mathbb{R} tout entier.

4.4 Régularité en la condition initiale

Dans cette section, on fait varier le couple (t_0, u_0) qui définit la condition initiale du problème (Cauchy). Cela définit une famille de solutions de l'équation différentielle « $u' = f(t, u)$ ». Lorsque f est C^2 , cette famille de solutions est différentiable par rapport à (t_0, u_0) . De plus, ses dérivées partielles peuvent être décrites comme les solutions d'un autre problème de Cauchy.

Pour simplifier les notations, on énonce d'abord ce résultat dans le cas où t_0 est fixé et seul u_0 varie. Le cas général est donné ensuite.

Théorème 4.10 : régularité par rapport à la condition initiale

Soient I un intervalle ouvert non-vide de \mathbb{R} , U un ouvert de \mathbb{R}^n et $f : I \times U \rightarrow \mathbb{R}^n$ une application. On suppose que f est de classe C^2 . Soit $t_0 \in I$. Pour tout $u_0 \in U$, on note $u_{u_0} : J_{u_0} \rightarrow U$ la solution maximale du problème de Cauchy

$$\begin{aligned} u'_{u_0} &= f(t, u_{u_0}), & (\text{Cauchy } u_0) \\ u_{u_0}(t_0) &= u_0. \end{aligned}$$

L'ensemble $\Omega = \{(u_0, t), u_0 \in U, t \in J_{u_0}\} \subset U \times I$ est ouvert et l'application

$$\begin{aligned} V : \quad \Omega &\rightarrow U \\ (u_0, t) &\rightarrow u_{u_0}(t) \end{aligned}$$

est de classe C^1 .

De plus, pour tout u_0 , $\frac{dV}{du_0}(u_0, \cdot) : J_{u_0} \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ est solution du problème de Cauchy suivant :

$$\begin{aligned} \frac{d}{dt} \left[\frac{dV}{du_0} \right] (u_0, t) &= \frac{df}{du}(t, V(u_0, t)) \circ \frac{dV}{du_0}(u_0, t), & (\text{Cauchy } \frac{dV}{du_0}) \\ \frac{dV}{du_0}(u_0, t_0) &= \text{Id}_{\mathbb{R}^n}. \end{aligned}$$

Remarque

Il ne faut pas tâcher d'apprendre par cœur le problème (Cauchy $\frac{dV}{du_0}$) dont $\frac{dV}{du_0}$ est solution. Il suffit de se souvenir du fait que V est C^1 . Ensuite, (Cauchy $\frac{dV}{du_0}$) se retrouve en dérivant (Cauchy u_0). En effet, (Cauchy u_0) peut se récrire en fonction de V comme

$$\begin{aligned} \frac{dV}{dt}(u_0, t) &= f(t, V(u_0, t)), \\ V(u_0, t) &= u_0. \end{aligned}$$

Dériver par rapport à u_0 les membres de gauche et de droite de chacune des deux égalités donne exactement (Cauchy $\frac{dV}{du_0}$).

Démonstration du théorème 4.10. Pour simplifier un peu, on va supposer que

f ne dépend pas de t . On peut faire cette hypothèse grâce au lemme qui suit (dont la démonstration se trouve dans l'annexe D.2) On note donc « $f(u)$ » plutôt que « $f(t, u)$ ».

Lemme 4.11

Si le théorème est vrai pour toutes les fonctions f indépendantes de t , il est également vrai pour toutes celles qui dépendent de t .

Le lemme qui suit simplifie davantage le problème en montrant qu'il suffit de démontrer la régularité de V au voisinage de chaque u_0 , pour des temps t proches de t_0 . Il est démontré dans l'annexe D.3.

Lemme 4.12

Supposons que

pour tout u_0 , Ω contient un voisinage de (u_0, t_0) , sur lequel V est C^1 et vérifie les équations (Cauchy $\frac{dV}{du_0}$). (4.7)

Alors Ω est un ouvert, V est C^1 sur tout Ω et y vérifie les équations (Cauchy $\frac{dV}{du_0}$).

Il ne reste qu'à vérifier que la propriété (4.7) est vérifiée. Soit $u_0 \in U$.
Première étape : V est définie sur un voisinage de (u_0, t_0) .

Soient $M_1, \epsilon > 0$ tels que $\overline{B}(u_0, \epsilon) \subset U$ et

$$\forall v \in B(u_0, \epsilon), \quad \|f(v)\|_2 \leq M_1.$$

La proposition qui suit, démontrée dans l'annexe D.4, implique que Ω contient $B(u_0, \frac{\epsilon}{2}) \times]t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1}[$.

Proposition 4.13

Pour tout $v \in B(u_0, \frac{\epsilon}{2})$,

$$]t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1}[\subset J_v.$$

De plus, pour tout $t \in]t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1}[$,

$$u_v(t) \in B(u_0, \epsilon).$$

Deuxième étape : V est lipschitzienne sur ce voisinage.

Pour tous $v, t \in B(u_0, \frac{\epsilon}{2}) \times]t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1}[$,

$$u'_v(t) = f(u_v(t)) \quad \Rightarrow \quad \|u'_v(t)\|_2 \leq M_1.$$

Donc, pour tout $v \in B(u_0, \frac{\epsilon}{2})$, u_v est M_1 -lipschitzienne sur $]t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1}[$, c'est-à-dire que V est M_1 -lipschitzienne par rapport à sa deuxième variable.

Soit $M_2 > 0$ tel que

$$\forall v \in B(u_0, \epsilon), \quad \left\| \frac{df}{du}(v) \right\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)} \leq M_2.$$

La fonction f est M_2 -lipschitzienne sur $B(u_0, \epsilon)$ d'après l'inégalité des accroissements finis. Ainsi, pour tous $v_1, v_2 \in B(u_0, \frac{\epsilon}{2})$, $t \in]t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1}[$,

$$\begin{aligned} \|u'_{v_1}(t) - u'_{v_2}(t)\|_2 &= \|f(u_{v_1}(t)) - f(u_{v_2}(t))\|_2 \\ &\leq M_2 \|u_{v_1}(t) - u_{v_2}(t)\|_2. \end{aligned}$$

En intégrant puis en utilisant l'inégalité triangulaire, on en déduit, pour tout $t \in [t_0; t_0 + \frac{\epsilon}{2M_1}[$,

$$\begin{aligned} \|u_{v_1}(t) - u_{v_2}(t)\|_2 &= \left\| u_{v_1}(t_0) - u_{v_2}(t_0) + \int_{t_0}^t (u'_{v_1}(s) - u'_{v_2}(s)) ds \right\|_2 \\ &\leq \|u_{v_1}(t_0) - u_{v_2}(t_0)\|_2 + \int_{t_0}^t \|u'_{v_1}(s) - u'_{v_2}(s)\|_2 ds \\ &\leq \|u_{v_1}(t_0) - u_{v_2}(t_0)\|_2 + \int_{t_0}^t M_2 \|u_{v_1}(s) - u_{v_2}(s)\|_2 ds. \end{aligned}$$

Ainsi, d'après le lemme de Gronwall (lemme D.1 en annexe), pour tout $t \in [t_0; t_0 + \frac{\epsilon}{2M_1}[$,

$$\|u_{v_1}(t) - u_{v_2}(t)\|_2 \leq \|u_{v_1}(t_0) - u_{v_2}(t_0)\|_2 e^{M_2(t-t_0)}$$

$$\begin{aligned}
&= \|v_1 - v_2\|_2 e^{M_2(t-t_0)} \\
&\leq \|v_1 - v_2\|_2 e^{\frac{\epsilon M_2}{2M_1}}.
\end{aligned}$$

Par symétrie, l'inégalité est aussi valable pour $t \in \left] t_0 - \frac{\epsilon}{2M_1}; t_0 \right]$, ce qui montre que V est $e^{\frac{\epsilon M_2}{2M_1}}$ -lipschitzienne par rapport à sa première variable sur $B(u_0, \frac{\epsilon}{2}) \times \left] t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1} \right]$. La fonction V est donc globalement lipschitzienne (et en particulier continue) sur cet ouvert.

Troisième étape : dérivabilité de V par rapport à t sur ce voisinage.

D'après sa définition, V est dérivable par rapport à sa deuxième variable et, pour tous v, t ,

$$\frac{dV}{dt}(v, t) = u'_v(t) = f(V(v, t)).$$

Comme f est continue sur U et V sur $B(u_0, \frac{\epsilon}{2}) \times \left] t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1} \right]$, la fonction $\frac{dV}{dt}$ est également continue sur ce dernier ouvert.

Quatrième étape : dérivabilité de V par rapport à u_0 sur ce voisinage.

Montrons que V admet une dérivée partielle par rapport à sa première variable, continue et solution du problème (Cauchy $\frac{dV}{du_0}$). Nous allons procéder « à l'envers », en considérant la solution du problème (Cauchy $\frac{dV}{du_0}$) et en montrant qu'elle est continue et qu'elle est la dérivée partielle de V par rapport à u_0 . Pour tout $v \in B(u_0, \frac{\epsilon}{2})$, soit donc $w_v : \tilde{I}_v \subset \left] t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1} \right[\rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ la solution maximale du problème

$$\begin{aligned}
w'_v(t) &= \frac{df}{du}(V(v, t)) \circ w_v(t) \\
w_v(t_0) &= \text{Id}_{\mathbb{R}^n}.
\end{aligned}$$

La solution maximale existe et est unique car, pour tout v , la fonction

$$(t, x) \in \left] t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1} \right[\times \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \rightarrow \frac{df}{du}(V(v, t)) \circ x$$

est lipschitzienne par rapport à x , avec une constante indépendante de t (puisque $\frac{df}{du} \circ V$ est continue et donc localement bornée).

Le même raisonnement que celui que nous venons de faire pour u_v dans la deuxième étape montre qu'il existe une constante $M_3 \geq M_1$ telle que, pour

tout $v \in B(u_0, \frac{\epsilon}{2})$, l'intervalle de définition de w_v contient

$$\left] t_0 - \frac{\epsilon}{2M_3}; t_0 + \frac{\epsilon}{2M_3} \right[$$

et que l'application $(v, t) \rightarrow w_v(t)$ est lipschitzienne et donc continue sur $B(u_0, \frac{\epsilon}{2}) \times \left] t_0 - \frac{\epsilon}{2M_3}; t_0 + \frac{\epsilon}{2M_3} \right[$ (c'est ce point de la démonstration qui utilise l'hypothèse selon laquelle f est C^2).

Montrons enfin que V est dérivable par rapport à sa première variable, et que, pour tous v, t dans l'ouvert considéré,

$$\frac{dV}{du_0}(v, t) = w_v(t).$$

Pour cela, nous allons effectuer une sorte de développement limité du problème (Cauchy u_0) à l'ordre 1 en u_0 .

Soient $v, h \in \mathbb{R}^n$ tels que $v, v + h \in B(u_0, \frac{\epsilon}{2})$. Considérons la fonction

$$\Delta : t \in \left] t_0 - \frac{\epsilon}{2M_3}; t_0 + \frac{\epsilon}{2M_3} \right[\rightarrow u_{v+h}(t) - u_v(t) - w_v(t)(h).$$

On a

$$\Delta(t_0) = (v + h) - v - \text{Id}_{\mathbb{R}^n}(h) = 0.$$

De plus, pour tout t ,

$$\begin{aligned} \Delta'(t) &= u'_{v+h}(t) - u'_v(t) - w'_v(t)(h) \\ &= f(u_{v+h}(t)) - f(u_v(t)) - \frac{df}{du}(u_v(t)) \circ w_v(t)(h) \\ &= \frac{df}{du}(u_v(t))(u_{v+h}(t) - u_v(t)) - \frac{df}{du}(u_v(t)) \circ w_v(t)(h) \\ &\quad + E(t) \\ &= \frac{df}{du}(u_v(t))(\Delta(t)) + E(t) \end{aligned}$$

avec $E(t) = f(u_{v+h}(t)) - f(u_v(t)) - \frac{df}{du}(u_v(t))(u_{v+h}(t) - u_v(t))$ et donc, par l'inégalité de Taylor avec reste intégral,

$$\|E(t)\|_2 \leq \frac{1}{2} \left(\sup_{\tilde{v} \in \bar{B}(u_0, \epsilon)} \left\| \frac{d^2 f}{du^2}(\tilde{v}) \right\|_{\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n))} \right) \|u_{v+h}(t) - u_v(t)\|_2^2.$$

Notons $C_1 = \frac{1}{2} \sup_{\tilde{v} \in \bar{B}(u_0, \epsilon)} \left\| \frac{d^2 f}{du^2}(\tilde{v}) \right\|_{\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n))}$ et C_2 la constante de Lipschitz de V par rapport à sa première variable (dont on a démontré l'existence il y a quelques paragraphes). Avec ces notations, on a, pour tout t ,

$$\|E(t)\|_2 \leq C_1 C_2 \|h\|_2^2$$

et donc

$$\left\| \Delta'(t) - \frac{df}{du}(u_v(t))(\Delta(t)) \right\|_2 \leq C_1 C_2 \|h\|_2^2.$$

En notant $C_3 = \sup_{\tilde{v} \in \bar{B}(u_0, \epsilon)} \left\| \frac{df}{du}(\tilde{v}) \right\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)}$, on en déduit

$$\|\Delta'(t)\|_2 \leq C_1 C_2 \|h\|_2^2 + C_3 \|\Delta(t)\|_2$$

puis, pour tout $t \in \left[t_0; t_0 + \frac{\epsilon}{2M_3} \right]$,

$$\begin{aligned} \|\Delta(t)\|_2 &= \left\| \Delta(t_0) + \int_{t_0}^t \Delta'(s) ds \right\|_2 \\ &= \left\| \int_{t_0}^t \Delta'(s) ds \right\|_2 \\ &\leq \int_{t_0}^t \|\Delta'(s)\|_2 ds \\ &\leq \int_{t_0}^t (C_1 C_2 \|h\|_2^2 + C_3 \|\Delta(s)\|_2) ds \\ &= C_1 C_2 \|h\|_2^2 (t - t_0) + \int_{t_0}^t C_3 \|\Delta(s)\|_2 ds. \end{aligned}$$

Le lemme de Gronwall (lemme D.1) nous permet donc d'affirmer que, pour tout $t \in \left[t_0; t_0 + \frac{\epsilon}{2M_3} \right]$,

$$\begin{aligned} \|\Delta(t)\|_2 &\leq C_1 C_2 \|h\|_2^2 (t - t_0) + C_1 C_2 C_3 \|h\|_2^2 \int_{t_0}^t e^{C_3(t-s)} (s - t_0) ds \\ &= \frac{C_1 C_2}{C_3} \|h\|_2^2 (e^{C_3(t-t_0)} - 1). \end{aligned}$$

Par symétrie, l'inégalité est aussi valable si $t \in \left] t_0 - \frac{\epsilon}{2M_3}; t_0 \right]$, à condition de remplacer « $e^{C_3(t-t_0)}$ » par « $e^{C_3|t-t_0|}$ » dans le membre de droite.

Si on pose $C_4 = \frac{C_1 C_2}{C_3} \left(e^{\frac{C_3 \epsilon}{2M_3}} - 1 \right)$, on a donc montré que, pour tous v, h tels que $v, v + h \in B(u_0, \frac{\epsilon}{2})$ et pour tout $t \in]t_0 - \frac{\epsilon}{2M_3}; t_0 + \frac{\epsilon}{2M_3}[$,

$$\|V(v + h, t) - V(v, t) - w_v(t)(h)\|_2 = \|\Delta(t)\|_2 \leq C_4 \|h\|_2^2.$$

La fonction V est donc dérivable par rapport à sa première variable, de dérivée $\frac{dV}{du_0}(v, t) = w_v(t)$.

Conclusion.

Nous avons vu que V était continue sur $B(u_0, \frac{\epsilon}{2}) \times]t_0 - \frac{\epsilon}{2M_3}; t_0 + \frac{\epsilon}{2M_3}[$, admettait sur cet ouvert des dérivées partielles par rapport à chacune de ses deux variables et que les dérivées partielles étaient continues. La fonction V est donc C^1 sur cet ouvert. À la quatrième étape, nous avons de plus montré que la dérivée partielle $\frac{dV}{du_0}$ était solution du problème (Cauchy $\frac{dV}{du_0}$). La propriété (4.7) est donc vraie. □

Théorème 4.14 : régularité, cas général

On garde les notations du théorème précédent ; f est toujours C^2 .

Pour toute paire $(t_0, u_0) \in I \times U$, on note $u_{t_0, u_0} : J_{t_0, u_0} \rightarrow U$ la solution maximale du problème de Cauchy

$$\begin{aligned} u'_{t_0, u_0} &= f(t, u_{t_0, u_0}), & (\text{Cauchy } (t_0, u_0)) \\ u_{t_0, u_0}(t_0) &= u_0. \end{aligned}$$

L'ensemble $\Omega = \{(t_0, u_0, t), t_0 \in I, u_0 \in U, t \in J_{t_0, u_0}\} \subset I \times U \times I$ est ouvert et l'application

$$\begin{aligned} V : \quad \Omega &\rightarrow U \\ (t_0, u_0, t) &\rightarrow u_{t_0, u_0}(t) \end{aligned}$$

est de classe C^1 .

De plus, les dérivées partielles de V sont solutions des problèmes de Cauchy suivants

$$\begin{aligned} \frac{d}{dt} \left[\frac{dV}{du_0} \right] (t_0, u_0, t) &= \frac{df}{du} (t, V(t_0, u_0, t)) \circ \frac{dV}{du_0} (t_0, u_0, t), \\ \frac{dV}{du_0} (t_0, u_0, t_0) &= \text{Id}_{\mathbb{R}^n}. \end{aligned}$$

$$\begin{aligned}\frac{d}{dt} \left[\frac{dV}{dt_0} \right] (t_0, u_0, t) &= \frac{df}{du}(t, V(t_0, u_0, t)) \left(\frac{dV}{dt_0}(t_0, u_0, t) \right), \\ \frac{dV}{dt_0}(t_0, u_0, t_0) &= -f(t_0, u_0).\end{aligned}$$

Ce théorème peut se déduire du précédent d'une manière similaire à la démonstration du lemme 4.11. On ne donne pas la démonstration ici.

Remarque

Un théorème encore plus général est vrai : on peut supposer que f est une fonction de trois variables et non deux, ce qui donne un problème de Cauchy de la forme

$$\begin{aligned}u' &= f(t, u, a), \\ u(t_0) &= u_0.\end{aligned}$$

Si f est C^2 , les solutions maximales de ce problème sont C^1 en (t_0, u_0, a) .

Chapitre 5

Résolution explicite dans des cas particuliers

5.1 Équations scalaires autonomes

Dans cette section, on considère une équation *scalaire* (l'inconnue u est à valeurs dans \mathbb{R} et non dans \mathbb{R}^n pour $n > 1$) et *autonome* (la fonction f ne dépend pas du temps). Il s'agit donc d'une égalité de la forme

$$u' = f(u), \tag{5.1}$$

pour une certaine fonction $f : U \rightarrow \mathbb{R}$, avec U un ouvert non-vide de \mathbb{R} . Dans toute la section, on suppose que f est localement lipschitzienne, de sorte que le théorème de Cauchy-Lipschitz s'applique. Nous allons décrire les solutions maximales de l'équation (5.1).

Commençons par les solutions les plus simples : les constantes.

Proposition 5.1

Supposons la fonction f de l'équation (5.1) localement lipschitzienne. Pour tout $u_0 \in U$, la fonction constante $u : t \in \mathbb{R} \rightarrow u_0$ est une solution maximale de l'équation différentielle si et seulement si $f(u_0) = 0$.

Démonstration. Soit $u_0 \in U$. Soit $u : t \in \mathbb{R} \rightarrow u_0$. Sa dérivée est nulle. Elle est donc solution de l'équation différentielle (5.1) si et seulement si

$$0 = f(u_0).$$

Lorsqu'elle est solution, elle est solution *maximale* car elle est définie sur \mathbb{R} tout entier. \square

Décrivons maintenant les solutions non-constantes en fonction des primitives de $\frac{1}{f}$. Considérons $u : J \rightarrow \mathbb{R}$ une solution maximale dont la dérivée n'est pas identiquement nulle. Soit $t_0 \in J$ tel que $u'(t_0) \neq 0$. Supposons pour simplifier que $f(u(t_0)) = u'(t_0) > 0$; un raisonnement très similaire est possible si $f(u(t_0)) < 0$.

Notons $] \alpha; \beta [$ l'intervalle maximal contenant $u(t_0)$ sur lequel f est strictement positive (avec éventuellement $\alpha = -\infty$ et $\beta = +\infty$).

Proposition 5.2

Pour tout $t \in J$, $u(t) \in] \alpha; \beta [$.

Démonstration. Raisonnons par l'absurde et supposons que ce n'est pas le cas. Puisque $u(t_0) \in] \alpha; \beta [$, la continuité de u et le théorème des valeurs intermédiaires entraînent qu'il existe $t_1 \in J$ tel que $u(t_1) = \alpha$ ou $u(t_1) = \beta$. Supposons par exemple $u(t_1) = \alpha$.

Alors u est solution du problème de Cauchy

$$\begin{aligned} u' &= f(u), \\ u(t_1) &= \alpha. \end{aligned}$$

La fonction constante $\tilde{u} : t \in \mathbb{R} \rightarrow \alpha$ est solution maximale de ce problème (en effet, $f(\alpha) = 0$, car $] \alpha; \beta [$ est un intervalle maximal sur lequel f est strictement positive). La solution maximale du problème étant unique, puisque f est localement lipschitzienne, $u = \tilde{u}$, c'est-à-dire que u est constante. C'est absurde. \square

Fixons $\Phi :] \alpha; \beta [\rightarrow \mathbb{R}$ une primitive de $\frac{1}{f}$: pour une constante C quelconque, on pose

$$\Phi(v) = C + \int_{u(t_0)}^v \frac{1}{f(s)} ds, \quad \forall v \in] \alpha; \beta [.$$

C'est une application continue, de dérivée strictement positive. Elle réalise donc un difféomorphisme vers son image, qui est un intervalle ouvert qu'on note $] \gamma; \delta [$.

On observe que, pour tout $t \in J$,

$$(\Phi \circ u)'(t) = \Phi'(u(t))u'(t) = \frac{u'(t)}{f(u(t))} = 1.$$

Donc, pour tout $t \in J$,

$$\Phi \circ u(t) = \Phi \circ u(t_0) + (t - t_0) = t - t_0 + C.$$

Ainsi, pour tout $t \in J$, $u(t) = \Phi^{-1}(t - t_0 + C)$.

Proposition 5.3

L'intervalle J est égal à $] \gamma + t_0 - C; \delta + t_0 - C[$.

Démonstration. Pour tout $t \in J$, puisque $\phi \circ u(t) = t - t_0 + C$, on doit avoir $t - t_0 + C \in] \gamma; \delta[$, donc $t \in] \gamma + t_0 - C; \delta + t_0 - C[$. Cela montre que $J \subset] \gamma + t_0 - C; \delta + t_0 - C[$.

Comme u est une solution maximale, elle est définie sur $] \gamma + t_0 - C; \delta + t_0 - C[$ tout entier. En effet, si ce n'était pas le cas, la fonction $\tilde{u} : t \in] \gamma + t_0 - C; \delta + t_0 - C[\rightarrow \Phi^{-1}(t - t_0 + C) \in U$ serait une solution de l'équation (5.1) qui la prolongerait strictement. \square

On a en fait obtenu le théorème suivant.

Théorème 5.4

Les solutions maximales non-constantes de l'équation (5.1) sont les fonctions de la forme

$$t \in] \gamma + D; \delta + D[\rightarrow \Phi^{-1}(t - D),$$

où Φ est une primitive de $\frac{1}{f}$, définie sur un intervalle maximal où f ne s'annule pas, $] \gamma; \delta[$ est l'image de Φ et $D \in \mathbb{R}$ est une constante arbitraire.

Exemple 5.5

Trouvons toutes les solutions maximales de l'équation différentielle

$$u' = -u^3.$$

La fonction $x \rightarrow -x^3$ s'annule uniquement en 0. La seule solution constante est donc $u \equiv 0$.

Cherchons maintenant les solutions non-constantes. Les intervalles maximaux où $x \rightarrow -x^3$ ne s'annule pas sont $] -\infty; 0[$ et $]0; +\infty[$. Sur ces intervalles, des primitives de $x \rightarrow \frac{1}{-x^3}$ sont

$$\Phi_1 : x \in] -\infty; 0[\rightarrow \frac{1}{2x^2},$$

$$\Phi_2 : x \in]0; +\infty[\rightarrow \frac{1}{2x^2}.$$

La première est une bijection entre $] -\infty; 0[$ et $]0; +\infty[$, de réciproque

$$\Phi_1^{-1} : x \in]0; +\infty[\rightarrow -\frac{1}{\sqrt{2x}} \in] -\infty; 0[$$

et la deuxième est une bijection entre $]0; +\infty[$ et $]0; +\infty[$, de réciproque

$$\Phi_2^{-1} : x \in]0; +\infty[\rightarrow \frac{1}{\sqrt{2x}} \in]0; +\infty[.$$

Les solutions maximales sont donc toutes les fonctions de la forme

$$u : t \in]D; +\infty[\rightarrow -\frac{1}{\sqrt{2(x-D)}}$$

$$\text{et } u : t \in]D; +\infty[\rightarrow \frac{1}{\sqrt{2(x-D)}}$$

pour D un réel quelconque.

5.2 Équations linéaires scalaires

Une équation différentielle linéaire scalaire est une équation de la forme

$$u'(t) = a(t)u(t) + b(t), \quad (5.2)$$

avec a, b des fonctions continues sur un intervalle $I \subset \mathbb{R}$. La fonction b peut être appelée « terme source ».

Commençons par résoudre cette équation dans le cas où b est nulle.

Proposition 5.6 : sans terme source

Soit $a : I \rightarrow \mathbb{R}$ une application continue, avec I un intervalle ouvert. Soit $A : I \rightarrow \mathbb{R}$ une primitive de a . Les solutions maximales de l'équation différentielle

$$u'(t) = a(t)u(t)$$

sont les fonctions de la forme $u : t \in I \rightarrow Ce^{A(t)}$, pour C un réel quelconque.

Démonstration. On vérifie qu'une fonction de la forme $t \rightarrow Ce^{A(t)}$ est nécessairement solution de l'équation. Elle est maximale car elle est définie sur I tout entier.

Réciproquement, si $u : J \rightarrow \mathbb{R}$ est une solution maximale, on pose $v : t \in J \rightarrow u(t)e^{-A(t)} \in \mathbb{R}$. Cette application est dérivable et, pour tout $t \in J$,

$$v'(t) = (u'(t) - A'(t)u(t))e^{-A(t)} = (u'(t) - a(t)u(t))e^{-A(t)} = 0.$$

Donc v est constante. Si on note C sa valeur, on a, pour tout $t \in J$, $u(t) = Ce^{A(t)}$. Puisque u est maximale, on doit avoir $J = I$; la fonction est donc bien de la forme indiquée. \square

Considérons maintenant l'équation (5.2) générale, sans supposer que b est nulle. Pour la résoudre, on utilise la méthode dite *de la variation de la constante*. Notons à nouveau $A : I \rightarrow \mathbb{R}$ une primitive de a . Pour une fonction dérivable $u : J \rightarrow \mathbb{R}$ avec J un sous-intervalle de I , on écrit u sous la forme

$$u(t) = v(t)e^{A(t)}$$

(en posant $v(t) = u(t)e^{-A(t)}$ pour tout t).

La fonction u est solution de l'équation si et seulement si, pour tout $t \in J$,

$$\begin{aligned} (v'(t) + a(t)v(t))e^{A(t)} &= u'(t) \\ &= a(t)u(t) + b(t) = a(t)v(t)e^{A(t)} + b(t), \end{aligned}$$

c'est-à-dire si et seulement si, pour tout t ,

$$v'(t) = b(t)e^{-A(t)},$$

c'est-à-dire si et seulement si, en notant B une primitive quelconque de $t \rightarrow b(t)e^{-A(t)}$, il existe un réel C tel que

$$v = C + B,$$

ce qui est équivalent à l'existence d'un réel C tel que, pour tout $t \in J$,

$$u(t) = Ce^{A(t)} + B(t)e^{A(t)}.$$

De ce raisonnement, on peut déduire le théorème suivant.

Théorème 5.7 : résolution des équations linéaires scalaires

Pour tout u_0 , la solution maximale du problème de Cauchy

$$\begin{aligned} u'(t) &= a(t)u(t) + b(t), \\ u(t_0) &= u_0, \end{aligned}$$

pour a, b des fonctions continues sur un intervalle ouvert I et u_0 un réel, est

$$u : t \in I \quad \rightarrow \quad u_0 e^{\int_{t_0}^t a(s) ds} + \int_{t_0}^t b(s) e^{\int_s^t a(\tau) d\tau} ds.$$

5.3 Équations linéaires en dimension générale

Dans cette section, on considère une équation différentielle linéaire de dimension $n \in \mathbb{N}^*$ quelconque, c'est-à-dire une équation de la forme

$$u'(t) = A(t)u(t) + b(t), \tag{5.3}$$

avec $A \in C^0(I, \mathbb{R}^{n \times n})$ et $b \in C^0(I, \mathbb{R}^n)$, pour I un intervalle de \mathbb{R} .

Proposition 5.8

Les solutions maximales de l'équation (5.3) sont globales (c'est-à-dire définies sur I tout entier).

Démonstration. La démonstration repose sur le théorème des bouts (théorème 4.7) ; elle est très similaire à celle de l'exemple 4.9.

Soit $u : J \rightarrow \mathbb{R}^n$ une solution maximale. Raisonnons par l'absurde et supposons que $J \neq I$, par exemple que $\sup J < \sup I$. Soit $\epsilon > 0$ tel que $[\sup J - \epsilon; \sup J + \epsilon] \subset I$. Soit $C > 0$ tel que, pour tout $t \in [\sup J - \epsilon; \sup J + \epsilon]$,

$$\|A(t)\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)}, \|b(t)\|_2 \leq C.$$

Une telle constante existe car A et b sont continues.

On en déduit que, pour tout t assez proche de $\sup J$,

$$\|u'(t)\|_2 \leq C(\|u(t)\|_2 + 1).$$

Notons $t_0 = \sup J - \epsilon$. Pour tout $t \in [t_0; \sup J[$,

$$\begin{aligned} \|u(t)\|_2 &= \left\| u(t_0) + \int_{t_0}^t u'(s) ds \right\|_2 \\ &\leq \|u(t_0)\|_2 + \int_{t_0}^t \|u'(s)\|_2 ds \\ &\leq \|u(t_0)\|_2 + \int_{t_0}^t C(\|u(s)\|_2 + 1) ds \\ &= \|u(t_0)\|_2 + C(t - t_0) + \int_{t_0}^t C\|u(s)\|_2 ds. \end{aligned}$$

Le lemme de Gronwall (lemme D.1 en annexe) nous permet alors d'affirmer que, pour tout $t \in [t_0; \sup J[$,

$$\|u(t)\|_2 \leq (\|u(t_0)\|_2 + 1) e^{C(t-t_0)} - 1 \leq (\|u(t_0)\|_2 + 1) e^{C\epsilon} - 1.$$

Donc u est bornée au voisinage de $\sup J$ et ne sort pas de tout compact de \mathbb{R}^n . Cela contredit le théorème des bouts. □

5.3.1 Sans terme source

Considérons d'abord l'équation sans terme source :

$$u'(t) = A(t)u(t), \tag{5.4}$$

avec $A \in C^0(I, \mathbb{R}^{n \times n})$.

Comme l'équation est linéaire en u , une combinaison linéaire de solutions est également solution : si $u_1, u_2 : I \rightarrow \mathbb{R}^n$ sont deux solutions et λ, μ sont des réels quelconques, $\lambda u_1 + \mu u_2$ est également une solution. En conséquent, pour tous $t_0, t \in I$, l'application qui, à un vecteur $u_0 \in \mathbb{R}^n$ associe $u_{u_0}(t)$, où u_{u_0} est la solution maximale du problème

$$u'(t) = A(t)u(t),$$

$$u(t_0) = u_0,$$

est linéaire. Notons $R(t, t_0) \in \mathbb{R}^{n \times n}$ sa matrice : pour tout u_0 ,

$$u_{u_0}(t) = R(t, t_0)u_0. \quad (5.5)$$

On appelle R la *résolvante* de l'équation (5.4).

Si on sait calculer la résolvante, alors on a accès (d'après l'équation (5.5)) à toutes les solutions maximales de notre équation différentielle (5.4). Malheureusement, en général, on ne sait pas calculer une expression explicite de R . Toutefois, on peut caractériser R comme la solution d'un certain problème de Cauchy. C'est ce qu'indique le théorème suivant.

Théorème 5.9

Pour tout $t_0 \in I$, $R(\cdot, t_0) : I \rightarrow \mathbb{R}^{n \times n}$ est la solution maximale du problème de Cauchy

$$\begin{aligned} \frac{dR}{dt}(t, t_0) &= A(t)R(t, t_0), \\ R(t_0, t_0) &= \text{Id}_n. \end{aligned}$$

Démonstration. Soit $t_0 \in I$ fixé. Soit $M : I \rightarrow \mathbb{R}^{n \times n}$ la solution maximale du problème de Cauchy en question :

$$\begin{aligned} M'(t) &= A(t)M(t), \\ M(t_0) &= \text{Id}_n. \end{aligned}$$

Elle est définie sur I tout entier d'après la proposition 5.8. Montrons que, pour tout $t \in I$, $M(t) = R(t, t_0)$.

D'après la définition de R (équation (5.5)), il faut montrer que, pour tout $u_0 \in \mathbb{R}^n$ et tout $t \in I$, $u_{u_0}(t) = M(t)u_0$. Fixons $u_0 \in \mathbb{R}^n$ et définissons $v : t \in I \rightarrow M(t)u_0$. Il s'agit d'une application différentiable, qui est solution du problème de Cauchy

$$\begin{aligned} v'(t) &= M'(t)u_0 = A(t)M(t)u_0 = A(t)v(t), \\ v(t_0) &= M(t_0)u_0 = u_0. \end{aligned}$$

Donc $v = u_{u_0}$ et on a bien, pour tout t , $u_{u_0}(t) = v(t) = M(t)u_0$. □

Remarque

Il est tentant de se dire, par analogie avec le cas scalaire, que la solution du problème

$$\begin{aligned}M'(t) &= A(t)M(t), \\M(t_0) &= \text{Id}_n\end{aligned}$$

est l'application $t \in I \rightarrow \exp\left(\int_{t_0}^t A(s)ds\right)$. Malheureusement, ceci n'est pas vrai (à moins que les matrices $A(s)$ ne commutent deux à deux), car, en général, pour $X, H \in \mathbb{R}^{n \times n}$, $d \exp(X)(H) \neq H \exp(X)$.

Avant de passer aux équations linéaires avec terme source, voici une propriété classique de la résolvante.

Proposition 5.10

Pour tous $t_1, t_2, t_3 \in I$, $R(t_3, t_2)R(t_2, t_1) = R(t_3, t_1)$.

Démonstration. Fixons $t_1, t_2, t_3 \in I$. Soit $u_1 \in \mathbb{R}^n$ quelconque. Montrons que

$$R(t_3, t_2)R(t_2, t_1)u_1 = R(t_3, t_1)u_1.$$

Notons $u_{u_1} : I \rightarrow \mathbb{R}^n$ la solution maximale du problème de Cauchy

$$\begin{aligned}u'_{u_1}(t) &= A(t)u_{u_1}(t), \\u_{u_1}(t_1) &= u_1.\end{aligned}$$

D'après la définition de R , $R(t_3, t_1)u_1 = u_{u_1}(t_3)$ et $R(t_2, t_1)u_1 = u_{u_1}(t_2)$.

Notons $u_2 = R(t_2, t_1)u_1$ et $u_{u_2} : I \rightarrow \mathbb{R}^n$ la solution maximale du problème de Cauchy

$$\begin{aligned}u'_{u_2}(t) &= A(t)u_{u_2}(t), \\u_{u_2}(t_2) &= u_2.\end{aligned}$$

D'après la définition de R , $R(t_3, t_2)R(t_2, t_1)u_1 = R(t_3, t_2)u_2 = u_{u_2}(t_3)$.

Or u_{u_1} est solution du problème de Cauchy qui définit u_{u_2} . En effet, $u_{u_1}(t_2) = R(t_2, t_1)u_1 = u_2$. Donc $u_{u_1} = u_{u_2}$ et

$$R(t_3, t_2)R(t_2, t_1)u_1 = u_{u_2}(t_3) = u_{u_1}(t_3) = R(t_3, t_1)u_1.$$

□

Corollaire 5.11

Pour tous $t_1, t_2 \in I$, $R(t_1, t_2)R(t_2, t_1) = R(t_1, t_1) = \text{Id}_n$ donc $R(t_2, t_1)$ est inversible, d'inverse $R(t_1, t_2)$.

5.3.2 Avec terme source

On revient maintenant à l'équation (5.3) générale, avec terme source :

$$u'(t) = A(t)u(t) + b(t). \quad (5.3)$$

Comme dans le cas scalaire, la méthode de variation de la constante permet d'en calculer les solutions. Soit $u : I \rightarrow \mathbb{R}^n$ une application quelconque. Soient $t_0 \in I$ et $v : I \rightarrow \mathbb{R}^n$ telle que, pour tout t ,

$$u(t) = R(t, t_0)v(t)$$

(c'est-à-dire qu'on pose $v(t) = R(t_0, t)u(t)$). La fonction u est solution de l'équation (5.3) si et seulement si, pour tout t ,

$$\begin{aligned} A(t)R(t, t_0)v(t) + R(t, t_0)v'(t) &= \frac{dR}{dt}(t, t_0)v(t) + R(t, t_0)v'(t) \\ &= u'(t) \\ &= A(t)u(t) + b(t) \\ &= A(t)R(t, t_0)v(t) + b(t). \end{aligned}$$

C'est équivalent au fait que, pour tout t , $R(t, t_0)v'(t) = b(t)$, c'est-à-dire au fait que v soit une primitive de $t \rightarrow R(t_0, t)b(t)$. La fonction u est donc solution si et seulement si il existe $v_0 \in \mathbb{R}^n$ tel que, pour tout $t \in I$,

$$v(t) = v_0 + \int_{t_0}^t R(t_0, s)b(s)ds,$$

ce qui est équivalent à

$$\begin{aligned} u(t) &= R(t, t_0)v_0 + \int_{t_0}^t R(t, t_0)R(t_0, s)b(s)ds \\ &= R(t, t_0)v_0 + \int_{t_0}^t R(t, s)b(s)ds. \end{aligned}$$

On en déduit le théorème suivant.

Théorème 5.12 : formule de Duhamel

Soient I un intervalle ouvert, $A \in C^0(I, \mathbb{R}^{n \times n})$, $b \in C^0(I, \mathbb{R}^n)$ et $u_0 \in \mathbb{R}^n$. La solution maximale du problème de Cauchy

$$\begin{aligned} u'(t) &= A(t)u(t) + b(t), \\ u(t_0) &= u_0 \end{aligned}$$

est

$$u : t \in I \rightarrow R(t, t_0)u_0 + \int_{t_0}^t R(t, s)b(s)ds.$$

Remarque

Si $n = 1$, la résolvante admet une expression explicite. En effet, pour tout t_0 , $R(\cdot, t_0)$ est la solution maximale du problème de Cauchy

$$\begin{aligned} \frac{dR}{dt}(t, t_0) &= A(t)R(t, t_0), \\ R(t_0, t_0) &= \text{Id}_1 = 1. \end{aligned}$$

(Notons bien que, si $n = 1$, A est une fonction à valeurs réelles.) Donc, pour tout t ,

$$R(t, t_0) = \exp\left(\int_{t_0}^t A(s)ds\right).$$

Si on remplace R par sa valeur dans la formule de Duhamel, on retrouve, comme il fallait s'y attendre, le théorème 5.7.

5.3.3 Coefficients constants

Exponentielle de matrices Lorsque A est une fonction constante, la résolvante admet une expression explicite. Pour la donner, il est nécessaire de rappeler la définition et les propriétés principales de l'exponentielle matricielle. L'exponentielle est définie de manière identique pour les matrices à coefficients réels et pour les matrices à coefficients complexes. On énonce ici la définition et les propriétés dans le cas général des coefficients complexes.

Définition 5.13 : exponentielle de matrice

Pour toute matrice $A \in \mathbb{C}^{n \times n}$, on définit

$$\exp(A) = \sum_{k=0}^{+\infty} \frac{A^k}{k!} \in \mathbb{C}^{n \times n}.$$

Cette définition est correcte, dans le sens que la série $\sum_{k=0}^{+\infty} \frac{A^k}{k!}$ converge dans $\mathbb{C}^{n \times n}$.

Proposition 5.14

1. Pour toute matrice $A \in \mathbb{C}^{n \times n}$, si les coefficients de A sont réels, alors les coefficients de $\exp(A)$ sont également réels.
2. Pour toutes $A, B \in \mathbb{C}^{n \times n}$, si A et B commutent (c'est-à-dire que $AB = BA$), alors

$$\exp(A + B) = \exp(A) \exp(B) = \exp(B) \exp(A).$$

3. Pour toutes $A, G \in \mathbb{C}^{n \times n}$ telles que G est inversible,

$$\exp(GAG^{-1}) = G \exp(A) G^{-1}.$$

4. Pour toute $A \in \mathbb{C}^{n \times n}$, l'application $h : t \in \mathbb{R} \rightarrow \exp(tA)$ est dérivable et

$$h'(t) = A \exp(tA) = \exp(tA)A, \quad \forall t \in \mathbb{R}.$$

Corollaire 5.15 : exponentielle d'une matrice diagonalisable

Soit $A \in \mathbb{C}^{n \times n}$. On suppose qu'il existe $G \in GL(n, \mathbb{C})$ et $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ tels que

$$A = G \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} G^{-1}.$$

Alors

$$\exp(A) = G \begin{pmatrix} e^{\lambda_1} & 0 & \dots & 0 \\ 0 & e^{\lambda_2} & & \\ \vdots & & \ddots & \\ 0 & & & e^{\lambda_n} \end{pmatrix} G^{-1}.$$

Démonstration. D'après la propriété 3 de la proposition 5.14,

$$A = G \exp \left[\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \right] G^{-1}.$$

De plus, pour tout $k \in \mathbb{N}$,

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n^k \end{pmatrix},$$

ce qui entraîne que

$$\begin{aligned} \exp \left[\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \right] &= \sum_{k=0}^{+\infty} \frac{1}{k!} \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n^k \end{pmatrix} \\ &= \begin{pmatrix} \sum_{k=0}^{+\infty} \frac{\lambda_1^k}{k!} & 0 & \dots & 0 \\ 0 & \sum_{k=0}^{+\infty} \frac{\lambda_2^k}{k!} & & \\ \vdots & & \ddots & \\ 0 & & & \sum_{k=0}^{+\infty} \frac{\lambda_n^k}{k!} \end{pmatrix} \\ &= \begin{pmatrix} e^{\lambda_1} & 0 & \dots & 0 \\ 0 & e^{\lambda_2} & & \\ \vdots & & \ddots & \\ 0 & & & e^{\lambda_n} \end{pmatrix}. \end{aligned}$$

□

Ce corollaire permet de calculer l'exponentielle de toute matrice diagonalisable. Pour les matrices qui ne sont pas diagonalisables, l'exponentielle peut être calculée à l'aide de la *décomposition de Dunford*. Indiquons brièvement les grandes lignes du calcul.

Soit $A \in \mathbb{R}^{n \times n}$ quelconque. Le point de départ de la méthode est d'écrire A sous la forme suivante :

$$A = G(D + N)G^{-1},$$

avec $G, D, N \in \mathbb{R}^{n \times n}$ des matrices telles que

- G est inversible ;
- D est diagonale ;
- N est nilpotente (c'est-à-dire qu'il existe $K \in \mathbb{N}^*$ tel que $N^K = 0$) ;
- N et D commutent.

C'est cette forme qu'on appelle *décomposition de Dunford*. Trouver G, D, N peut se faire en calculant les *sous-espaces caractéristiques* de A ; c'est au-delà du programme de ce cours.

Ensuite, la propriété 5.14 permet d'écrire

$$\exp(A) = G \exp(D + N)G^{-1} = G \exp(D) \exp(N)G^{-1}.$$

L'exponentielle de D se calcule comme dans le corollaire 5.15. Pour calculer $\exp(N)$, on utilise directement la définition : comme N est nilpotente, la somme infinie de la définition devient une somme finie, qu'on peut calculer explicitement. En notant K le plus petit entier tel que $N^K = 0$, on a en effet

$$\exp(N) = \sum_{k=0}^{+\infty} \frac{N^k}{k!} = \sum_{k=0}^{K-1} \frac{N^k}{k!}.$$

Coefficients constants Considérons le problème de Cauchy suivant, à coefficients constants

$$\begin{aligned} u'(t) &= Au(t) + b, \\ u(t_0) &= u_0. \end{aligned} \tag{5.6}$$

avec $A \in \mathbb{R}^{n \times n}$, $b, u_0 \in \mathbb{R}^n$.

Proposition 5.16

Pour tout $t_0 \in \mathbb{R}$, la résolvante de l'équation (5.6) vérifie

$$R(t, t_0) = \exp((t - t_0)A), \quad \forall t \in \mathbb{R}.$$

Démonstration. Pour tout t_0 , d'après le théorème 5.9, $R(\cdot, t_0)$ est la solution maximale de

$$\begin{aligned}\frac{dR}{dt}(t, t_0) &= AR(t, t_0), \\ R(t_0, t_0) &= \text{Id}_n.\end{aligned}$$

Il suffit de vérifier que $(t \in \mathbb{R} \rightarrow \exp((t - t_0)A))$ est cette solution maximale. En fait, il suffit de vérifier que $(t \in \mathbb{R} \rightarrow \exp((t - t_0)A))$ est solution du problème de Cauchy : si elle l'est, elle est nécessairement maximale puisqu'elle est définie sur \mathbb{R} .

Elle vérifie bien la condition initiale : $\exp((t_0 - t_0)A) = \exp(0_{n \times n}) = \text{Id}_n$.

De plus, d'après la propriété 4 et la proposition 5.14, cette application est dérivable et, en tout $t \in \mathbb{R}$, sa dérivée vaut

$$A \exp((t - t_0)A),$$

de sorte que $(t \in \mathbb{R} \rightarrow \exp((t - t_0)A))$ vérifie aussi la première égalité du problème de Cauchy. \square

Cette expression explicite pour la résolvante permet de résoudre explicitement le problème de Cauchy (5.6) en appliquant la formule de Duhamel.

Corollaire 5.17

La solution maximale du problème (5.6) est

$$u : t \in \mathbb{R} \quad \rightarrow e^{(t-t_0)A}u_0 + \int_{t_0}^t e^{(s-t_0)A}b \, ds.$$

Lorsque A est inversible, cela se simplifie en

$$u : t \in \mathbb{R} \quad \rightarrow e^{(t-t_0)A}u_0 + (e^{(t-t_0)A} - \text{Id}_n) A^{-1}b.$$

Chapitre 6

Équilibres des équations autonomes

6.1 Définitions

La notion de « point d'équilibre » a surtout un sens pour les problèmes *autonomes*, c'est-à-dire pour les problèmes (**Cauchy**) où f ne dépend pas de t . Dans ce chapitre, la fonction $f : U \rightarrow \mathbb{R}^n$ n'a donc qu'une seule variable et le problème de Cauchy devient

$$\begin{aligned} u' &= f(u), \\ u(t_0) &= u_0. \end{aligned} \quad (\text{Autonome})$$

On suppose que f est localement lipschitzienne, de façon à ce que le théorème de Cauchy-Lipschitz soit valide.

6.1.1 Flot

Définition 6.1 : flot de l'équation (Autonome)

Pour tout u_0 , on note u_{u_0} la solution maximale du problème (**Autonome**) pour $t_0 = 0$, et I_{u_0} son intervalle de définition. Pour tout $t \in I_{u_0}$, on définit

$$\phi_t(u_0) = u_{u_0}(t).$$

On appelle $(\phi_t)_{t \in \mathbb{R}}$ le *flot* de l'équation différentielle.

Remarque

L'ensemble de définition de ϕ_t dépend a priori de t . Pour tout t , il s'agit de

$$\{u_0 \in U, t \in I_{u_0}\}.$$

La façon la plus intuitive de se représenter le flot est la suivante. Imaginons que l'équation différentielle $u' = f(u)$ décrit l'évolution d'une certaine quantité physique (la position ou l'orientation d'un objet dans l'espace, par exemple). Pour tout $t \in \mathbb{R}$, ϕ_t représente l'action de cette évolution lorsqu'on la laisse se dérouler pendant t instants de temps : dans notre exemple, si un objet a une position u_0 à un instant de référence 0, il aura une position $\phi_t(u_0)$ à l'instant t .

Nous avons déjà aperçu le flot dans la section 4.4, avec une notation différente (il s'agissait de la fonction V). Les résultats de cette section impliquent d'ailleurs que, lorsque f est de classe C^2 , l'application ϕ_t est, pour tout t , définie sur un ouvert et de classe C^1 .

Dans la mesure où nous nous intéressons ici à une équation autonome, définir le flot en utilisant comme point de référence $t_0 = 0$ n'est pas une limitation : comme l'indique la proposition suivante, la solution du problème (**Autonome**) s'exprime en fonction de $(\phi_t)_{t \in \mathbb{R}}$ quel que soit t_0 .

Proposition 6.2

Pour tous $t_0 \in \mathbb{R}, u_0 \in U$, la solution maximale du problème (**Autonome**) est

$$t \in I_{u_0} + t_0 \quad \rightarrow \quad \phi_{t-t_0}(u_0) = u_{u_0}(t - t_0).$$

Démonstration. Notons $v_1 : I_1 \rightarrow U$ la solution maximale de (**Autonome**) et

$$v_2 : t \in I_2 \stackrel{\text{déf}}{=} I_{u_0} + t_0 \rightarrow u_{u_0}(t - t_0).$$

Pour tout t , $v_2'(t) = u'_{u_0}(t - t_0) = f(v_2(t))$ et $v_2(t_0) = u_{u_0}(0) = u_0$. On en déduit que v_2 est solution de (**Autonome**), donc

$$I_2 \subset I_1 \text{ et } v_2 = v_1 \text{ sur } I_2.$$

De même, l'application $t \in I_1 - t_0 \rightarrow v_1(t + t_0)$ est solution de l'équation (Autonome) lorsqu'on remplace t_0 par 0. Comme $u_{u_0} = v_2(\cdot + t_0)$ est solution maximale de cette équation,

$$I_1 - t_0 \subset I_{u_0} = I_2 - t_0 \text{ et } v_1(\cdot + t_0) = v_2(\cdot + t_0) \text{ sur } I_1 - t_0.$$

On en déduit $I_1 = I_2$ et $v_1 = v_2$. □

6.1.2 Portrait de phase

Définition 6.3 : orbites

On appelle *orbite* d'un point $u_0 \in U$ par le flot $(\phi_t)_{t \in \mathbb{R}}$ de l'équation (Autonome) l'ensemble

$$\mathcal{O}_{u_0} \stackrel{\text{d\'ef}}{=} \{\phi_t(u_0), t \in I_{u_0}\}.$$

L'ensemble des orbites forme une « partition » de U , c'est-à-dire que tous les points appartiennent à une orbite (tout point appartient au moins à sa propre orbite) et que deux orbites sont toujours soit disjointes (elles n'ont aucun point commun) soit confondues.¹ On appelle cette partition le *portrait de phase* de l'équation (Autonome).

Exemple 6.4

Considérons la fonction

$$\begin{aligned} f : \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (x, y) &\rightarrow (x(1-x), (1-2x)y) \end{aligned}$$

et l'équation autonome associée :

$$\begin{aligned} x' &= x(1-x), \\ y' &= (1-2x)y, \\ (x(0), y(0)) &= (x_0, y_0). \end{aligned}$$

1. En effet, si, pour deux points $u_0, u_1 \in U$, $\mathcal{O}_{u_0} \cap \mathcal{O}_{u_1} \neq \emptyset$, cela signifie qu'il existe $t_0 \in I_{u_0}, t_1 \in I_{u_1}$ tels que $\phi_{t_0}(u_0) = \phi_{t_1}(u_1)$. Avec le même raisonnement que dans la démonstration de la proposition 6.2, on voit que $I_{u_0} + t_1 - t_0 = I_{u_1}$ et, pour tout $t \in I_{u_0}$, $\phi_t(u_0) = \phi_{t+t_1-t_0}(u_1)$, ce qui entraîne que $\mathcal{O}_{u_0} = \mathcal{O}_{u_1}$.

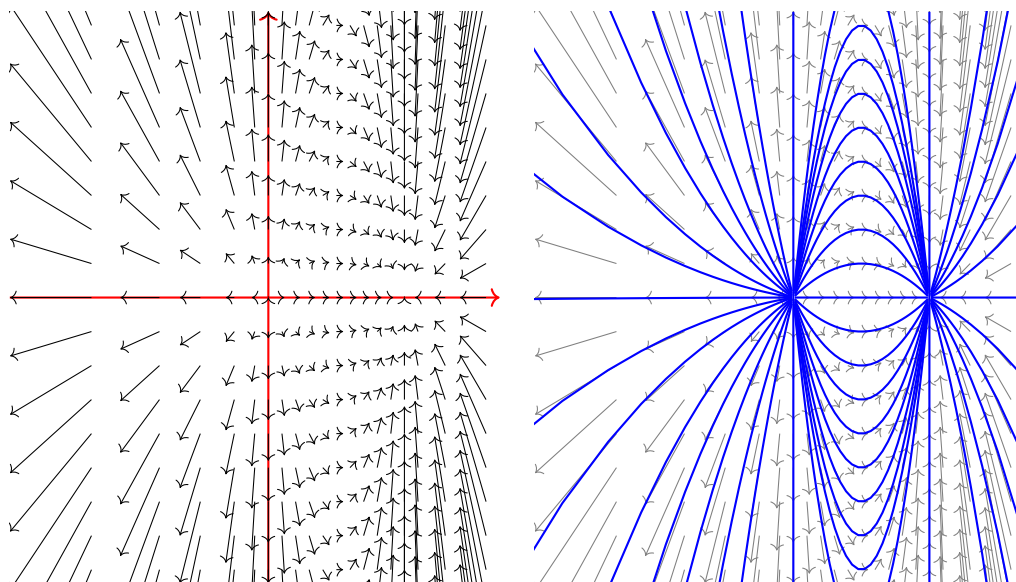


FIGURE 6.1 – À gauche, le champ de vecteurs $f(x, y) = (x(1-x), (1-2x)y)$; à droite, le portrait de phase correspondant.

Il se trouve que les solutions de cette équation peuvent être calculées de manière explicite :

- si $x_0 = 0$, alors $(x(t), y(t)) = (0, y_0 e^t)$ pour tout $t \in \mathbb{R}$;
- si $x_0 = 1$, alors $(x(t), y(t)) = (1, y_0 e^{-t})$ pour tout $t \in \mathbb{R}$;
- si $x_0 \neq 0, 1$ alors $(x(t), y(t)) = \left(\frac{x_0 e^t}{x_0 e^t + 1 - x_0}, \frac{y_0 e^t}{(x_0 e^t + 1 - x_0)^2} \right)$ pour tout $t \in I_{u_0}$, avec

$$\begin{aligned}
 I_{u_0} &= \begin{cases}]-\infty; \ln\left(1 + \frac{1}{|x_0|}\right) [& \text{si } x_0 < 0, \\ \mathbb{R} & \text{si } 0 < x_0 < 1, \\]\ln\left(1 - \frac{1}{x_0}\right); +\infty [& \text{si } 1 < x_0. \end{cases}
 \end{aligned}$$

Cela permet de tracer les orbites, comme sur la figure 6.1. Dans cette figure, la partie gauche représente la fonction f : en un certain nombre de points (x, y) du plan, on a dessiné $f(x, y)$ (qui est un vecteur de \mathbb{R}^2)

sous la forme d'une flèche, d'origine positionnée en (x, y) (c'est-à-dire que la flèche relie le point (x, y) et le point $(x, y) + f(x, y)$).^a La partie droite montre, superposées à la représentation de f , un certain nombre d'orbites. On remarque que les flèches sont tangentes aux orbites (c'est une conséquence du fait que les orbites sont des courbes données par des fonctions u satisfaisant $u' = f(u)$).

a. En fait, f a été légèrement normalisée pour des raisons de lisibilité : la flèche relie (x, y) et $(x, y) + 0, 2f(x, y)$. Sans le coefficient de normalisation, les flèches se chevauchent.

6.1.3 Équilibres

Définition 6.5 : équilibre

Un point $u_0 \in U$ est un *équilibre* de l'équation différentielle (**Autonome**) si $f(u_0) = 0$ (autrement dit, si l'application constante de valeur u_0 est solution de (**Autonome**)).

Nous allons nous intéresser spécifiquement au comportement des solutions de l'équation (**Autonome**) au voisinage des équilibres. Informellement, on dira qu'un équilibre est *stable* si toute trajectoire partant d'un point u_0 assez proche de l'équilibre reste proche de l'équilibre et qu'il est *asymptotiquement stable* si toute trajectoire partant d'un point u_0 assez proche de l'équilibre converge vers cet équilibre.

Définition 6.6 : stabilité

Si $u_0 \in U$ est un équilibre de l'équation (**Autonome**), on dit que u_0 est *stable* si, pour tout voisinage V_0 de u_0 , il existe un voisinage $V_1 \subset U$ de u_0 tel que

- pour tout $u_1 \in V_1$, $\phi_t(u_1)$ est défini pour tout $t \in \mathbb{R}^+$ (c'est-à-dire $\mathbb{R}^+ \subset I_{u_1}$);
- pour tous $u_1 \in V_1, t \in \mathbb{R}^+, \phi_t(u_1) \in V_0$.

On dit que u_0 est *asymptotiquement stable* s'il est stable et, de plus, il existe un voisinage $V_2 \subset U$ de u_0 tel que, pour tout $u_2 \in V_2$,

$$\phi_t(u_2) \xrightarrow{t \rightarrow +\infty} u_0.$$

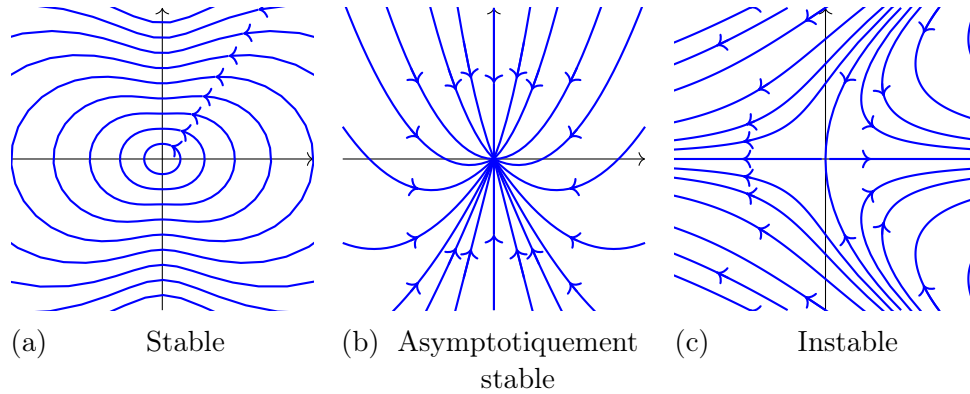


FIGURE 6.2 – Représentation graphique des trajectoires du système (Autonome) avec $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dans trois cas où $(0, 0)$ est un équilibre.

Si u_0 n'est pas stable, on dit qu'il est *instable*.

Une illustration de ces notions se trouve sur la figure 6.2.

6.2 Équations linéaires

Dans cette section, on étudie la stabilité d'un équilibre dans le cas d'une équation différentielle linéaire à coefficients constants :

$$u' = Au + b, \quad (6.1)$$

avec $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$.

Supposons que cette application a un équilibre z_0 . Quitte à translater,² on peut supposer que $z_0 = 0$ et donc que $0 = Az_0 + b = b$. L'équation est alors simplement

$$u' = Au. \quad (6.2)$$

On rappelle que, conformément au corollaire 5.17, le flot de tout $u_0 \in \mathbb{R}^n$ est

$$\phi_t(u_0) = \exp(tA)u_0, \quad \forall t \in \mathbb{R}.$$

Il faut donc étudier $\exp(tA)$.

2. « Translater » signifie considérer l'équation différentielle $v' = Av + b + Az_0$ au lieu de (6.1). Ses solutions sont les fonctions de la forme $u - z_0$, pour toute fonction u solution de (6.1). Cette équation admet 0 pour équilibre.

6.2.1 Cas diagonalisable

Traitons d'abord le cas où A est diagonalisable sur \mathbb{C} : il existe des nombres complexes $\lambda_1, \dots, \lambda_n$ et une matrice inversible $G \in \mathbb{R}^{n \times n}$ telle que

$$A = G \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} G^{-1}.$$

Pour tout $t \in \mathbb{R}$, d'après le corollaire 5.15,

$$\exp(tA) = G \begin{pmatrix} e^{t\lambda_1} & 0 & \dots & 0 \\ 0 & e^{t\lambda_2} & & \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & e^{t\lambda_n} \end{pmatrix} G^{-1}.$$

Fixons un vecteur $u_0 \in \mathbb{R}^n$. Notons

$$G^{-1}u_0 = \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix}.$$

Pour tout t ,

$$\phi_t(u_0) = \exp(tA)u_0 = G \begin{pmatrix} g_1 e^{t\lambda_1} \\ \vdots \\ g_n e^{t\lambda_n} \end{pmatrix}. \quad (6.3)$$

Théorème 6.7

Le point 0 est un équilibre stable de l'équation (6.2) si et seulement si

$$\operatorname{Re}(\lambda_k) \leq 0, \quad \forall k \in \{1, \dots, n\}.$$

C'est un équilibre asymptotiquement stable si et seulement si

$$\operatorname{Re}(\lambda_k) < 0, \quad \forall k \in \{1, \dots, n\}.$$

Démonstration. Supposons d'abord que

$$\operatorname{Re}(\lambda_k) \leq 0, \quad \forall k \in \{1, \dots, n\}$$

et montrons que 0 est un équilibre stable.

Pour tout $t \geq 0$,

$$|e^{t\lambda_k}| = e^{t\operatorname{Re}(\lambda_k)} |e^{it\operatorname{Im}(\lambda_k)}| = e^{t\operatorname{Re}(\lambda_k)} \leq 1, \quad \forall k \in \{1, \dots, n\}.$$

D'après l'équation (6.3), on a alors, pour tout u_0 et tout $t \geq 0$,

$$\begin{aligned} \|\phi_t(u_0)\|_2 &\leq \|G^{-1}\| \left\| \begin{pmatrix} g_1 e^{t\lambda_1} \\ \vdots \\ g_n e^{t\lambda_n} \end{pmatrix} \right\|_2 \\ &\leq \|G^{-1}\| \sqrt{|g_1|^2 + \dots + |g_n|^2} \\ &= \|G^{-1}\| \|Gu_0\|_2 \\ &\leq \|G^{-1}\| \|G\| \|u_0\|_2. \end{aligned} \tag{6.4}$$

Cette équation permet de démontrer la stabilité. En effet, considérons un voisinage $V_0 \subset \mathbb{R}^n$ de 0 quelconque. Soit $R > 0$ tel que $B(0, R) \subset V_0$. Définissons

$$V_1 = B\left(0, \frac{R}{\|G\| \|G^{-1}\|}\right).$$

D'après l'équation (6.4), pour tout $u_0 \in V_1$, $\phi_t(u_0) \in B(0, R) \subset V_0$ pour tout $t \geq 0$. Cela démontre la stabilité.

Supposons maintenant que

$$\operatorname{Re}(\lambda_k) < 0, \quad \forall k \in \{1, \dots, n\}$$

et montrons que 0 est un équilibre asymptotiquement stable. On a déjà montré qu'il était stable ; il faut montrer en outre qu'il existe un voisinage de 0 dont les trajectoires par le flot convergent toutes vers 0. Le raisonnement est similaire au précédent. En effet, pour tout k , puisque $\operatorname{Re}(\lambda_k) < 0$,

$$|e^{t\lambda_k}| = e^{t\operatorname{Re}(\lambda_k)} |e^{it\operatorname{Im}(\lambda_k)}| = e^{t\operatorname{Re}(\lambda_k)} \xrightarrow{t \rightarrow +\infty} 0,$$

donc $e^{t\lambda_k} \rightarrow 0$. En conséquence, pour tout u_0 ,

$$g_k e^{t\lambda_k} \xrightarrow{t \rightarrow +\infty} 0, \quad \forall k \in \{1, \dots, n\}.$$

L'équation (6.3) entraîne donc que $\phi_t(u_0) \xrightarrow{t \rightarrow +\infty} 0$ pour tout point initial u_0 . L'équilibre est asymptotiquement stable.

Supposons maintenant qu'il existe $k \in \{1, \dots, n\}$ tel que

$$\operatorname{Re}(\lambda_k) > 0$$

et montrons que 0 est un équilibre instable. Pour cela, nous allons montrer que tout voisinage de 0 contient un point u_0 tel que $\|\phi_t(u_0)\| \rightarrow +\infty$ quand $t \rightarrow +\infty$. Soit donc V un voisinage de 0.

Soit $u_0 \in \mathbb{R}^n$ tel que $g_k \neq 0$. Un tel vecteur u_0 existe : si la k -ème ligne de G^{-1} (qu'on note $(G^{-1})_{k,:}$) n'est pas imaginaire pure, on peut prendre $u_0 = \operatorname{Re}((G^{-1})_{k,:})$ (car alors $\operatorname{Re}((G^{-1}u_0)_k) \neq 0$, donc $g_k \neq 0$). Si toutes ses coordonnées sont imaginaires pures, on peut prendre $u_0 = \operatorname{Im}((G^{-1})_{k,:})$ (car alors $\operatorname{Im}((G^{-1}u_0)_k) \neq 0$, donc $g_k \neq 0$).

Quitte à multiplier u_0 par une constante assez petite, on peut supposer que $u_0 \in V$. D'après l'équation (6.3), $\|G^{-1}\phi_t(u_0)\|_2 \rightarrow +\infty$ quand $t \rightarrow +\infty$. En effet, la k -ème coordonnée de ce vecteur est $g_k e^{t\lambda_k}$, et

$$|g_k e^{t\lambda_k}| = |g_k| e^{t\operatorname{Re}(\lambda_k)} \xrightarrow{t \rightarrow +\infty} +\infty.$$

Or, pour tout t , $\|\phi_t(u_0)\|_2 \geq \frac{\|G^{-1}\phi_t(u_0)\|_2}{\|G^{-1}\|}$. Donc $\|\phi_t(u_0)\|_2 \rightarrow +\infty$ quand $t \rightarrow +\infty$, ce qui finit de démontrer l'instabilité.

De manière similaire, supposons qu'il existe $k \in \{1, \dots, n\}$ tel que

$$\operatorname{Re}(\lambda_k) \geq 0$$

et montrons que 0 n'est pas asymptotiquement stable. On considère à nouveau un voisinage V quelconque de 0 et un point $u_0 \in V$ tel que $g_k \neq 0$. Alors

$$|g_k e^{t\lambda_k}| = |g_k| e^{t\operatorname{Re}(\lambda_k)} \not\rightarrow 0 \quad \text{quand } t \rightarrow +\infty,$$

donc $\|\phi_t(u_0)\|_2 \not\rightarrow 0$ quand $t \rightarrow +\infty$, de sorte qu'il existe au moins un point de V dont la trajectoire par le flot de l'équation (6.2) ne tend pas vers 0. \square

6.2.2 Cas non-diagonalisable

Dans cette sous-section, nous allons étendre les résultats précédents au cas où A n'est pas diagonalisable sur \mathbb{C} . Un résultat classique d'algèbre linéaire

affirme que A est trigonalisable et même, plus précisément, que A peut s'écrire sous la forme

$$A = G \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & B_K \end{pmatrix} G^{-1},$$

où, pour tout $k \in \{1, \dots, K\}$, B_k est une matrice carrée, dont on note $n_k \times n_k$ la dimension, de la forme

$$B_k = \begin{pmatrix} \lambda_k & \star & \dots & \star \\ 0 & \lambda_k & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \dots & 0 & \lambda_k \end{pmatrix},$$

pour un certain $\lambda_k \in \mathbb{C}$. On note N_k la partie triangulaire strictement supérieure de B_k , de sorte que $B_k = \lambda_k \text{Id}_{n_k} + N_k$.

Pour tout vecteur $u_0 \in \mathbb{R}^n$, notons

$$G^{-1}u_0 = \begin{pmatrix} g_1 \\ \vdots \\ g_K \end{pmatrix},$$

où, cette fois, les g_1, \dots, g_K sont des vecteurs, de longueurs n_1, n_2, \dots, n_K . De manière analogue à l'équation (6.3), la proposition 5.14 entraîne que, pour tout $t \geq 0$,

$$\phi_t(u_0) = \exp(tA)u_0 = G \begin{pmatrix} \exp(tB_1)g_1 \\ \vdots \\ \exp(tB_K)g_K \end{pmatrix}. \quad (6.5)$$

Il va falloir calculer les $\exp(tB_k)$. Pour tout k , $B_k = \lambda_k \text{Id}_{n_k} + N_k$ et, comme $\lambda_k \text{Id}_{n_k}$ et N_k commutent,

$$\exp(tB_k) = \exp(t\lambda_k \text{Id}_{n_k}) \exp(tN_k) = e^{t\lambda_k} \exp(tN_k).$$

Comme N_k est nilpotente, $t \rightarrow \exp(tN_k)$ est une fonction polynomiale. Cette fonction est constante (égale à Id_{n_k}) si N_k est nulle et non-constante sinon.³

Nous avons maintenant tous les résultats préliminaires nécessaires pour énoncer et démontrer le résultat de stabilité suivant.

3. Une justification rapide de la non-constance est que sa dérivée $t \rightarrow N_k \exp(tN_k)$ n'est pas nulle si $N_k \neq 0$.

Théorème 6.8

Le point 0 est un équilibre stable de l'équation (6.2) si et seulement si, pour tout k ,

$$(\operatorname{Re}(\lambda_k) < 0) \quad \text{ou} \quad (\operatorname{Re}(\lambda_k) = 0 \text{ et } N_k = 0).$$

C'est un équilibre asymptotiquement stable si et seulement si, pour tout k ,

$$\operatorname{Re}(\lambda_k) < 0.$$

Démonstration. Supposons que, pour tout $k \in \{1, \dots, K\}$,

$$(\operatorname{Re}(\lambda_k) < 0) \quad \text{ou} \quad (\operatorname{Re}(\lambda_k) = 0 \text{ et } N_k = 0).$$

Montrons que 0 est un équilibre stable. De la même manière que dans la démonstration du théorème 6.7, il suffit de montrer l'existence d'une constante $C > 0$ telle que, pour tout $u_0 \in \mathbb{R}^n$ et tout $t \geq 0$,

$$\|\phi_t(u_0)\|_2 \leq C \|u_0\|. \quad (6.6)$$

Pour tout k et pour tout t , puisque $\exp(tB_k) = e^{t\lambda_k} \exp(tN_k)$,

$$\|\exp(tB_k)\| = |e^{t\lambda_k}| \|\exp(tN_k)\| = e^{t\operatorname{Re}(\lambda_k)} \|\exp(tN_k)\|.$$

Pour tout k , si $\operatorname{Re}(\lambda_k) < 0$,

$$e^{t\operatorname{Re}(\lambda_k)} \|\exp(tN_k)\| \xrightarrow{t \rightarrow +\infty} 0$$

par croissance dominée : le terme exponentiel $e^{t\operatorname{Re}(\lambda_k)}$ tend vers 0, tandis que le terme $\|\exp(tN_k)\|$ est majoré par un polynôme en t . Comme $t \rightarrow e^{t\operatorname{Re}(\lambda_k)} \|\exp(tN_k)\|$ est continue, le fait qu'elle tende vers 0 en $+\infty$ entraîne qu'elle est bornée sur \mathbb{R}^+ . Soit M_k un majorant.

Pour tout k , si $\operatorname{Re}(\lambda_k) = 0$ et $N_k = 0$, on a pour tout t

$$\|\exp(tB_k)\| = |e^{t\lambda_k}| = 1.$$

On note alors $M_k = 1$.

Finalement, on définit $M = \max(M_1, \dots, M_K)$. D'après l'équation (6.5), on a, pour tout $u_0 \in \mathbb{R}^n$ et tout $t \geq 0$,

$$\|\phi_t(u_0)\|_2 \leq M \|G\| \sqrt{\|g_1\|_2^2 + \dots + \|g_K\|_2^2}$$

$$= M \|G\| \|G^{-1}u_0\|_2 \leq M \|G\| \|G^{-1}\| \|u_0\|_2.$$

Cela démontre l'équation (6.6) et donc la stabilité.

Le raisonnement est similaire, mais plus simple, pour montrer la stabilité asymptotique. Supposons que, pour tout $k \in \{1, \dots, K\}$,

$$\operatorname{Re}(\lambda_k) < 0.$$

On vient de démontrer qu'alors, l'équilibre était stable. Il faut montrer que 0 admet un voisinage dont toutes les trajectoires par le flot tendent vers 0. Comme on l'a déjà vu, pour tout k ,

$$\|\exp(tB_k)\| \xrightarrow{t \rightarrow +\infty} 0.$$

Ainsi, pour tout $u_0 \in \mathbb{R}^n$, d'après l'équation (6.5),

$$\|\phi_t(u_0)\|_2 \xrightarrow{t \rightarrow +\infty} 0.$$

Cela conclut.

Supposons maintenant qu'il n'est pas vrai que, pour tout $k \in \{1, \dots, K\}$,

$$(\operatorname{Re}(\lambda_k) < 0) \quad \text{ou} \quad (\operatorname{Re}(\lambda_k) = 0 \text{ et } N_k = 0)$$

et montrons que 0 est instable. L'hypothèse entraîne que, pour un certain k ,

$$(\operatorname{Re}(\lambda_k) > 0) \quad \text{ou} \quad (\operatorname{Re}(\lambda_k) = 0 \text{ et } N_k \neq 0).$$

Fixons un tel k . Soit $V \subset \mathbb{R}^n$ un voisinage quelconque de 0.

Commençons par supposer que $\operatorname{Re}(\lambda_k) > 0$. Soit $u_0 \in \mathbb{R}^n$ tel que $g_k \neq 0$. Quitte à multiplier u_0 par une constante, on peut supposer que $u_0 \in V$. Pour tout t assez grand,

$$\|\exp(tN_k)g_k\|_2 \geq \|g_k\|_2.$$

En effet, $t \rightarrow \exp(tN_k)g_k$ est une fonction polynomiale. Soit elle est non-constante, et alors $\|\exp(tN_k)g_k\|_2 \rightarrow +\infty$ quand $t \rightarrow +\infty$, soit elle est constante et alors, pour tout t , $\|\exp(tN_k)g_k\|_2 = \|\exp(0N_k)g_k\|_2 = \|g_k\|_2$.

Donc

$$\|\exp(tB_k)g_k\|_2 = e^{t\operatorname{Re}(\lambda_k)} \|\exp(tN_k)g_k\|_2 \xrightarrow{t \rightarrow +\infty} +\infty.$$

D'après l'équation (6.5), $\|\phi_t(u_0)\|_2 \xrightarrow{t \rightarrow +\infty} +\infty$, c'est-à-dire que $(\phi_t(u_0))_{t \in \mathbb{R}^+}$ ne reste dans aucun voisinage de 0 : l'équilibre est instable.

Supposons maintenant que $\operatorname{Re}(\lambda_k) = 0$ et $N_k \neq 0$. Soit $u_0 \in \mathbb{R}^n$ tel que $N_k g_k \neq 0$ (un tel u_0 existe, par un raisonnement similaire à un argument de la démonstration du théorème 6.7). On peut supposer qu'il appartient à V . Alors $t \rightarrow \exp(tN_k)g_k$ est une fonction polynomiale non-constante (sa dérivée en 0 vaut $N_k g_k \neq 0$) et

$$\|\exp(tN_k)g_k\|_2 \xrightarrow{t \rightarrow +\infty} +\infty.$$

On a alors aussi $\|\exp(tB_k)g_k\|_2 = \|\exp(tN_k)g_k\|_2 \xrightarrow{t \rightarrow +\infty} +\infty$, ce qui entraîne que $\|\phi_t(u_0)\|_2 \xrightarrow{t \rightarrow +\infty} +\infty$ et achève de démontrer l'instabilité.

Pour finir, supposons qu'il existe $k \in \{1, \dots, K\}$ tel que $\operatorname{Re}(\lambda_k) \geq 0$ et montrons que l'équilibre n'est pas asymptotiquement stable. Si $\operatorname{Re}(\lambda_k) > 0$ ou $\operatorname{Re}(\lambda_k) = 0$ et $N_k \neq 0$, alors l'équilibre n'est pas stable, comme on vient de le démontrer. Le seul cas qui reste à traiter est $\operatorname{Re}(\lambda_k) = 0$ et $N_k = 0$. Soit $V \subset \mathbb{R}^n$ un voisinage quelconque de 0.

Soit $u_0 \in V$ tel que $g_k \neq 0$. Alors, pour tout $t \geq 0$,

$$\|\exp(tB_k)g_k\|_2 = \|e^{t\lambda_k}g_k\|_2 = \|g_k\|_2.$$

Donc, d'après l'équation (6.5), $\|\phi_t(u_0)\|_2 \not\rightarrow 0$ quand $t \rightarrow +\infty$. L'équilibre n'est pas asymptotiquement stable. \square

Dans le cas où l'équilibre n'est pas stable, on peut raffiner le raisonnement précédent pour déterminer quelles trajectoires du flot tendent vers 0. Le résultat auquel on aboutit (et qu'on ne démontrera pas) s'énonce le plus simplement lorsque A est *hyperbolique*, conformément à la définition suivante.

Définition 6.9 : hyperbolicité

On dit que A est *hyperbolique* si toutes ses valeurs propres complexes sont de partie réelle non-nulle :

$$\operatorname{Re}(\lambda_k) \neq 0, \quad \text{pour tout } k \in \{1, \dots, K\}.$$

Théorème 6.10 : espaces stable et instable

Supposons que A est hyperbolique. Définissons

$$E^s = \{u_0 \in \mathbb{R}^n \text{ tq } g_k = 0 \text{ pour tout } k \text{ tq } \operatorname{Re}(\lambda_k) > 0\};$$

$$E^u = \{u_0 \in \mathbb{R}^n \text{ tq } g_k = 0 \text{ pour tout } k \text{ tq } \operatorname{Re}(\lambda_k) < 0\}.$$

(On appelle respectivement ces espaces sous-espaces *stable* et *instable* de A .)

Alors

$$E^s = \{u_0 \in \mathbb{R}^n, \phi_t(u_0) \xrightarrow{t \rightarrow +\infty} 0\},$$

$$E^u = \{u_0 \in \mathbb{R}^n, \phi_t(u_0) \xrightarrow{t \rightarrow -\infty} 0\}.$$

De plus, ces espaces sont supplémentaires : $\mathbb{R}^n = E^s \oplus E^u$.

6.2.3 Représentation graphique en dimension 2

Dans cette sous-section, on représente explicitement les trajectoires lorsque A est une matrice hyperbolique de taille 2×2 . On distingue les cas de la manière suivante.

1. A est diagonalisable à valeurs propres réelles. Dans ce cas, quitte à effectuer un changement de base,⁴ on peut supposer que

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

avec $\lambda_1 \leq \lambda_2$ les valeurs propres. Le flot d'un point $u_0 = (x_0, y_0)$ est

$$\phi_t(u_0) = (x_0 e^{\lambda_1 t}, y_0 e^{\lambda_2 t}), \quad \forall t \in \mathbb{R}.$$

On peut remarquer, pour tracer plus facilement le portrait de phase, que le flot de u_0 est inclus dans le graphe de la fonction

$$x \in \mathbb{R} \rightarrow \frac{y_0}{|x_0|^{\lambda_2/\lambda_1}} |x|^{\lambda_2/\lambda_1} \in \mathbb{R}.$$

(Remarquons que λ_2/λ_1 peut être positif ou négatif, selon que λ_1 et λ_2 sont ou non de même signe ; cela influe beaucoup sur l'allure du graphe.)

- (a) $0 < \lambda_1 \leq \lambda_2$: voir la figure 6.3a. Toutes les trajectoires divergent (à part celle qui reste en 0).

4. Le changement de base n'est pas nécessairement orthogonal. Cela déforme donc un peu la figure, sans en changer les principales propriétés.

- (b) $\lambda_1 < 0 < \lambda_2$: voir la figure 6.3b. On observe l'espace stable E^s (la droite des abscisses) et l'espace instable E^u (la droite des ordonnées).
- (c) $\lambda_1 \leq \lambda_2 < 0$: voir la figure 6.3c. C'est un cas asymptotiquement stable : toutes les trajectoires convergent vers 0.
2. A est diagonalisable à valeurs propres non-réelles. Notons $\lambda \in \mathbb{C}$ l'une des valeurs propres. L'autre est $\bar{\lambda}$. On peut montrer que, quitte à effectuer un bon changement de base,

$$A = \begin{pmatrix} \operatorname{Re}(\lambda) & \operatorname{Im}(\lambda) \\ -\operatorname{Im}(\lambda) & \operatorname{Re}(\lambda) \end{pmatrix}.$$

On peut vérifier que, pour tout t ,

$$\exp(tA) = e^{t\operatorname{Re}(\lambda)} \begin{pmatrix} \cos(t\operatorname{Im}(\lambda)) & \sin(t\operatorname{Im}(\lambda)) \\ -\sin(t\operatorname{Im}(\lambda)) & \cos(t\operatorname{Im}(\lambda)) \end{pmatrix},$$

ce qui est la composée d'une rotation d'angle $t\operatorname{Im}(\lambda)$ et d'une homothétie de rapport $\exp(t\operatorname{Re}(\lambda))$.

- (a) $\operatorname{Re}(\lambda) > 0$: voir la figure 6.3d. Toutes les trajectoires divergent (à part celle qui reste en 0).
- (b) $\operatorname{Re}(\lambda) < 0$: voir la figure 6.3e. C'est un cas asymptotiquement stable : toutes les trajectoires convergent vers 0.
3. A n'est pas diagonalisable.

Dans ce cas, A n'a qu'une seule valeur propre (pour tout n , une matrice de taille $n \times n$ avec n valeurs propres distinctes est diagonalisable). Cette valeur propre est donc réelle (si elle admet une valeur propre z non-réelle, elle admet aussi \bar{z} pour valeur propre et a donc en fait deux valeurs propres distinctes). Donc A est trigonalisable sur \mathbb{R} . En fait, en choisissant un bon changement de base, on peut supposer que

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

où λ est la valeur propre.

Alors, pour tout t ,

$$\exp(tA) = \exp(t\lambda\operatorname{Id}_2) \exp\left(t \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right) = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}.$$

Le flot d'un point $u_0 = (x_0, y_0)$ est

$$\phi_t(u_0) = ((x_0 + ty_0)e^{\lambda t}, y_0 e^{\lambda t}), \quad \forall t \in \mathbb{R}.$$

- (a) $\lambda > 0$: voir la figure 6.3f. Toutes les trajectoires divergent (à part celle qui reste en 0).
- (b) $\lambda < 0$: voir la figure 6.3g. C'est un cas asymptotiquement stable : toutes les trajectoires convergent vers 0.

6.3 Équations non-linéaires

Dans cette section, on revient à l'équation (Autonome) en toute généralité, sans supposer que f est linéaire. On énonce et on démontre en partie un théorème qui généralise une partie des résultats que nous avons vus dans le cas linéaire.

Théorème 6.11

Supposons que la fonction f de l'équation (Autonome) est de classe C^1 . Soit $u_0 \in U$ un équilibre.

Si toutes les valeurs propres (sur \mathbb{C}) de la matrice jacobienne $Jf(u_0)$ ont une partie réelle strictement négative, alors u_0 est asymptotiquement stable.

Si une valeur propre de la matrice jacobienne $Jf(u_0)$ est de partie réelle strictement positive, alors u_0 est instable.

Démonstration partielle. On va se contenter de démontrer la stabilité asymptotique dans le cas où toutes les valeurs propres sont de partie réelle strictement négative.

Quitte à traduire, on peut supposer $u_0 = 0$ et on fait cette hypothèse dans la suite pour alléger les notations.

Supposons que toutes les valeurs propres de $Jf(0)$ ont une partie réelle strictement négative et montrons que 0 est asymptotiquement stable. Le principe de la démonstration est d'exhiber ce qu'on appelle une *fonction de Lyapunov* du système, c'est-à-dire une fonction de U vers \mathbb{R} , qui décroît le long des trajectoires de l'équation (Autonome). Cette propriété de décroissante garantit que les ensembles de sous-niveau de la fonction de Lyapunov

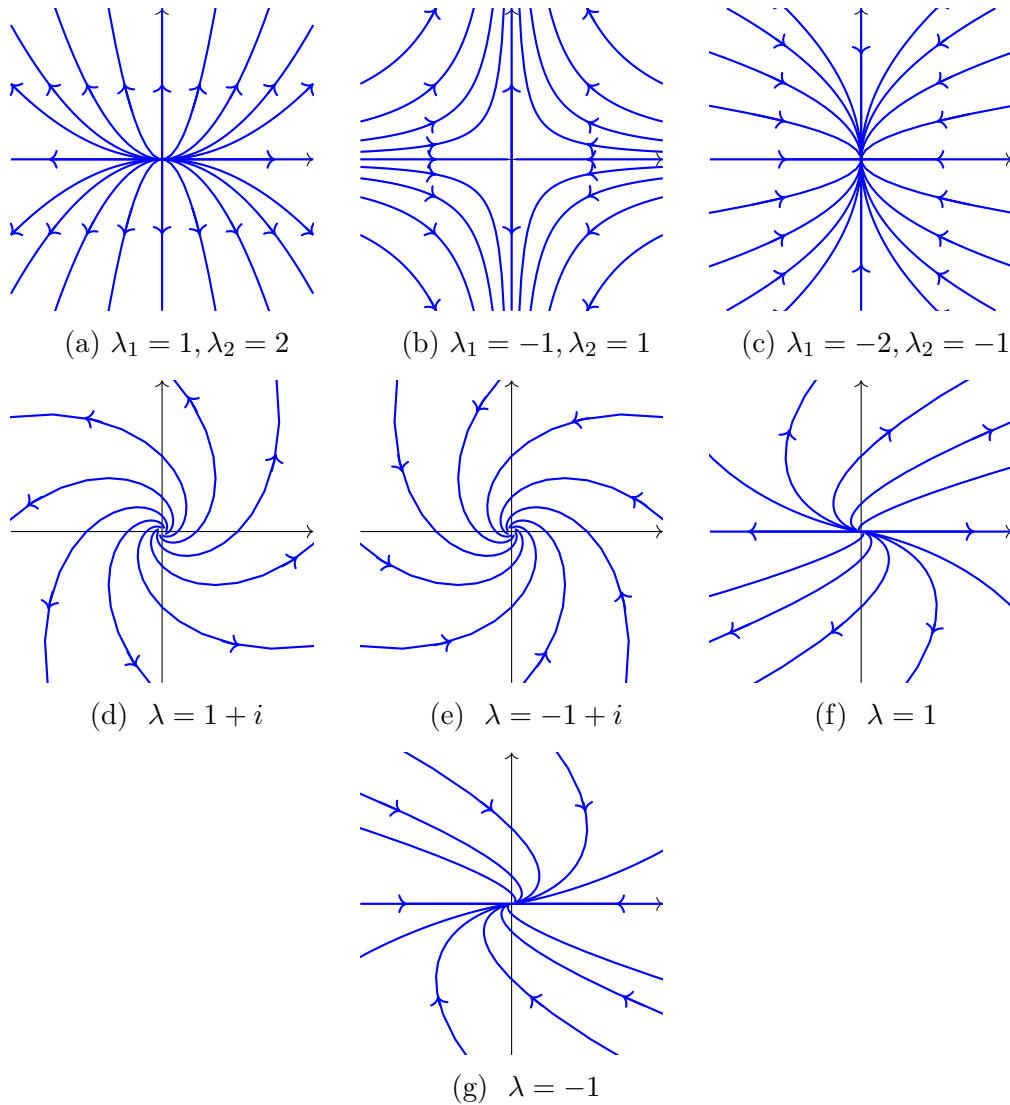


FIGURE 6.3 – Représentation du flot de l'équation (6.2) pour différentes matrices hyperboliques.

sont stables par le flot de l'équation différentielle. Si ces ensembles de sous-niveau forment une « base » de voisinages de 0⁵ (et ce sera le cas), l'équilibre est stable. En étudiant un peu plus précisément la vitesse de décroissance de la fonction de Lyapunov, on peut même montrer la stabilité asymptotique.

Nous allons construire cette fonction de Lyapunov comme une fonction quadratique, définie en fonction de $Jf(0)$. Pour cela, on trigonalise $Jf(0)$ sur \mathbb{C} : on fixe $G \in GL(n, \mathbb{C})$ telle que

$$Jf(0) = G(D + N)G^{-1},$$

avec D une matrice diagonale (dont les coefficients diagonaux sont les valeurs propres de $Jf(0)$) et N une matrice triangulaire supérieure stricte.

Soit $\mu = \max_{k=1, \dots, K} \operatorname{Re}(D_{k,k}) < 0$.

On va commencer par se ramener au cas où $\|N\| < \frac{|\mu|}{2}$. Définissons, pour ϵ assez petit (on précisera après à quel point ϵ doit être petit),

$$H = \begin{pmatrix} 1 & & & \\ & \epsilon^{-1} & & \\ & & \dots & \\ & & & \epsilon^{-n} \end{pmatrix}.$$

Alors

$$Jf(0) = GH(H^{-1}DH + H^{-1}NH)H^{-1}G^{-1} = GH(D + H^{-1}NH)(GH)^{-1}$$

et, pour tous $i, j \in \{1, \dots, n\}$,

$$(H^{-1}NH)_{ij} = \frac{H_{jj}}{H_{ii}} N_{ij},$$

de sorte que $(H^{-1}NH)_{ij} = 0$ si $i \geq j$ (c'est-à-dire que $H^{-1}NH$ est triangulaire supérieure stricte) et, si $i > j$,

$$|(H^{-1}NH)_{ij}| \leq \epsilon |N_{ij}|.$$

Ainsi, pour ϵ assez proche de 0, $H^{-1}NH$ peut être arbitrairement proche de 0. Quitte à remplacer G par GH et N par $H^{-1}NH$, on peut donc supposer que

$$\|N\| < \frac{|\mu|}{2}.$$

5. c'est-à-dire si tout voisinage de 0 contient un ensemble de sous-niveau

Notre fonction de Lyapunov sera $(u \in U \rightarrow \|G^{-1}u\|_2^2)$. Le long d'une trajectoire $(\phi_t(u_0))$, sa dérivée est, en un point t , $2\text{Re} \langle G^{-1}\phi_t(u_0), G^{-1}f(\phi_t(u_0)) \rangle$. On commence donc par majorer $\text{Re} \langle G^{-1}u, G^{-1}f(u) \rangle$, pour tout $u \in U$:

$$\begin{aligned}
& \text{Re} \left(\langle G^{-1}u, G^{-1}f(u) \rangle \right) \\
&= \text{Re} \left(\langle G^{-1}u, G^{-1}(f(0) + Jf(0)(u) + o(\|u\|)) \rangle \right) \\
&= \text{Re} \left(\langle G^{-1}u, G^{-1}G(D + N)G^{-1}u \rangle \right) + o(\|u\|^2) \\
&= \text{Re} \left(\langle G^{-1}u, (D + N)G^{-1}u \rangle \right) + o(\|u\|^2) \\
&= \sum_{k=1}^K \text{Re}(D_{k,k})|(G^{-1}u)_k|^2 + \text{Re} \left(\langle G^{-1}u, NG^{-1}u \rangle \right) + o(\|u\|^2) \\
&\leq \mu \|G^{-1}u\|_2^2 + \|N\| \|G^{-1}u\|_2^2 + o(\|u\|^2) \\
&\leq \frac{\mu}{2} \|G^{-1}u\|_2^2 + o(\|u\|^2) \\
&= \left(\frac{\mu}{2} + o(1) \right) \|G^{-1}u\|_2^2.
\end{aligned}$$

En conséquent, il existe $\eta > 0$ tel que, pour tout $u \in B(0, \eta)$,

$$\text{Re} \left(\langle G^{-1}u, G^{-1}f(u) \rangle \right) \leq \frac{\mu}{4} \|G^{-1}u\|_2^2. \quad (6.7)$$

(Rappelons que μ est négatif : les deux termes de l'inégalité sont donc négatifs.)

Nous pouvons maintenant démontrer la stabilité asymptotique. Commençons par la stabilité. Soit $V \subset U$ un voisinage quelconque de 0. Montrons qu'il existe $W \subset U$ un voisinage de 0 tel que, pour tout $u_1 \in W$, $\phi_t(u_1)$ est défini et appartient à V pour tout $t \in \mathbb{R}^+$.

Soit $W = \{u \in \mathbb{R}^n \text{ tq } \|G^{-1}u\|_2 < \zeta\}$, avec $\zeta > 0$ une constante choisie suffisamment petite pour que $W \subset V \cap B(0, \eta)$ (W est ce qu'on appelle un *ensemble de sous-niveau* de $(u \in U \rightarrow \|G^{-1}u\|_2^2)$). C'est un voisinage ouvert de 0. Soit $u_1 \in W$ quelconque. Alors, pour tout $t \geq 0$,

$$\begin{aligned}
\frac{d}{dt} \|G^{-1}\phi_t(u_1)\|_2^2 &= 2\text{Re} \left(\left\langle G^{-1}\phi_t(u_1), \frac{d}{dt} [G^{-1}\phi_t(u_1)] \right\rangle \right) \\
&= 2\text{Re} \left(\langle G^{-1}\phi_t(u_1), G^{-1}f(\phi_t(u_1)) \rangle \right).
\end{aligned}$$

D'après l'équation (6.7), pour tout $t \geq 0$ tel que $G^{-1}\phi_t(u_1) \in W$,

$$\frac{d}{dt} \|G^{-1}\phi_t(u_1)\|_2^2 \leq \frac{\mu}{2} \|G^{-1}\phi_t(u_1)\|_2^2 \leq 0 \quad (6.8)$$

(on dit que $(u \in U \rightarrow \|G^{-1}u\|_2^2)$ est une *fonction de Lyapunov* sur W).

Notons $t_0 \in \mathbb{R}^+ \cup \{+\infty\}$ le plus grand réel (éventuellement infini) tel que, pour tout $t \in [0; t_0[$, $\phi_t(u_1)$ est défini et appartient à W . Comme W est borné, $\phi_t(u_1)$ ne sort pas de tout compact au voisinage de t_0 . Donc, si $t_0 < +\infty$, $\phi_{t_0}(u_1)$ est défini (d'après le théorème des bouts). Ainsi qu'on vient de le voir, l'application $(t \rightarrow \|G^{-1}\phi_t(u_1)\|_2^2)$ est décroissante sur $]0; t_0[$. Elle est également continue : si $t_0 < +\infty$, on doit avoir

$$\|G^{-1}\phi_{t_0}(u_1)\|_2 \leq \|G^{-1}\phi_0(u_1)\|_2 < \zeta.$$

Donc $G^{-1}\phi_{t_0}(u_1) \in W$. Comme W est ouvert et comme les solutions maximales de (Autonome) sont définies sur des ouverts, il existe $t_1 > t_0$ tel que, pour tout $t \in [0; t_1[$, $\phi_t(u_1)$ est défini et appartient à W . C'est en contradiction avec la définition de t_0 . On voit donc qu'il serait absurde que $t_0 < +\infty$. Donc $t_0 = +\infty$ et, pour tout $t \in \mathbb{R}^+$, $\phi_t(u_1)$ est défini et appartient à W (ainsi qu'à V , puisque $W \subset V$). Cela achève de démontrer la stabilité.

La stabilité asymptotique suit des mêmes arguments. Définissons en effet W comme précédemment (pour n'importe quel voisinage $V \subset U$ de 0) et considérons à nouveau $u_1 \in W$ quelconque. D'après ce qu'on vient de voir, l'inégalité (6.8) est vraie pour tout $t \geq 0$. Donc, pour tout $t \geq 0$,

$$\frac{d}{dt} \ln (\|G^{-1}\phi_t(u_1)\|_2^2) \leq \frac{\mu}{2},$$

ce qui entraîne que, pour tout $t \geq 0$,

$$\|G^{-1}\phi_t(u_1)\|_2^2 \leq \|G^{-1}\phi_0(u_1)\|_2^2 e^{-\frac{\mu}{2}t}.$$

Donc $\|G^{-1}\phi_t(u_1)\|_2 \xrightarrow{t \rightarrow +\infty} 0$ et $\|\phi_t(u_1)\|_2 \xrightarrow{t \rightarrow +\infty} 0$. Cela conclut pour la stabilité asymptotique. □

6.4 L'exemple du pendule

Dans cette dernière section, on étudie le portrait de phase, les équilibres et la stabilité d'une équation différentielle particulière, qui modélise un pendule.

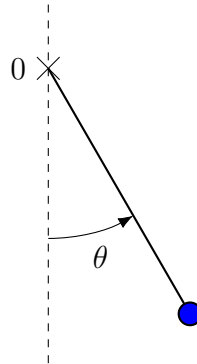


FIGURE 6.4 – Représentation schématique du pendule.

6.4.1 Justification de l'équation

On considère un pendule, c'est-à-dire une petite masse, qu'on imagine suspendue à un fil rigide. Le fil est fixé à un axe autour duquel il peut tourner vers la gauche ou la droite (pas vers l'avant ou l'arrière : le fil reste dans un plan). Pour tout $t \in \mathbb{R}$, on note $\theta(t)$ l'angle (positif ou négatif) entre le fil rigide et la verticale, à l'instant de temps t . Ce système est représenté sur la figure 6.4.

On imagine que le pendule n'est soumis à aucune autre force que la tension du fil (c'est la force qui assure que le pendule reste accroché au fil) et la gravité. C'est très simpliste : dans la réalité, il y aurait nécessairement des forces de frottement. Notons m la masse du pendule et R la longueur du fil. Si on prend pour origine le point de contact entre l'axe et le fil, les coordonnées du pendule dans son plan de mouvement sont, à tout instant $t \in \mathbb{R}$,

$$(R \sin(\theta(t)), -R \cos(\theta(t))).$$

La vitesse est la dérivée de la position,

$$v(t) \stackrel{\text{déf}}{=} (R\theta'(t) \cos(\theta(t)), R\theta'(t) \sin(\theta(t))), \quad \text{pour tout } t \in \mathbb{R},$$

et l'accélération est la dérivée de la vitesse,

$$a(t) \stackrel{\text{déf}}{=} (-R(\theta'(t))^2 \sin(\theta(t)) + R\theta''(t) \cos(\theta(t)), \\ R(\theta'(t))^2 \cos(\theta(t)) + R\theta''(t) \sin(\theta(t))), \quad \text{pour tout } t \in \mathbb{R}.$$

La force exercée par la gravité est représentée par le vecteur

$$(0, -mg),$$

où g est la constante de gravitation universelle. La force de tension n'a pas de formule explicite directe mais on sait qu'elle est exercée sur le pendule dans la direction du fil : pour tout t , il existe $k(t) \in \mathbb{R}$ tel que cette force est représentée par le vecteur

$$(-k(t) \sin(\theta(t)), k(t) \cos(\theta(t))).$$

La deuxième loi de Newton permet d'écrire, pour tout t ,

$$(0, -mg) + (-k(t) \sin(\theta(t)), k(t) \cos(\theta(t))) = ma(t).$$

Ainsi,

$$\begin{aligned} -k(t) \sin(\theta(t)) &= -R(\theta'(t))^2 \sin(\theta(t)) + R\theta''(t) \cos(\theta(t)); \\ -mg + k(t) \cos(\theta(t)) &= R(\theta'(t))^2 \cos(\theta(t)) + R\theta''(t) \sin(\theta(t)). \end{aligned}$$

En multipliant la première ligne par $\cos(\theta(t))$, la seconde par $\sin(\theta(t))$, puis en sommant, on obtient

$$-mg \sin(\theta(t)) = R\theta''(t).$$

Pour simplifier les notations, on suppose que $mg = R$, ce qui conduit à l'équation suivante :

$$\theta''(t) = -\sin(\theta(t)).$$

Ceci est une équation de deuxième ordre. Pour arriver à une équation de la forme (**Autonome**), on suit la remarque précédent le théorème de Cauchy-Lipschitz (4.1) : on introduit la fonction $u : t \in \mathbb{R} \rightarrow (\theta(t), \theta'(t)) \in \mathbb{R}^2$. Elle vérifie l'équation

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

avec $f : (u_1, u_2) \in \mathbb{R}^2 \rightarrow (u_2, -\sin(u_1))$.

On peut d'ores et déjà remarquer que les solutions maximales de (**Pendule**) sont définies sur \mathbb{R} , en vertu de la propriété énoncée à l'exemple 4.9.⁶

6. En effet, pour tout (u_1, u_2) , comme $|\sin(u_1)| \leq |u_1|$, on a que $\|f(u_1, u_2)\|_2 \leq \|(u_1, u_2)\|_2$.

6.4.2 Équilibres

Les zéros de f (et donc les équilibres du système (**Pendule**)) sont les points de \mathbb{R}^2 de la forme

$$(k\pi, 0)$$

pour tous les entiers $k \in \mathbb{Z}$. Lorsque k est pair, cela correspond à la position « tout en bas » du pendule ; lorsque k est impair, au contraire, cela correspond à la position « tout en haut ».

L'intuition physique nous dit que la position du bas (k pair) est stable (si le pendule est tout en bas et qu'on le bouge légèrement, il va revenir vers la position d'équilibre et pas s'en écarter) tandis que celle du haut (k impair) est instable (si le fil est vertical, avec le pendule au-dessus de l'axe, une petite perturbation va plutôt avoir pour effet de faire tomber le pendule que de le ramener à cette position d'équilibre).

Pour le démontrer, on peut tenter d'appliquer le théorème 6.11. Pour tout $k \in \mathbb{Z}$, la matrice jacobienne de f en $(k\pi, 0)$ est

$$Jf(k\pi, 0) = \begin{pmatrix} 0 & 1 \\ (-1)^{k+1} & 0 \end{pmatrix}.$$

On vérifie que les valeurs propres de cette matrice sont i et $-i$ si k est pair, 1 et -1 si k est impair. Comme $\operatorname{Re}(1) > 0$, on peut affirmer que l'équilibre $(k\pi, 0)$ est instable pour tout k impair.

En revanche, si k est pair, on ne peut rien déduire du théorème 6.11 : i et $-i$ sont de partie réelle nulle.

6.4.3 Intégrale première et portrait de phase

Les trajectoires de l'équation (**Pendule**) n'admettent pas d'expression explicite. Toutefois, on peut les étudier de manière relativement précise, et aussi étudier la stabilité des équilibres de la forme $(k\pi, 0)$ pour k pair, grâce à un outil bien pratique : une *intégrale première*. Il s'agit d'une fonction qui est constante le long des trajectoires du système, de sorte que les orbites sont incluses dans ses lignes de niveau.

Dans notre cas, l'intégrale première la plus naturelle est

$$F : (u_1, u_2) \in \mathbb{R}^2 \rightarrow -\cos(u_1) + \frac{u_2^2}{2}.$$

Il s'agit bien d'une intégrale première car, si u est une solution de l'équation (**Pendule**), alors, pour tout t ,

$$\begin{aligned}(F \circ u)'(t) &= u_1'(t) \sin(u_1(t)) + u_2'(t)u_2(t) \\ &= u_2(t) \sin(u_1(t)) - \sin(u_1(t))u_2(t) \\ &= 0,\end{aligned}$$

c'est-à-dire que $F \circ u$ est une fonction constante.

À quoi ressemblent les lignes de niveau de F ? Elles sont représentées sur la figure 6.5.

- Si $F_0 < -1$, $\{u, F(u) = F_0\} = \emptyset$, car $F(u_1, u_2) = -\cos(u_1) + \frac{u_2^2}{2} \geq -\cos(u_1) \geq -1$ pour tout $(u_1, u_2) \in \mathbb{R}^2$.
- Si $F_0 = -1$, $\{u, F(u) = F_0\} = \{(2k\pi, 0), k \in \mathbb{Z}\}$; l'ensemble de niveau est discret.
- Si $-1 < F_0 < 1$, $\{u, F(u) = F_0\}$ est une union de courbes fermées, identiques les unes aux autres à translation près par un multiple de $(2\pi, 0)$.
- Si $F_0 = 1$, $\{u, F(u) = F_0\}$ peut s'écrire comme l'union de deux courbes (régulières) qui s'intersectent, créant ainsi une singularité, en les points de la forme $(k\pi, 0)$ pour k impair.
- Si $F_0 > 1$, $\{u, F(u) = F_0\} = \{(u_1, u_2), u_2 = \pm\sqrt{2(F_0 + \cos(u_1))}\}$. Cet ensemble a deux composantes connexes, toutes deux non-bornées; l'une est incluse dans le demi-plan supérieur et l'autre dans le demi-plan inférieur.

Savoir que les trajectoires de l'équation (**Pendule**) sont incluses dans les lignes de niveau de F nous permet de démontrer le théorème suivant.

Théorème 6.12

Les solutions maximales constantes de l'équation (**Pendule**) sont les applications $(t \in \mathbb{R} \rightarrow (k\pi, 0))$ pour tout $k \in \mathbb{Z}$.

Soit $u : \mathbb{R} \rightarrow \mathbb{R}^2$ une solution maximale non-constante. Notons $F_0 = F(u(0))$.

- Si $F_0 < 1$, la fonction u est périodique. De plus, il existe $k \in \mathbb{Z}$ un entier tel que u_1 parcourt alternativement le segment $[2k\pi - \arccos(-F_0); 2k\pi + \arccos(-F_0)]$ dans le sens croissant et dans le sens décroissant.

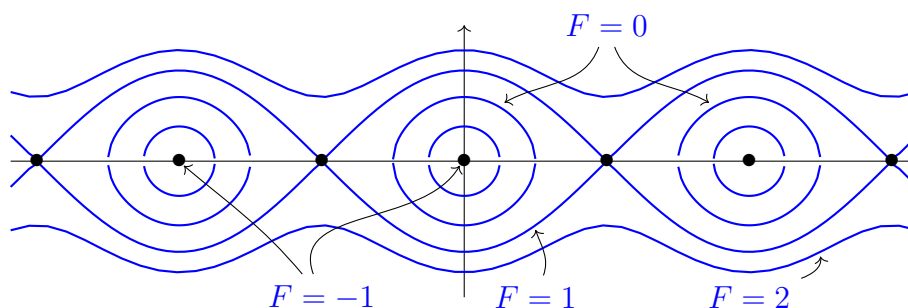


FIGURE 6.5 – Lignes de niveau de la fonction F ; les points noirs représentent les équilibres $(k\pi, 0)$, $k \in \mathbb{Z}$.

- Si $F_0 > 1$, la fonction u n'est pas périodique et u_1 diverge. Toutefois, il existe $T > 0$ tel que

$$u(t + T) = u(t) + (2\pi, 0), \quad \text{pour tout } t \in \mathbb{R}$$

ou bien

$$u(t + T) = u(t) - (2\pi, 0), \quad \text{pour tout } t \in \mathbb{R}.$$

- Si $F_0 = 1$, il existe $k \in \mathbb{Z}$ un entier tel que

$$u(t) \xrightarrow{t \rightarrow -\infty} ((2k - 1)\pi, 0) \text{ et } u(t) \xrightarrow{t \rightarrow +\infty} ((2k + 1)\pi, 0)$$

$$\text{ou bien } u(t) \xrightarrow{t \rightarrow -\infty} ((2k + 1)\pi, 0) \text{ et } u(t) \xrightarrow{t \rightarrow +\infty} ((2k - 1)\pi, 0).$$

Avant de démontrer une partie de ce théorème, discutons la signification physique des trajectoires décrites par le théorème. Le cas $F_0 < 1$ correspond à des mouvements d'oscillation périodiques autour de la position d'équilibre « du bas », entre les angles $-\arcsin(-F_0)$ et $\arcsin(-F_0)$. Le cas $F_0 > 1$ correspond à des mouvements de rotation autour de l'axe : partant (par exemple) du bas avec une vitesse assez élevée, le pendule atteint la position d'équilibre « du haut », la dépasse, redescend de l'autre côté et recommence.

Le cas $F_0 = 1$ est assez particulier. Il s'agit en effet des trajectoires « limites » entre les deux régimes décrits précédemment : si le pendule est lancé avec exactement la bonne impulsion il peut a priori tendre vers la position d'équilibre « du haut », à une vitesse qui tend vers 0 de telle manière que le

pendule n'atteint pas cette position du haut en temps fini mais se contente de converger vers elle. Ce sont des trajectoires qu'on n'observe jamais dans la réalité.

Démonstration partielle du théorème. L'affirmation concernant les solutions constantes est due au fait que les points $(k\pi, 0)$ pour $k \in \mathbb{Z}$ sont les seuls zéros de la fonction $f : (u_1, u_2) \rightarrow (u_2, -\sin(u_1))$.

Pour le reste, nous allons nous contenter de démontrer le premier point. Les autres découlent d'arguments un peu similaires. Supposons donc que $F_0 < 1$. En fait, $F_0 \in]-1; 1[$: F ne prend pas de valeurs inférieures à -1 et on ne peut pas avoir $F_0 = -1$, sinon u serait constante (les points atteignant la valeur -1 sont des équilibres).

Notons $u = (u_1, u_2)$. La fonction u_2 n'est pas constante (sinon, $\sin(u_1) = -u_2'$ doit être constante, de valeur nulle, donc u_1 est aussi constante, c'est-à-dire que u est constante). Il existe donc $t_0 \in \mathbb{R}$ tel que $u_2(t_0) \neq 0$. Fixons un tel point. Supposons par exemple que $u_2(t_0) > 0$ (on pourrait faire exactement le même raisonnement si $u_2(t_0) < 0$).

On remarque tout d'abord que u est bornée. En effet, pour tout t ,

$$-\cos(u_1(t)) + \frac{u_2(t)^2}{2} = F(u(t)) = F_0$$

donc $u_2(t)^2 \leq 2(F_0 + \cos(u_1(t))) \leq 2(F_0 + 1)$. De plus, $u_1(t)$ ne prend aucune valeur de la forme $k\pi$ avec $k \in \mathbb{Z}$ impair (car, pour tout t , $\cos(u_1(t)) = -F_0 + \frac{u_2(t)^2}{2} \geq -F_0 > -1$). D'après le théorème des valeurs intermédiaires, u_1 reste donc comprise dans un intervalle de la forme $]k\pi; (k+2)\pi[$ pour un certain $k \in \mathbb{Z}$ impair.

Montrons d'abord qu'il est impossible que $u_2(t) > 0$ pour tout $t \geq t_0$. Par l'absurde, on suppose que $u_2(t) > 0$ pour tout $t \geq t_0$. Comme $u_1'(t) = u_2(t)$ pour tout t , la fonction u_1 est croissante sur $[t_0; +\infty[$. Nous avons vu qu'elle était bornée. Elle admet donc une limite $\theta_{+\infty}$ en $+\infty$.

Si $\sin(\theta_{+\infty}) > 0$, alors $u_2'(t) = -\sin(u_1(t)) < -\frac{1}{2}\sin(\theta_{+\infty})$ pour tout t assez grand. En conséquence, $u_2 \rightarrow -\infty$ en $+\infty$, ce qui contredit le fait que u est bornée. De même, si $\sin(\theta_{+\infty}) < 0$, on aboutit à une contradiction.

On doit donc avoir $\sin(\theta_{+\infty}) = 0$, c'est-à-dire $\theta_{+\infty} = k\pi$ pour $k \in \mathbb{Z}$. Il est impossible que k soit impair (sinon, $\cos(u_1(t)) \xrightarrow{t \rightarrow +\infty} -1$, or on a déjà vu que $\cos(u_1(t)) \geq -F_0 > -1$ pour tout t). Donc k est pair et

$$\cos(u_1(t)) \xrightarrow{t \rightarrow +\infty} 1.$$

Pour tout $t \geq t_0$, $u_2(t) = \sqrt{2(F_0 + \cos(u_1(t)))}$; en conséquent,

$$u_2(t) \xrightarrow{t \rightarrow +\infty} \sqrt{2(F_0 + 1)}.$$

En particulier, $u_2(t) > \sqrt{F_0 + 1}$ pour tout t assez grand, ce qui implique que $u_1'(t) = u_2(t) > \sqrt{F_0 + 1}$, donc $u_1 \xrightarrow{+\infty} +\infty$, ce qui est à nouveau une contradiction.

On a donc bien montré qu'il est impossible que $u_2(t) > 0$ pour tout $t \geq t_0$. De même, on peut montrer qu'il est impossible que $u_2(t) > 0$ pour tout $t \leq t_0$.

Notons t_0^- le plus grand réel inférieur à t_0 tel que $u_2(t_0^-) = 0$ et t_0^+ le plus petit réel supérieur à t_0 tel que $u_2(t_0^+) = 0$. On doit avoir

$$-\cos(u_1(t_0^-)) = -\cos(u_1(t_0^+)) = F_0.$$

Cela signifie que $u_1(t_0^-)$ et $u_1(t_0^+)$ sont de la forme $2k\pi - \arcsin(-F_0)$ ou $2k\pi + \arcsin(-F_0)$, pour un certain $k \in \mathbb{Z}$ (entier k qui n'est pas nécessairement le même pour t_0^- et t_0^+).

Pour tout t , $\cos(u_1(t)) = \frac{u_2(t)^2}{2} - F_0 \geq -F_0$. Comme u_1 est strictement croissante à droite de t_0^- (puisque $u_1' = u_2$), on ne peut pas avoir $u_1(t_0^-) = 2k\pi + \arcsin(-F_0)$, pour un certain $k \in \mathbb{Z}$ (sinon, comme \cos est strictement décroissante au voisinage de $2k\pi + \arcsin(-F_0)$, on aurait $\cos(u_1(t)) < -F_0$ pour tout t légèrement supérieur à t_0^-). Il existe donc k_- tel que

$$u_1(t_0^-) = 2k_- \pi - \arcsin(-F_0).$$

Un raisonnement identique montre qu'il existe k_+ tel que

$$u_1(t_0^+) = 2k_+ \pi + \arcsin(-F_0).$$

On doit avoir $k_- \leq k_+$ car $u_1(t_0^-) < u_1(t_0^+)$. On ne peut pas avoir $k_- < k_+$ sinon, par le théorème des valeurs intermédiaires, il existerait t tel que $u_1(t) = (2k_+ - 1)\pi$ et alors $\cos(u_1(t)) = -1 < -F_0$. Donc $k_- = k_+$. Notons k cette valeur commune.

Pour conclure, on va montrer que, pour tout $t \in \mathbb{R}$,

$$\begin{aligned} u_1(t + t_0^+ - t_0^-) &= 4k\pi - u_1(t); \\ u_2(t + t_0^+ - t_0^-) &= -u_2(t). \end{aligned} \tag{6.9}$$

Supposons dans l'immédiat que ces relations sont vraies et déduisons-en le résultat demandé. Pour tout t , on doit alors avoir

$$u_1(t + 2(t_0^+ - t_0^-)) = 4k\pi - u_1(t + t_0^+ - t_0^-) = u_1(t);$$

$$u_2(t + 2(t_0^+ - t_0^-)) = -u_2(t + t_0^+ - t_0^-) = u_2(t).$$

Cela montre que u est $2(t_0^+ - t_0^-)$ -périodique.

De plus, nous avons vu que u_1 parcourait $[2k\pi - \arcsos(-F_0); 2k\pi + \arcsos(-F_0)]$ dans le sens croissant sur $[t_0^-; t_0^+]$. La relation $u_1(t + t_0^+ - t_0^-) = 4k\pi - u_1(t)$ montre qu'elle parcourt cet intervalle dans le sens décroissant sur $[t_0^+; 2t_0^+ - t_0^-]$. Ensuite, la $2(t_0^+ - t_0^-)$ -périodicité montre que u_1 parcourt à nouveau $[2k\pi - \arcsos(-F_0); 2k\pi + \arcsos(-F_0)]$ dans le sens croissant sur $[2t_0^+ - t_0^-; 3t_0^+ - 2t_0^-]$ et ainsi de suite.

Reste à montrer (6.9). Pour cela, on définit

$$\begin{aligned} v : \mathbb{R} &\rightarrow \mathbb{R}^2 \\ t &\rightarrow (4k\pi - u_1(t), -u_2(t)). \end{aligned}$$

L'application v est solution de l'équation (Pendule) : pour tout t ,

$$\begin{aligned} v_1'(t) &= -u_1'(t) = -u_2(t) = v_2(t) \\ v_2'(t) &= -u_2'(t) = \sin(u_1(t)) = -\sin(4k\pi - u_1(t)) = -\sin(v_1(t)). \end{aligned}$$

L'application $t \in \mathbb{R} \rightarrow u(t + (t_0^+ - t_0^-))$ est également solution (comme toutes les translations de u). Ces deux solutions sont maximales, puisqu'elles sont définies sur \mathbb{R} . Elles coïncident en t_0^- :

$$\begin{aligned} v(t_0^-) &= (4k\pi - u_1(t_0^-), -u_2(t_0^-)) \\ &= (4k\pi - (2k\pi - \arcsos(-F_0)), 0) \\ &= (2k\pi + \arcsos(-F_0), 0) \\ &= u(t_0^+) \\ &= u(t_0^- + (t_0^+ - t_0^-)). \end{aligned}$$

Par unicité de la solution maximale d'un problème de Cauchy dont la fonction f est localement uniformément lipschitzienne, on doit avoir $v(t) = u(t + t_0^+ - t_0^-)$ pour tout $t \in \mathbb{R}$, ce qui démontre l'équation (6.9). □

Le portrait de phase est dessiné sur la figure 6.6. Sur cette figure, on voit bien l'instabilité des points critiques $(k\pi, 0)$ pour $k \in \mathbb{Z}$ impair (certaines trajectoires s'en éloignent bien qu'elles soient parties d'extrêmement près). La figure permet également de conjecturer, conformément à l'intuition physique discutée plus tôt, que les points critiques $(k\pi, 0)$ pour $k \in \mathbb{Z}$ pair sont stables.

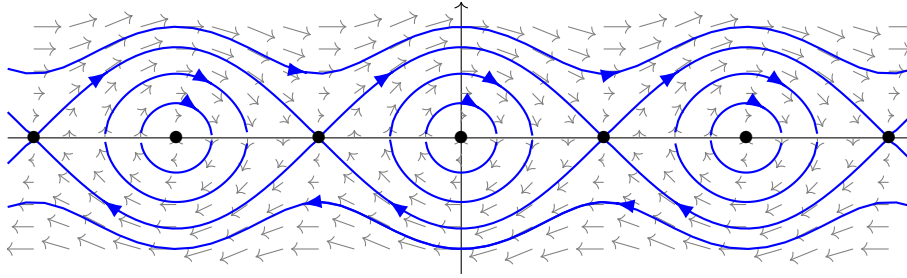


FIGURE 6.6 – Portrait de phase de l'équation (Pendule).

Théorème 6.13

Pour tout $k \in \mathbb{Z}$ pair, $(k\pi, 0)$ est un équilibre stable du système.

Démonstration. Effectuons la démonstration pour $k = 0$ (cela simplifie légèrement les notations mais ne change rien à l'argument).

Soit $V \subset \mathbb{R}^2$ un voisinage de $(0, 0)$. Soit $\eta > 0$ tel que $] -\eta; \eta[^2 \subset V$. Définissons le voisinage de 0

$$W = \left\{ u \in \mathbb{R}^2, F(u) < \min \left(-1 + \frac{\eta^2}{2}, -\cos(\eta) \right) \right\} \cap]-\eta; \eta[^2.$$

Pour toute solution de (Pendule) telle que $u(0) \in W$, on a $u(t) \in W \subset V$ pour tout $t \geq 0$ (et d'ailleurs même pour tout $t \in \mathbb{R}$).

En effet, comme $F \circ u$ est constante, on a pour tout $t \in \mathbb{R}$ $F(u(t)) = F(u(0)) < \min \left(-1 + \frac{\eta^2}{2}, -\cos(\eta) \right)$. De plus, il n'existe pas $t \in \mathbb{R}$ tel que $u_1(t) = \pm\eta$ ou $u_2(t) = \pm\eta$. En effet, si $u_1(t) = \pm\eta$,

$$F(u(t)) \geq -\cos(u_1(t)) = -\cos(\eta) \geq \min \left(-1 + \frac{\eta^2}{2}, -\cos(\eta) \right)$$

et si $u_2(t) = \pm\eta$,

$$F(u(t)) \geq -1 + \frac{u_2(t)^2}{2} = -1 + \frac{\eta^2}{2} \geq \min \left(-1 + \frac{\eta^2}{2}, -\cos(\eta) \right).$$

Dans les deux cas, c'est impossible. Puisque u est continue, on doit avoir $u(t) \in]-\eta; \eta[^2$ pour tout $t \in \mathbb{R}$, ce qui finit de montrer que $u(t) \in W$. \square

Annexe A

Connexité

Définition A.1 : connexité

Soit M un sous-ensemble de \mathbb{R}^n , pour un certain $n \in \mathbb{N}^*$. On dit que M est *connexe* s'il n'existe pas $U_1, U_2 \subset M$ vérifiant toutes les propriétés suivantes :

- U_1 et U_2 sont des ouverts de M^a ;
- $U_1 \cap U_2 = \emptyset$;
- $U_1 \cup U_2 = M$.

a. c'est-à-dire que $U_1 = M \cap V_1$ pour un certain ouvert V_1 de \mathbb{R}^n , et de même pour U_2

Proposition A.2 : définition alternative de la connexité

Soit M un sous-ensemble de \mathbb{R}^n , pour un certain $n \in \mathbb{N}^*$. L'ensemble M est connexe si et seulement si tous les sous-ensembles Ω de M qui sont simultanément ouverts et fermés dans M vérifient la propriété suivante :

$$\Omega = \emptyset \quad \text{ou} \quad \Omega = M.$$

Exemple A.3

Un intervalle de \mathbb{R} est toujours connexe.
L'union de deux intervalles ouverts disjoints non-vides n'est jamais connexe.

Définition A.4 : composante connexe

Soit M un sous-ensemble de \mathbb{R}^n , pour un certain $n \in \mathbb{N}^*$. Pour tout sous-ensemble A de M , on dit que A est une *composante connexe* de M s'il vérifie les deux propriétés suivantes :

- A est connexe ;
- A est un sous-ensemble connexe maximal de M , c'est-à-dire que pour tout $B \subset M$ connexe, si $A \subset B$, alors $A = B$.

Exemple A.5

Soient $(I_k)_{k \in E}$ une collection (finie ou infinie) d'intervalles ouverts non-vides de \mathbb{R} , disjoints deux à deux. Notons

$$M = \bigcup_{k \in E} I_k.$$

Les composantes connexes de M sont les I_k .

Proposition A.6

Soit M un sous-ensemble de \mathbb{R}^n , pour un certain $n \in \mathbb{N}^*$.
Les composantes connexes de M sont disjointes deux à deux.
De plus, M est l'union de ses composantes connexes.

Proposition A.7

Soient M_1, M_2 deux sous-ensembles, respectivement, de \mathbb{R}^{n_1} et \mathbb{R}^{n_2} .
Supposons qu'il existe

$$\phi : M_1 \rightarrow M_2$$

un homéomorphisme de M_1 vers M_2 .

Pour tout $A \subset M_1$,

- A est connexe si et seulement si $\phi(A)$ est connexe ;
- A et $\phi(A)$ ont le même nombre de composantes connexes.

Annexe B

Fonctions régulières à valeurs imposées

Lemme B.1

Pour tout segment $[a; b] \subset \mathbb{R}$, il existe une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ de classe C^∞ telle que

- $f(x) = 0$ pour tout $x \in]-\infty; a] \cup [b; +\infty[$;
- $f(x) > 0$ pour tout $x \in]a; b[$.

Démonstration. Soient a, b fixés, avec $a < b$. Définissons

$$g : \mathbb{R} \rightarrow \mathbb{R} \\ x \rightarrow \begin{cases} 0 & \text{si } x \leq 0, \\ e^{-\frac{1}{x}} & \text{si } x > 0. \end{cases}$$

C'est une fonction C^∞ , qui vérifie, pour tout $x \in \mathbb{R}$,

$$g(x) = 0 \text{ si } x \leq 0, \\ g(x) > 0 \text{ si } x > 0.$$

Définissons, pour tout $x \in \mathbb{R}$,

$$f(x) = g(x - a)g(b - x).$$

Cette fonction est C^∞ . De plus,

- pour tout $x \in]-\infty; a]$, $x - a \leq 0$ donc $g(x - a) = 0$ et $f(x) = 0$;

- pour tout $x \in [b; +\infty[$, $b - x \leq 0$ donc $g(b - x) = 0$ et $f(x) = 0$;
- pour tout $x \in]a; b[$, $x - a > 0$ et $b - x > 0$ donc $g(x - a) > 0$ et $g(b - x) > 0$, d'où $f(x) > 0$.

□

Corollaire B.2

Pour tout segment $[a; b] \subset \mathbb{R}$, il existe une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ de classe C^∞ telle que

- $f(x) = 0$ pour tout $x \in]-\infty; a]$;
- $f(x) \in [0; 1]$ pour tout $x \in]a; b[$;
- $f(x) = 1$ pour tout $x \in [b; +\infty[$.

Démonstration. Soit $F : \mathbb{R} \rightarrow \mathbb{R}$ comme dans le lemme précédent. Posons

$$f : t \in \mathbb{R} \rightarrow \frac{\int_{-\infty}^t F(s) ds}{\int_{-\infty}^b F(s) ds}.$$

C'est bien une fonction C^∞ . Elle est nulle sur $] - \infty; a]$ car F l'est. Elle est constante sur $[b; +\infty[$ car F est nulle sur $[b; +\infty[$; sa valeur est $f(b) = 1$. De plus, comme F est positive, f est croissante; on a donc $f(x) \in [0; 1]$ pour tout $x \in]a; b[$. □

Proposition B.3

Soient $c_1 < c_2 < \dots < c_S$ et $d_1 < d_2 < \dots < d_S$ des réels arbitraires, pour un $S \geq 2$. Il existe un C^∞ -difféomorphisme $\psi : [c_1; c_S] \rightarrow [d_1; d_S]$ tel que, pour tout $k = 1, \dots, S$,

$$\psi(c_k) = d_k.$$

Le même résultat est vrai si $d_1 > d_2 > \dots > d_S$.

Démonstration. On va définir ψ comme primitive d'une fonction $p : [c_1; c_S] \rightarrow [d_1; d_S]$ bien choisie.

Pour tout $k = 1, \dots, S - 1$, soit $f_k : \mathbb{R} \rightarrow \mathbb{R}$ une fonction de classe C^∞ qui vaut 0 sur $] - \infty; c_k] \cup [c_{k+1}; +\infty[$ et qui est strictement positive sur $]c_k; c_{k+1}[$.

Le lemme B.1 nous assure qu'il en existe. Quitte à la multiplier par une constante strictement positive bien choisie, on peut supposer que

$$\int_{c_k}^{c_{k+1}} f_k(t) dt = 1.$$

Fixons un réel $\epsilon > 0$ tel que, pour tout $k = 1, \dots, S-1$,

$$\epsilon < \frac{d_{k+1} - d_k}{c_{k+1} - c_k}.$$

Définissons maintenant

$$p = \epsilon + \sum_{k=1}^{S-1} (d_{k+1} - d_k - \epsilon(c_{k+1} - c_k)) f_k,$$

$$\psi : x \in [c_1; c_S] \rightarrow d_1 + \int_{d_1}^x p(t) dt.$$

Ces deux fonctions sont C^∞ . On observe que, pour tout $k = 1, \dots, S-1$,

$$\begin{aligned} \psi(c_{k+1}) - \psi(c_k) &= \int_{c_k}^{c_{k+1}} [\epsilon + (d_{k+1} - d_k - \epsilon(c_{k+1} - c_k)) f_k(t)] dt \\ &= d_{k+1} - d_k. \end{aligned}$$

Cela permet de montrer par récurrence que, pour tout $k = 1, \dots, S$,

$$\psi(c_k) = d_k.$$

De plus, la fonction ψ est strictement croissante (sa dérivée, p , est toujours supérieure à ϵ). Comme elle est continue, elle réalise un homéomorphisme de $[c_1; c_S]$ vers $[\psi(c_1); \psi(c_S)] = [d_1; d_S]$. De plus, comme sa dérivée ne s'annule pas, sa réciproque est de classe C^∞ . C'est donc un C^∞ -difféomorphisme, comme on le souhaitait.

Montrons enfin que le résultat reste vrai si on n'a pas $d_1 < d_2 < \dots < d_S$ mais $d_1 > d_2 > \dots > d_S$. Le résultat qu'on vient de démontrer assure l'existence d'un C^∞ -difféomorphisme $\psi : [c_1; c_S] \rightarrow [-d_1; -d_S]$ tel que, pour tout $k = 1, \dots, S$,

$$\psi(c_k) = -d_k.$$

Alors $-\psi$ est un C^∞ -difféomorphisme de $[c_1; c_S]$ sur $[d_1; d_S]$ vérifiant les égalités voulues.

□

Proposition B.4

Soient $[a_1; a_2]$ et $[b_1; b_2]$ deux segments de \mathbb{R} non réduits à un point. Pour tout $k \in \mathbb{N}$, pour tous réels $\gamma_1^{(1)}, \dots, \gamma_1^{(k)}, \gamma_2^{(1)}, \dots, \gamma_2^{(k)}$ tels que

$$\gamma_1^{(1)} > 0 \quad \text{et} \quad \gamma_2^{(1)} > 0,$$

il existe un C^∞ -difféomorphisme ψ de $[a_1; a_2]$ vers $[b_1; b_2]$ tel que, pour tout $k' = 1, \dots, k$,

$$\psi^{(k')}(a_1) = \gamma_1^{(k')} \quad \text{et} \quad \psi^{(k')}(a_2) = \gamma_2^{(k')}.$$

Démonstration. On va définir ψ comme primitive d'une fonction p bien choisie.

Soient d'abord $q_1, q_2 : \mathbb{R} \rightarrow \mathbb{R}$ des fonctions de classe C^∞ telles que, pour tout $k' = 0, \dots, k-1$,

$$q_1^{(k')}(a_1) = \gamma_1^{(k'+1)} \quad \text{et} \quad q_2^{(k')}(a_2) = \gamma_2^{(k'+1)}.$$

(De telles fonctions existent. On peut définir $q_1 : x \rightarrow \sum_{k'=0}^{k-1} \frac{\gamma_1^{(k'+1)}}{k'!} (x - a_1)^{k'}$ et $q_2 : x \rightarrow \sum_{k'=0}^{k-1} \frac{\gamma_2^{(k'+1)}}{k'!} (x - a_2)^{k'}$, par exemple.)

Soit $\epsilon > 0$ tel que

$$q_1 > 0 \text{ sur } [a_1; a_1 + \epsilon] \text{ et } q_2 > 0 \text{ sur } [a_2 - \epsilon; a_2].$$

Un tel ϵ existe car $q_1(a_1) = \gamma_1^{(1)} > 0$ et $q_2(a_2) = \gamma_2^{(1)} > 0$ et q_1, q_2 sont continues. Quitte à réduire un peu ϵ , on peut de plus supposer que

$$\int_{a_1}^{a_1 + \epsilon} q_1(s) ds < \frac{b_2 - b_1}{2},$$

$$\int_{b_1 - \epsilon}^{b_1} q_2(s) ds < \frac{b_2 - b_1}{2}.$$

Soient $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$ des fonctions C^∞ (qui existent, d'après le corollaire B.2) telles que

- f_1 vaut 1 sur $] -\infty; a_1 + \frac{\epsilon}{2}]$, 0 sur $[a_1 + \epsilon; +\infty[$ et a des valeurs dans $[0; 1]$ sur $] a_1 + \frac{\epsilon}{2}; a_1 + \epsilon [$;

— f_2 vaut 0 sur $] -\infty; a_2 - \epsilon]$, 1 sur $[a_2 - \frac{\epsilon}{2}; +\infty[$ et a des valeurs dans $[0; 1]$ sur $]a_2 - \epsilon; a_2 - \frac{\epsilon}{2}[$.

Soit enfin $g : \mathbb{R} \rightarrow \mathbb{R}$ une fonction C^∞ qui vaut 0 sur $] -\infty; a_1 + \frac{\epsilon}{2}] \cup [a_2 - \frac{\epsilon}{2}; +\infty[$ et est strictement positive sur $]a_1 + \frac{\epsilon}{2}; a_2 - \frac{\epsilon}{2}[$ (conformément au lemme B.1). Quitte à multiplier g par une constante bien choisie, on peut supposer que

$$\int_{-\infty}^{+\infty} g(s) ds = 1.$$

On pose

$$p = f_1 q_1 + f_2 q_2 + \alpha g,$$

avec

$$\begin{aligned} \alpha &= b_2 - b_1 - \int_{a_1}^{a_2} (f_1 q_1 + f_2 q_2) \\ &= b_2 - b_1 - \int_{a_1}^{a_1 + \epsilon} f_1(s) q_1(s) ds - \int_{a_2 - \epsilon}^{a_2} f_2(s) q_2(s) ds \\ &\geq b_2 - b_1 - \int_{a_1}^{a_1 + \epsilon} q_1(s) ds - \int_{a_2 - \epsilon}^{a_2} q_2(s) ds \\ &> 0. \end{aligned}$$

La fonction p est C^∞ . Elle coïncide avec q_1 sur $] -\infty; a_1 + \frac{\epsilon}{2}]$ (car f_1 vaut 1 sur cet intervalle, tandis que f_2 et g valent 0). En particulier,

$$p^{(k')}(a_1) = \gamma_1^{(k'+1)} \quad \forall k' = 0, \dots, k-1. \quad (\text{B.1})$$

De même,

$$p^{(k')}(a_2) = \gamma_2^{(k'+1)} \quad \forall k' = 0, \dots, k-1. \quad (\text{B.2})$$

De plus, p est strictement positive sur $[a_1; a_2]$: sur $[a_1; a_1 + \frac{\epsilon}{2}]$, $f_1 q_1$ est strictement positive tandis que $f_2 q_2$ et g sont nulles, donc leur somme est strictement positive ; sur $]a_1 + \frac{\epsilon}{2}; a_1 + \epsilon]$, $f_1 q_1$ est positive, g est strictement positive et $f_2 q_2$ est nulle, donc leur somme est strictement positive, et ainsi de suite.

Enfin, d'après la définition de α ,

$$\int_{a_1}^{a_2} p(s) ds = b_2 - b_1.$$

Posons alors

$$\psi(x) = b_1 + \int_{a_1}^x p(s) ds.$$

C'est une application C^∞ , telle que $\psi(a_1) = b_1$ et $\psi(a_2) = b_2$. Sa dérivée est strictement croissante; elle réalise donc un C^∞ -difféomorphisme de $[a_1; a_2]$ vers $[b_1; b_2]$. En outre, elle vérifie les égalités

$$\psi^{(k')}(a_1) = \gamma_1^{(k')} \quad \text{et} \quad \psi^{(k')}(a_2) = \gamma_2^{(k')}$$

car sa dérivée, p , vérifie les équations (B.1) et (B.2).

□

Annexe C

Démonstrations pour la section 3.2

C.1 Démonstration de la proposition 3.17

Soit M une sous-variété connexe de \mathbb{R}^n , de classe C^k (pour un certain entier $k \in \mathbb{N}^*$). Nous allons montrer que deux points x_1, x_2 de M sont nécessairement reliés par un chemin (c'est-à-dire par une fonction continue et C^1 par morceaux, suivant la définition donnée au début de la section 3.2).

Soient $x_1, x_2 \in M$ fixés. Définissons

$$\mathcal{A} = \{y \in M, \text{ il existe un chemin reliant } x_1 \text{ et } y\}.$$

C'est un ensemble non-vidé : comme x_1 est relié à lui-même par les chemins constants de valeur x_1 , x_1 appartient à \mathcal{A} .

Montrons que \mathcal{A} est ouvert dans M . Soit $y \in \mathcal{A}$ quelconque. Soit $\gamma : [0; A] \rightarrow M$ un chemin reliant x_1 et y .

Soient U un voisinage de y dans \mathbb{R}^n , V un ouvert de \mathbb{R}^d et $f : V \rightarrow \mathbb{R}^n$ une application de classe C^k , réalisant un homéomorphisme sur son image, telle que

$$M \cap U = f(V).$$

(Il en existe, d'après la définition « immersion » des sous-variétés.) Notons a l'antécédant de y par f . Quitte à restreindre un peu V , on peut supposer que $V = B(a, \epsilon)$ pour un certain $\epsilon > 0$.

Montrons que $M \cap U \subset \mathcal{A}$. Soit $y' \in M \cap U$ quelconque. Notons $a' \in V$ son antécédant par f . Définissons

$$\begin{aligned} \tilde{\gamma} : [0; A+1] &\rightarrow M \\ t &\rightarrow \begin{cases} \gamma(t) & \text{si } t \in [0; A] \\ f((A+1-t)a + (t-A)a') & \text{si } t \in [A; A+1]. \end{cases} \end{aligned}$$

Cette application est bien définie : comme $a' \in V = B(a, \epsilon)$, le segment reliant a à a' est inclus dans $B(a, \epsilon)$, ce qui entraîne que $(A+1-t)a + (t-A)a' \in B(a, \epsilon)$ pour tout t . Elle est à valeurs dans M et C^1 par morceaux. Elle est de plus continue : elle l'est sur $[0; A]$ et $[A; A+1]$. De plus, elle admet en A une limite à gauche,

$$\gamma(A) = y$$

et une limite à droite,

$$f(a) = y,$$

qui coïncident. Elle est donc continue en A . En conclusion, c'est un chemin entre $\tilde{\gamma}(0) = x_1$ et $\tilde{\gamma}(A+1) = f(a') = y'$. Donc $y' \in \mathcal{A}$.

Ainsi, l'ensemble \mathcal{A} contient $M \cap U$, qui est un voisinage de y . Cela montre bien que \mathcal{A} est ouvert dans M .

Montrons maintenant que \mathcal{A} est un fermé de M , avec des arguments assez similaires. Soit $y \in M$ appartenant à l'adhérence de \mathcal{A} (c'est-à-dire limite d'une suite de points de \mathcal{A}). Montrons que $y \in \mathcal{A}$.

Définissons U, V, f comme dans la partie précédente de la démonstration. Soit à nouveau $a \in V$ l'antécédant de y par f et supposons que $V = B(a, \epsilon)$.

Puisque $M \cap U$ est un voisinage de y et puisque y est adhérent à \mathcal{A} , il existe un élément y' de $M \cap U$ qui appartient à \mathcal{A} . Fixons-le pour la suite, notons $a' \in V$ son antécédant par f et $\gamma : [0; A] \rightarrow M$ un chemin entre x_1 et y' . On définit

$$\begin{aligned} \tilde{\gamma} : [0; A+1] &\rightarrow M \\ t &\rightarrow \begin{cases} \gamma(t) & \text{si } t \in [0; A] \\ f((A+1-t)a' + (t-A)a) & \text{si } t \in [A; A+1]. \end{cases} \end{aligned}$$

De même que précédemment, on vérifie que $\tilde{\gamma}$ est un chemin, qui relie, cette fois-ci, x_1 et y . Donc $y \in \mathcal{A}$.

Ainsi, on a montré que \mathcal{A} était ouvert et fermé dans M , et non-vidé. Puisque M est connexe, on doit avoir $\mathcal{A} = M$. En particulier, \mathcal{A} contient x_2 : il existe un chemin entre x_1 et x_2 .

C.2 Démonstration de la proposition 3.21

Pour démontrer la proposition 3.21, on a recours au théorème d'Arzela-Ascoli qui donne une sorte de critère de compacité pour les parties de l'ensemble des fonctions continues entre deux espaces métriques. Ce théorème utilise la notion d'équicontinuité, dont la définition suit.

Définition C.1 : équicontinuité

Soient $(X, d_X), (Y, d_Y)$ deux espaces métriques. On note $\mathcal{C}(X, Y)$ l'ensemble des fonctions continues de X dans Y . Soit $\mathcal{A} \subset \mathcal{C}(X, Y)$. On dit que \mathcal{A} est *équicontinue* si, pour tous $x_0 \in X$ et $\epsilon > 0$, il existe $\eta > 0$ tel que

$$\forall x \in X, \quad (d_X(x_0, x) < \eta) \Rightarrow (\forall f \in \mathcal{A}, d_Y(f(x_0), f(x)) < \epsilon).$$

Proposition C.2

Soit $c > 0$.

Soient $(X, d_X), (Y, d_Y)$ deux espaces métriques. Une famille $\mathcal{A} \subset \mathcal{C}(X, Y)$ ne contenant que des fonctions c -lipschitziennes est équicontinue.

Démonstration. Pour tout $x_0 \in X$ et tout $\epsilon > 0$, si on pose $\eta = \epsilon/c$, alors, pour tout $x \in X$,

$$(d_X(x_0, x) < \eta) \Rightarrow (\forall f \in \mathcal{A}, d_Y(f(x_0), f(x)) \leq cd_X(x_0, x) < c\eta = \epsilon).$$

□

Théorème C.3 : Arzela-Ascoli [Paulin, 2009, Thm 5.31]

Soient $(X, d_X), (Y, d_Y)$ deux espaces métriques. Soit $\mathcal{A} \subset \mathcal{C}(X, Y)$. Supposons que

- \mathcal{A} est équicontinue ;
- pour tout $x \in X$, $\overline{\{f(x), f \in \mathcal{A}\}}$ est un compact de Y .

Alors toute suite $(f_n)_{n \in \mathbb{N}}$ d'éléments de \mathcal{A} admet une sous-suite qui converge uniformément sur tout compact vers une certaine fonction $g \in \mathcal{C}(X, Y)$.

Démonstration de la proposition 3.21. On utilise le théorème d'Arzela-Ascoli avec $X = [0; D/c]$ et $Y = \mathbb{R}^n$ (munis des distances usuelles). Notons \mathcal{A} l'ensemble des fonctions γ_n , dont on restreint l'intervalle de définition à $[0; D/c]$.¹ Il s'agit d'un ensemble de fonctions c -lipschitziennes (car $\|\gamma'_n(t)\|_2 = c$ pour tous n et t), donc équicontinu.

Pour tous $t \in [0; D/c]$ et $n \in \mathbb{N}$,

$$\|\gamma_n(t) - x_1\|_2 = \|\gamma_n(t) - \gamma_n(0)\|_2 \leq ct$$

donc, pour tout t , $\overline{\{f(t), f \in \mathcal{A}\}} = \overline{\{\gamma_n(t), n \in \mathbb{N}\}} \subset \overline{B}(x_1, ct)$. Ainsi, cet ensemble est fermé et borné dans \mathbb{R}^n : il est compact.

Les hypothèses du théorème d'Arzela-Ascoli sont donc vérifiées. On en déduit qu'il existe $\delta \in C^0([0; D/c], \mathbb{R}^n)$ et $\rho : \mathbb{N} \rightarrow \mathbb{N}$ une extraction telles que

$$\gamma_{\rho(n)} \xrightarrow{n \rightarrow +\infty} \delta \text{ uniformément sur tout compact.}$$

Comme $[0; D/c]$ est compact, la convergence « uniforme sur tout compacte » est simplement la convergence uniforme. On a donc bien

$$\|\gamma_{\rho(n)} - \delta\|_\infty \xrightarrow{n \rightarrow +\infty} 0.$$

Pour tout $t \in [0; D/c]$, $(\gamma_{\rho(n)}(t))_{n \in \mathbb{N}}$ est une suite d'éléments de M . Comme M est (par hypothèse) fermée dans \mathbb{R}^n , sa limite $\delta(t)$ appartient également à M . La fonction δ est donc bien à images dans M .

Elle est c -lipschitzienne : pour tous t_1, t_2 ,

$$\begin{aligned} \|\delta(t_1) - \delta(t_2)\|_2 &= \lim_{n \rightarrow +\infty} \|\gamma_{\rho(n)}(t_1) - \gamma_{\rho(n)}(t_2)\|_2 \\ &\leq \lim_{n \rightarrow +\infty} c \|t_1 - t_2\|_2 \\ &= c \|t_1 - t_2\|_2. \end{aligned}$$

Calculons enfin les valeurs de δ en 0 et D/c . En 0,

$$\delta(0) = \lim_{n \rightarrow +\infty} \gamma_{\rho(n)}(0) = \lim_{n \rightarrow +\infty} x_1 = x_1.$$

D'autre part, pour tout n ,

$$\|\delta(D/c) - x_2\|_2 \leq \|\delta(D/c) - \gamma_{\rho(n)}(D/c)\|_2$$

1. Pour tout n , γ_n est définie sur $[0; \ell(\gamma_n)/c]$ et $\ell(\gamma_n)/c \geq D/c$.

$$\begin{aligned}
& + \|\gamma_{\rho(n)}(D/c) - x_2\|_2 \\
= & \|\delta(D/c) - \gamma_{\rho(n)}(D/c)\|_2 \\
& + \|\gamma_{\rho(n)}(D/c) - \gamma_{\rho(n)}(\ell(\gamma_{\rho(n)})/c)\|_2 \\
\leq & \|\delta(D/c) - \gamma_{\rho(n)}(D/c)\|_2 \\
& + \|D - \ell(\gamma_{\rho(n)})\|_2.
\end{aligned}$$

Par passage à la limite de l'inégalité, on doit avoir $\|\delta(D/c) - x_2\|_2 = 0$, c'est-à-dire

$$\delta(D/c) = x_2.$$

□

C.3 Démonstration de la proposition 3.29

Pour tout $t \in [0; A]$,

$$\begin{aligned}
\text{dist}_M(f(0), f(t)) & \leq \ell(f|_{[0;t]}) \\
& = \int_0^t \|f'(s)\|_2 ds \\
& = \int_0^t (\|f'(0)\|_2 + o(1)) ds \\
& = t\|f'(0)\|_2 + o(t).
\end{aligned}$$

D'autre part, on a vu en démontrant la proposition 3.18 que

$$\begin{aligned}
\text{dist}_M(f(0), f(t)) & \geq \|f(0) - f(t)\|_2 \\
& = t\|f'(0)\|_2 + o(t).
\end{aligned}$$

Par encadrement, on obtient le résultat.

Annexe D

Compléments pour le chapitre 4

D.1 Lemme de Gronwall

Lemme D.1 : Gronwall

Soient $t_0 \leq T \in \mathbb{R}$, $a, c, u \in C^0([t_0; T], \mathbb{R})$. Supposons que $a \geq 0$ et, pour tout $t \in [t_0; T]$,

$$u(t) \leq c(t) + \int_{t_0}^t a(s)u(s)ds.$$

Alors, pour tout $t \in [t_0; T]$,

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

Le lemme est vrai aussi si $T < t_0$, à condition de remplacer le segment « $[t_0; T]$ » par « $[T; t_0]$ » et d'échanger les bornes dans chaque intégrale.

Démonstration. Cette démonstration a été vue en TD. □

D.2 Démonstration du lemme 4.11

Supposons que le théorème est vrai pour toutes les fonctions indépendantes de t et démontrons-le pour une fonction $f : (t, u) \in I \times U \rightarrow f(t, u) \in \mathbb{R}^n$ générale. Le principe est, à l'aide d'un changement de variable, d'écrire

les solutions maximales du problème (Cauchy u_0) comme les solutions maximales d'un autre problème, défini par une fonction indépendante de t , auquel on peut donc appliquer le théorème.

Pour tous $t_1 \in I, u_0 \in U$, définissons $\tilde{u}_{(t_1, u_0)} : J_{(t_1, u_0)} \rightarrow I \times U$ la solution maximale du problème

$$\begin{aligned} \tilde{u}'_{(t_1, u_0)} &= g(\tilde{u}_{(t_1, u_0)}), \\ \tilde{u}_{(t_1, u_0)}(t_0) &= (t_1, u_0), \end{aligned} \tag{D.1}$$

où $g : I \times U \rightarrow \mathbb{R}^{n+1}$ est la fonction telle que $g(x) = (1, f(t, v))$ pour tout $x = (t, v) \in I \times U$.

Pour tout u_0 , observons que $\tilde{u}_{(t_0, u_0)}$ est l'application

$$\begin{aligned} J_{u_0} &\rightarrow I \times U \\ t &\rightarrow (t, u_{u_0}(t)). \end{aligned} \tag{D.2}$$

En effet, on vérifie qu'elle est solution du problème (D.1). Il s'agit de plus de la solution maximale. En effet, soient $T_{u_0} : J_{(t_0, u_0)} \rightarrow I$ et $u_{u_0}^{pro} : J_{(t_0, u_0)} \rightarrow U$ telles que, pour tout $t \in J_{(t_0, u_0)}$,

$$\tilde{u}_{(t_0, u_0)}(t) = (T_{u_0}(t), u_{u_0}^{pro}(t)).$$

La définition du problème (D.1) entraîne que T_{u_0} est de dérivée constante égale à 1 ; comme cette fonction vaut t_0 en t_0 , on a, pour tout t , $T_{u_0}(t) = t$ et

$$u_{u_0}^{pro}(t) = f(T_{u_0}(t), u_{u_0}^{pro}(t)) = f(t, u_{u_0}^{pro}(t)).$$

Donc $u_{u_0}^{pro}$ est solution du même problème de Cauchy que u_{u_0} . Puisque u_{u_0} est solution maximale, $J_{(t_0, u_0)} \subset J_{u_0}$. Donc l'application définie à l'équation (D.2) est bien solution maximale du problème (D.1).

La fonction g du problème (D.1) n'a qu'un seul argument. Le théorème est donc vrai pour ce problème. L'ensemble $\tilde{\Omega} \stackrel{\text{déf}}{=} \{((t_1, u_0), t), t_1 \in I, u_0 \in U, t \in J_{(t_1, u_0)}\}$ est donc ouvert. L'application

$$\begin{aligned} W : \quad \tilde{\Omega} &\rightarrow I \times U \\ ((t_1, u_0), t) &\rightarrow \tilde{u}_{(t_1, u_0)}(t) \end{aligned}$$

est donc C^1 .

Or Ω est égal à $\{(u_0, t) \text{ tq } ((t_0, u_0), t) \in \tilde{\Omega}\}$. C'est donc l'antécédant de $\tilde{\Omega}$ par une application continue : c'est un ouvert. De plus, pour tout $(u_0, t) \in \Omega$,

$$V(u_0, t) = u_{u_0}(t) = [\tilde{u}_{(t_0, u_0)}(t)]_{2:n+1} = [W((t_0, u_0), t)]_{2:n+1},$$

où la notation « $2 : n + 1$ » désigne le vecteur constitué des deuxième, troisième, ..., $(n + 1)$ -ème coordonnées d'un élément de \mathbb{R}^{n+1} . Ainsi, le caractère C^1 de V est une conséquence du fait que W est C^1 .

Enfin, lorsqu'on sait que V est C^1 , on obtient le problème de Cauchy (Cauchy $\frac{dV}{du_0}$) en dérivant le problème de Cauchy (Cauchy u_0).

D.3 Démonstration du lemme 4.12

Supposons que la propriété (4.7) est vraie. Fixons u_0 et montrons que, pour tout $t \in J_{u_0}$,

$$\Omega \text{ contient un voisinage de } (u_0, t) \text{ sur lequel } V \text{ est } C^1 \text{ et vérifie} \quad (\text{D.3}) \\ \text{les équations (Cauchy } \frac{dV}{du_0}\text{).}$$

D'après l'hypothèse (4.7), le point t_0 satisfait la propriété (D.3). Soit J'_{u_0} l'ensemble des points de J_{u_0} satisfaisant la propriété (D.3). Il faut montrer que $J'_{u_0} = J_{u_0}$.

L'ensemble J'_{u_0} est non-vidé (il contient t_0) et ouvert : si t vérifie la propriété (D.3), notons H un voisinage de (u_0, t) comme dans la propriété. Pour tout t' assez proche de t , H est également un voisinage de (u_0, t') sur lequel V est C^1 et vérifie les équations (Cauchy $\frac{dV}{du_0}$). Donc $t' \in J'_{u_0}$.

Montrons que J'_{u_0} est fermé dans J_{u_0} . Comme J_{u_0} est connexe (c'est un intervalle), cela terminera la démonstration. Soit $t \in J_{u_0}$ un élément de l'adhérence de J'_{u_0} . Montrons que $t \in J'_{u_0}$.

Il faut montrer que V est définie et C^1 au voisinage de (u_0, t) . Pour cela, on va utiliser la proposition suivante, qui donne V au voisinage de (u_0, t) en fonction des valeurs de V aux voisinages d'autres points. Sa démonstration est à la fin de la section ; c'est cette partie-là de la démonstration qui utilise l'hypothèse selon laquelle f ne dépend pas de t .

Proposition D.2

Pour tous $v \in U$, $s, t' \in \mathbb{R}$ tels que $(v, t') \in \Omega$ et $(V(v, t'), t_0 + (s - t')) \in \Omega$, on a que (v, s) appartient à Ω et

$$V(v, s) = V(V(v, t'), t_0 + (s - t')).$$

Pour conclure la démonstration, il suffit de trouver t' tel que V est définie et C^1 au voisinage de (u_0, t') et V est définie et C^1 au voisinage de $(V(u_0, t'), t_0 + (t - t'))$: ensuite, d'après la proposition D.2, l'application V au voisinage de (u_0, t) s'écrit comme composée de deux applications bien définies et C^1 ; elle est donc C^1 . Si elle est C^1 , elle vérifie nécessairement les équations (Cauchy $\frac{dV}{du_0}$) (par dérivation de (Cauchy u_0)).

On sait que V est définie et C^1 sur un voisinage de $(V(u_0, t), t_0)$, d'après l'hypothèse (4.7). Or $(V(u_0, t'), t_0 + (t - t')) \rightarrow (V(u_0, t), t_0)$ lorsque $t' \rightarrow t$ donc, pour tout t' assez proche de t , V est définie et C^1 sur un voisinage de $(V(u_0, t'), t_0 + (t - t'))$.

De plus, comme t est dans l'adhérence de J'_{u_0} , il existe des $t' \in J_{u_0}$ arbitrairement proches de t qui appartiennent à J'_{u_0} , c'est-à-dire que V est définie au voisinage de (u_0, t') et C^1 . Il suffit de choisir un tel t' , assez proche de t pour que V soit aussi définie et C^1 sur un voisinage de $(V(u_0, t'), t_0 + (t - t'))$.

Démonstration de la proposition D.2. Soient $v \in U$, $s, t' \in \mathbb{R}$ tels que $(v, t') \in \Omega$ et $(V(v, t'), t_0 + (s - t')) \in \Omega$.

Vérifions que $J_v = J_{V(v, t')} + t' - t_0$ et, pour tout $\tau \in J_v$, $u_v(\tau) = u_{V(v, t')}(t_0 + \tau - t')$. Définissons

$$\psi : \tau \in J_{V(v, t')} + t' - t_0 \rightarrow u_{V(v, t')}(t_0 + \tau - t').$$

Les fonctions u_v et ψ sont toutes deux solutions du problème de Cauchy

$$\begin{aligned} u' &= f(u), \\ u(t') &= V(v, t'). \end{aligned}$$

Elles sont de plus maximales (car si on pouvait les prolonger, u_v et $u_{V(v, t')}$ admettraient également un prolongement qui serait solution du problème (Cauchy u_0) et ne seraient donc pas maximales). D'après le théorème de Cauchy-Lipschitz, elles sont égales, comme on l'avait annoncé.

Pour $\tau = s$, l'égalité entre u_v et ψ donne bien

$$V(v, s) = u_v(s) = u_{V(v, t')}(t_0 + s - t') = V(V(v, t'), t_0 + (s - t')).$$

□

D.4 Démonstration de la proposition 4.13

La démonstration est assez similaire à celle de la proposition 4.8.

Soit $v \in B(u_0, \frac{\epsilon}{2})$.

On vérifie d'abord que, pour tout $t \in J_v \cap]t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1}[$, $u_v(t) \in B(u_0, \epsilon)$. Par l'absurde, supposons qu'il existe $t \in J_v \cap]t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1}[$ tel que $\|u_v(t) - u_0\|_2 \geq \epsilon$. Par symétrie, on peut supposer qu'il existe de tels t dans la moitié droite de l'intervalle, $[t_0; t_0 + \frac{\epsilon}{2M_1}[$. Définissons t_1 comme l'infimum des réels t vérifiant cette propriété.

Par continuité de u_v , on a $\|u_v(t_1) - u_0\|_2 \geq \epsilon$. En revanche, pour tout $t \in [t_0; t_1[$, $u_v(t) \in B(u_0, \epsilon)$ et donc

$$\|u'_v(t_1)\|_2 = \|f(u_v(t_1))\|_2 \leq M_1.$$

Donc u_v est M_1 -lipschitzienne sur $[t_0; t_1[$ et

$$\begin{aligned} \|u_v(t_1) - u_0\|_2 &\leq \|u_v(t_1) - u_v(t_0)\|_2 + \|u_v(t_0) - u_0\|_2 \\ &\leq M_1|t_1 - t_0| + \|v - u_0\|_2 \\ &< \epsilon. \end{aligned}$$

On a obtenu une contradiction.

L'inclusion $]t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1}[\subset J_v$ provient du théorème des bouts. En effet, si $\sup J_v < t_0 + \frac{\epsilon}{2M_1}$, l'application u_v doit sortir de tout compact au voisinage de $\sup J_v$, ce qui est en contradiction avec le fait que u_v reste dans $B(u_0, \epsilon)$ sur $J_v \cap]t_0 - \frac{\epsilon}{2M_1}; t_0 + \frac{\epsilon}{2M_1}[$. De même si $\inf J_v > t_0 - \frac{\epsilon}{2M_1}$.

Bibliographie

Sylvie Benzoni-Gavage. *Calcul différentiel et équations différentielles : Cours et exercices corrigés*. Dunod, 2010.

F. Paulin. Topologie, analyse et calcul différentiel. https://www.imo.universite-paris-saclay.fr/~frederic.paulin/notescours/cours_analyseI.pdf, 2009.