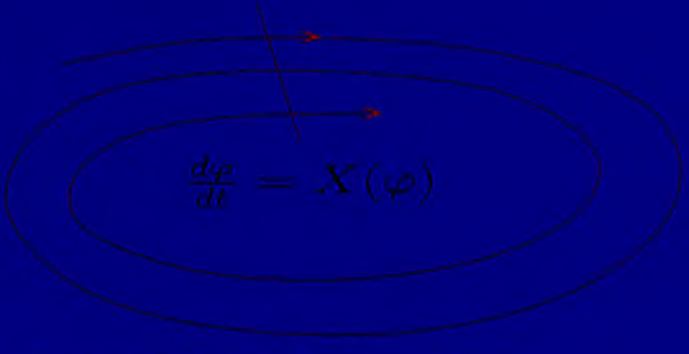
COLLECTION ENSEIGNEMENT SUP //// Mathématiques

L3M1

Des équations différentielles aux systèmes dynamiques I

THÉORIE ÉLÉMENTAIRE DES ÉQUATIONS DIFFÉRENTIELLES
AVEC ÉLÉMENTS DE TOPOLOGIE DIFFÉRENTIELLE

$$df(a)[h] = h.f'(a)$$



Robert Roussarie et Jean Roux

EDP SCIENCES

DES ÉQUATIONS DIFFÉRENTIELLES AUX SYSTÈMES DYNAMIQUES

ES

Tome 1

Théorie élémentaire des équations différentielles avec éléments de topologie différentielle

Robert Roussarie et Jean Roux

Collection dirigée par Daniel Guin

EDP SCIENCES

17, avenue du Hoggar
Parc d'activités de Courtabœuf, BP 112
91944 Les Ulis Cedex A, France

Illustration de couverture : La formule est l'expression de la différentielle d'une application dans un cas particulier. La figure est une illustration du théorème de Poincaré-Bendixson, avec le comportement en spirale de l'orbite par un point x non récurrent du champ.

Imprimé en France

ISBN : 978-2-7598-0512-9

Impression & brochage **sepec** - France

Numéro d'impression : 04545111250 - Dépôt légal : décembre 2011

IMPRIM'VERT®

TABLE DES MATIÈRES

AVANT-PROPOS

La motivation initiale de cet ouvrage a été la transcription de cours donnés oralement aux ingénieurs de la direction des Études et Recherches d'EDF. Depuis la première mouture du texte, la maturation a été longue mais, citons Héraclite (cinquième siècle avant notre ère), « Le temps est un enfant qui joue, en déplaçant les pions ». Ces cours proposaient une introduction au calcul différentiel, quelques éléments de la théorie qualitative des équations différentielles et quelques prolongements sur des idées plus récentes concernant les systèmes dynamiques. Lorsqu'il a été question de passer de l'exposé oral à une version écrite, il est apparu qu'il serait bon d'en étoffer le contenu. Pour ce faire, nous avons utilisé la matière d'un autre cours consacré aux équations différentielles, non publié jusqu'alors et enseigné par le premier auteur au niveau de la licence de Mathématique à l'université de Bourgogne. Dans l'intention de rendre le contenu plus autonome, nous avons aussi décidé d'y adjoindre des prérequis sur le calcul différentiel ainsi que des notions plus avancées de topologie différentielle. Les quelques aperçus sur les systèmes dynamiques présentés lors des premiers cours oraux ont alors pu être développés en une introduction plus conséquente à ce vaste sujet.

Nous sommes ainsi arrivés à un ouvrage structuré en deux tomes, avec comme idée directrice de faire en sorte que le texte se suffise à lui-même et soit le plus progressif possible. L'ensemble peut se lire et s'utiliser à plusieurs niveaux. On peut se limiter au tome 1, comportant deux parties (I, II), pour trouver un cours classique sur les équations différentielles, abordable dans le cadre de la licence de Mathématique, ou une initiation à des notions de base indispensables aux applications. Le tome 1 permet de rendre cette initiation autonome, indépendante d'une formation universitaire en Mathématique. Le tome 2 est une ouverture vers la théorie moderne des systèmes dynamiques. Il peut être utilisé dans le cadre d'un master de Mathématique ou de Physique. Il peut aussi être employé avec profit par toute personne cherchant des compléments sur certains aspects récents de la théorie des systèmes dynamiques.

Comme il est dit plus haut, la partie I, assez courte, est consacrée à des prérequis de calcul différentiel et de topologie différentielle. Cette partie ne contient pas de longs développements ; on y trouvera peu de démonstrations. On se contente d'y définir les termes et notions de base utilisés par la suite, concernant aussi bien le calcul différentiel dans un espace euclidien (différentielle d'une application, développement limité, formules de Taylor, étude d'une fonction au voisinage d'un point critique) que la topologie différentielle (variétés, espace tangent, plongements), cadre naturel pour développer la théorie des systèmes dynamiques et même celle des simples équations différentielles. Illustrons notre propos. Un système différentiel défini dans un espace euclidien, mais astreint à laisser invariantes des intégrales premières, (par exemple un système mécanique intégrable), va se restreindre à des sous-variétés de cet espace euclidien (par exemple à des sphères). De même, un système bi-périodique dans le plan s'étudie au mieux s'il est considéré comme un système sur le tore T^2. Il est donc important de pouvoir disposer du langage de base de la topologie différentielle. Dans cette partie, on a aussi introduit le théorème de Sard, car celui-ci sert en particulier de base à la notion de propriété générique, notion indispensable pour la théorie des systèmes dynamiques.

La partie II peut être considérée comme la matière d'un cours classique sur la théorie qualitative des équations différentielles, par exemple dans l'esprit du livre de Lefschetz [15], c'est-à-dire dans la tradition qui a été initiée par Poincaré [22]. Les démonstrations sont souvent données avec quelques détails (cependant, on n'établit pas la différentiabilité des trajectoires, mais seulement leur existence et leur unicité en tant que courbes continues, ce qui suffit à donner une idée assez précise de la nature du théorème de Cauchy, à la source de la théorie qualitative des équations différentielles). On développe dans cette partie les principales notions qualitatives de base dérivant d'une notion centrale, celle du flot d'un champ de vecteurs. Le texte inclut par exemple le théorème de Poincaré-Bendixson relatif à la trivialité des récurrences dans le plan, avec comme préalable une preuve du théorème de Jordan pour une courbe simple plongée dans le plan, le théorème du voisinage tubulaire, la notion d'indice pour un point singulier isolé pour un champ en dimension 2, la notion d'application de Poincaré pour une orbite périodique, celle de suspension d'un difféomorphisme et un chapitre sur les notions de stabilité de points singuliers et d'orbites périodiques de champs et de difféomorphismes.

Le flot d'un champ de vecteurs est un exemple de système dynamique et le tome 2 est une introduction plus systématique à l'étude des systèmes dynamiques. Nous n'avons pas voulu développer ni même introduire tous les aspects de la théorie moderne de ces systèmes. Le tome 2 ne propose donc pas une présentation systématique des systèmes dynamiques et certains domaines essentiels sont à peine

abordés, comme par exemple la théorie ergodique, ou bien celle des systèmes hamiltoniens. On a voulu au contraire se concentrer sur quelques thèmes de nature assez topologique et les développer précisément. On trouvera ainsi un chapitre consacré à l'exposition des idées de René Thom sur les notions de généricité et de transversalité, un chapitre sur les propriétés locales au voisinage des singularités hyperboliques, un chapitre consacré à la stabilité structurelle et deux autres à la théorie des bifurcations. Un dernier chapitre est consacré au modèle de Lorenz. Il est l'occasion de jeter un pont entre les systèmes dynamiques en dimension infinie et leur réduction éventuelle en dimension finie.

Les références au tome 1 sont toujours précédées par un I ou II selon la partie concernée à l'intérieur de celui-ci ; les références au tome 2 sont toujours précédées par un III. Les références internes, que ce soit dans une partie du tome 1 ou dans le tome 2, sont faites sans la mention du chiffre romain.

Le deuxième auteur remercie le LMD (Laboratoire de météorologie dynamique) puis le CERES (Centre d'enseignement et de recherche sur l'environnement et la société) de l'École normale supérieure, de lui avoir permis l'écriture de cet ouvrage, en collaboration avec Robert Roussarie.

R. Roussarie, J. Roux
5 mai 2011

Première partie

Éléments de topologie différentielle

1

PRÉLIMINAIRES DE CALCUL DIFFÉRENTIEL

1.1. Différentielle

1.1.1. Définitions

Soit f une application d'un ouvert U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ où n et p sont deux entiers quelconques ≥ 1.

Définition 1.1. On dit que f est différentiable en $a \in U$ s'il existe une application linéaire L de $\mathbb{R}^n$ dans $\mathbb{R}^p$ telle que

$$f(a+h) - f(a) = L[h] + ||h||\varepsilon(h) \tag{1.1}$$

pour $h \in \mathbb{R}^n$ suffisamment voisin de $0 \in \mathbb{R}^n$; $\varepsilon(h)$ est une application définie au voisinage de $0 \in \mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ et $\varepsilon(h) \to 0$ avec $h \to 0$; $\| \cdot \|$ désigne une norme quelconque de $\mathbb{R}^n$, par exemple la norme euclidienne. L'application f est dite *différentiable* si elle est différentiable en chaque point a de U.

Une application différentiable en un point a est donc une application dont l'accroissement en ce point peut être approché par une application linéaire. Une application différentiable est nécessairement continue.

Remarque 1.1. (a) On désignera par $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ l'espace linéaire des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$ *est représentée* par une matrice à p lignes et n colonnes $(L_{ij})_{ij}$ où $i \in \{1, \ldots, p\}$ est l'indice de ligne et $j \in \{1, \ldots, n\}$ est l'indice de colonne. Si h est un vecteur de $\mathbb{R}^n$, on

désigne par $L[h] \in \mathbb{R}^p$ l'évaluation de L sur h. Explicitement, si $h = (h_1, \ldots, h_n)^T$ et $L[h] = (v_1, \ldots, v_p)^T$, on a $v_i = \sum_{j=1}^n L_{ij} h_j$ pour $i = 1, \ldots, p$. Ainsi $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ s'identifie à l'espace $\mathrm{Mat}(p, n)$ des matrices à p lignes et n colonnes.

Il est important de noter que cette identification est liée au fait que l'on utilise les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$, choix qui semble aller de soi. Il sera cependant indispensable de pouvoir changer de bases et, dans ce cas, l'identification se fera par un isomorphisme différent : la matrice représentative sera modifiée par composition de matrices carrées, à droite et à gauche. Il faudra donc bien distinguer l'espace linéaire $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ de l'espace isomorphe (après choix de bases) des matrices $\mathrm{Mat}(p, n)$. Pour une application différentiable au point a, on distingue la *différentielle* appartenant à $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ de la (matrice) jacobienne représentative dans $\mathrm{Mat}(p, n)$. Pour simplifier et par abus de notations, on les notera toutes les deux par $df(a)$, mais il faudra être attentif au fait que la jacobienne est une matrice qui dépend du choix des bases. Dans le cas où l'on utilisera une base autre que la base canonique, on réservera la notation $df(a)$ pour la différentielle et l'on choisira un nom différent pour la jacobienne, par exemple $Jf(a)$.

(b) $\mathbb{R}^n$ est considéré comme l'espace vectoriel des vecteurs colonnes à n composantes dans $\mathbb{R}$. Le symbole T désignant la transposition des matrices, la transposée de la ligne $(h_1, \ldots, h_n)$ est le vecteur colonne $(h_1, \ldots, h_n)^T \in \mathbb{R}^n$ (les vecteurs lignes sont les éléments du dual : $\mathbb{R}^{n*}$). C'est cette notation qui a été utilisée ci-dessus. Cependant cette convention rigoureuse serait difficile à maintenir tout au long du texte. Elle nous conduirait par exemple à noter par : $f((x_1, \ldots, x_n)^T)$ la fonction à n variables notée usuellement par $f(x_1, \ldots, x_n)$. *Sauf lorsqu'on voudra mettre l'accent sur le fait que le vecteur considéré est un vecteur colonne* (colonne d'une matrice, vecteur considéré comme matrice colonne dans un produit matriciel, etc.) on notera par $(h_1, \ldots, h_n)$ le vecteur de $\mathbb{R}^n$ de composantes : $h_1, \ldots, h_n$. Il n'y aura pas en général de confusion possible et ce petit abus de notation allègera le texte.

(c) *Notations de Landau :* une application de la forme $||h|| \cdot \varepsilon(h)$ est représentée par le symbole $o(||h||)$. Ce symbole désigne donc toute application *négligeable devant* $||h||$, c'est-à-dire dont le quotient par $||h||$ tend vers 0 quand h tend vers 0. Plus généralement, on peut remplacer $||h||$ par une fonction positive test quelconque φ définie au voisinage de 0 dans $\mathbb{R}^n$: le symbole $o(\varphi)$ désigne une application arbitraire, définie au voisinage de $0 \in \mathbb{R}^n$ de la forme $\varphi(h)\varepsilon(h)$ où $\varepsilon(h)$ est une application définie au voisinage de $0 \in \mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ et tendant vers 0 lorsque h tend vers 0. Par exemple $o(1)$ désigne une application qui tend vers 0. On introduit aussi le symbole $O(\varphi)$ pour désigner toute application $F(h)$ pour laquelle il existe une constante M_F telle que l'on ait $||F(h)|| \leq M_F \varphi$ dans un voisinage de $0 \in \mathbb{R}^n$ (ici $\| \cdot \|$ est une norme de

$\mathbb{R}^p$). L'intérêt des notations de Landau est d'éviter d'avoir à donner des noms au reste d'une formule, comme on l'a fait dans la définition de la différentielle par exemple. D'autre part, ces symboles sont régis par des règles de calcul évidentes : $o(\varphi) + o(\varphi) \subset o(\varphi)$, $o(1)O(\varphi) \subset o(\varphi)$, $o(\varphi)O(\eta) \subset o(\varphi\eta)$, etc. (remarquez que les symboles $o(\varphi)$, $O(\varphi)$ désignent des classes d'applications, d'où l'utilisation du symbole $\subset$ ci-dessus).

Il est facile de montrer que l'application linéaire L de la formule (1.1) est uniquement définie. Cela conduit à la définition :

Définition 1.2. Supposons qu'une application f soit différentiable en un point a. L'unique approximation linéaire L intervenant dans la formule (1.1) est appelée *différentielle de f en a*. On la note $df(a)$. Son évaluation sur un vecteur $h \in \mathbb{R}^n$ sera notée $df(a)[h] \in \mathbb{R}^p$. On dit qu'une application f de $U \subset \mathbb{R}^n$ dans $\mathbb{R}^p$ est *différentiable (sur U)* si elle est différentiable en tout $x \in U$. Sa *différentielle* est l'application $x \in U \mapsto df(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$.

Remarque 1.2. Il est très facile de vérifier qu'une application différentiable sur un ouvert U est aussi continue sur U. Par contre, il existe des fonctions continues qui ne sont pas différentiables. Par exemple la fonction $\sqrt{|x|}$ est continue sur $\mathbb{R}$, mais n'est pas différentiable en $x = 0$.

Lorsque $n = 1$, l'application f est une courbe de $\mathbb{R}^p$. On pourra supposer que le domaine de définition U est un intervalle ouvert $I \in \mathbb{R}$. Dans ce cas, la différentielle va s'identifier avec la multiplication du nombre h par le *vecteur dérivé*, vecteur que l'on note $f'(a) = \frac{df}{dt}(a) \in \mathbb{R}^p$ (t désigne la coordonnée sur I). On peut définir directement le vecteur dérivé comme limite :

$$f'(a) = \lim_{(h \to 0,\, h \neq 0)} \frac{f(a+h) - f(a)}{h}. \tag{1.2}$$

La relation avec la différentielle est donnée par

$$df(a)[h] = hf'(a). \tag{1.3}$$

(Attention à la position de la variable h : c'est un scalaire qui doit être mis devant le vecteur $f'(a)$!)

On peut évidemment identifier dans ce cas différentielle et vecteur dérivé grâce à l'application $\rho : L \in \mathcal{L}(\mathbb{R}, \mathbb{R}^p) \mapsto L[1] \in \mathbb{R}^p$, qui est un isomorphisme linéaire : on a $f'(x) = \rho(df(x))$.

Il est à remarquer que cet isomorphisme est canonique, c'est-à-dire uniquement défini par le fait que $\{1\}$ est une base distinguée de $\mathbb{R}$: seul $\mathbb{R}$, parmi tous les espaces vectoriels sur $\mathbb{R}$, a une telle base distinguée.

En toute dimension n, on peut se ramener à ne considérer que des vecteurs de $\mathbb{R}^p$ en introduisant les dérivées directionnelles.

Définition 1.3. Soit f définie sur un ouvert $U \in \mathbb{R}^n$ et $u \in \mathbb{R}^n$. On appelle *dérivée directionnelle de f dans la direction u* le vecteur dérivé en $t = 0$ de la courbe $t \mapsto f(a + tu)$ (autrement dit de la restriction de f à la droite $t \mapsto a + tu$). On note cette dérivée : $f'_u(a)$.

En particulier, on peut considérer les directions des vecteurs de la base canonique de $\mathbb{R}^n$: $e_i = (0, \ldots, 1, \ldots, 0)$ dont les composantes sont nulles sauf un 1 en i-ème position, pour $i = 1, \ldots, n$. On appelle *dérivée partielle dans la i-ème direction* (ou plus simplement : *d'indice i*) la dérivée directionnelle $f'_{e_i}(a)$. Elle est notée $\frac{\partial f}{\partial x_i}(a) \in \mathbb{R}^p$, si $(x_1, \ldots, x_n)$ sont les coordonnées de $\mathbb{R}^n$, ou simplement $\partial_i f(a)$ s'il n'y a pas d'ambiguïté sur le nom des coordonnées.

Si une dérivée partielle $\frac{\partial f}{\partial x_i}(x)$ est définie pour tout point x, on a une nouvelle application $\frac{\partial f}{\partial x_i}$ définie sur U. Un cas très particulier de la formule de composition des différentielles (qui sera rappelée plus loin) est la relation suivante entre différentielle et dérivées partielles.

Proposition 1.1. *Supposons que f soit différentiable en $a \in U$. Alors pour tout $u \in \mathbb{R}^n$, la dérivée directionnelle $f'_u(a)$ existe et on a :*

$$f'_u(a) = df(a)[u], \tag{1.4}$$

d'où il suit l'expression de la différentielle en termes des dérivées partielles d'ordre 1 :

$$df(a)[(h_1, \ldots, h_n)] = \sum_{i=1}^{n} h_i \partial_i f(a). \tag{1.5}$$

Inversement, si les n dérivées partielles existent et sont continues, *la différentielle existe alors en tout point et est donnée par (1.5) [6].*

Remarque 1.3. Remarquez qu'il n'est pas vrai qu'une application soit différentiable si et seulement si toutes les dérivées partielles existent. La fonction de $\mathbb{R}^2$ définie par $f(x, y) = \frac{xy}{x^2+y^2}$ pour $(x, y) \neq (0, 0)$ et $f(0, 0) = 0$ est continue avec des dérivées partielles définies en tout point, mais n'est pas différentiable en $(0, 0)$. En effet, un calcul direct montre que les dérivées directionnelles existent en $(0, 0)$

et ont pour expression $f'_{(u_1,u_2)}(0,0) = \frac{u_1 u_2}{u_1^2+u_2^2}$, si $(u_1,u_2) \neq 0$. L'application $u \mapsto f'_u(0,0)$ n'est pas linéaire, ce qui serait le cas si la fonction f était différentiable en $(0,0)$, d'après (1.4).

L'expression (1.5) est capitale et nous allons l'expliciter plus loin. Disons déjà qu'elle permet en pratique de calculer la différentielle $df(a)$ en exprimant que sa matrice représentative a pour colonnes les vecteurs $\partial_i f(a)$.

On peut itérer la définition de différentielle : si une application f de U dans $\mathbb{R}^p$ est différentiable sur U, elle définit une nouvelle application df de U à valeurs dans $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p) \sim \mathbb{R}^{np}$, et ainsi de suite. Ceci permet de définir les notions de k-différentiabilité et de fonction de classe $\mathcal{C}^k$.

Définition 1.4. On dit qu'une application $f : U \mapsto \mathbb{R}^p$ est k fois différentiable si l'on peut définir par récurrence les différentielles successives $d^{s+1}f = d(d^s f)$ pour tout $s, 0 \leq s \leq k-1$ (conventionnellement, on pose $d^0 f = f$). Cette définition a un sens pour $k = +\infty$. On dit alors que la fonction est indéfiniment différentiable.

On dit qu'une application est (de classe) $\mathcal{C}^k$, pour $1 \leq k \leq +\infty$, si elle est k fois différentiable et que toutes les différentielles successives jusqu'à l'ordre k sont continues. On dira que f est $\mathcal{C}^\infty$ si elle est $\mathcal{C}^k$ pour tout k. On étend cette définition à $k = 0$ en disant qu'une application $\mathcal{C}^0$ est une application continue. Enfin une application analytique réelle sera dite de classe $\mathcal{C}^\omega$.

Remarque 1.4. Une application est de classe $\mathcal{C}^\infty$ si et seulement si elle est de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$. Comme une application différentiable est continue, il revient au même de dire qu'une application de classe $\mathcal{C}^\infty$ est une application indéfiniment différentiable. Dans le chapitre 3, nous définirons les dérivées partielles d'ordre supérieur. Cela nous permettra de généraliser la proposition 1.1.

1.1.2. Expressions de la différentielle

Nous allons examiner l'expression de la différentielle en partant de cas particuliers pour aller au cas général.

(a) *Cas $n = 1$, p quelconque.* Nous avons déjà traité de ce cas. Ici, il n'y a qu'une direction, et la différentielle de $f : t \in I \mapsto \mathbb{R}^p$ va être équivalente à la donnée du vecteur dérivé $f' = \frac{df}{dt}$. La relation entre ces deux notions est précisée dans la formule (1.3). La simplicité de ce cas particulier par rapport au cas général tient au fait que si f est différentiable, l'application qui à chaque x

fait correspondre le vecteur dérivé $f'(x)$ est à nouveau une application de I à valeurs dans $\mathbb{R}^p$. On peut itérer la définition de l'application vecteur dérivé : on a l'application dérivée seconde $f'' = \frac{d^2f}{dt^2}$ et ainsi de suite. Nous allons faire un grand usage des vecteurs dérivés par la suite. Si la variable t est le temps, l'application sera appelée une trajectoire (elle représente un mouvement) ; le vecteur dérivée s'appelle la *vitesse*, le vecteur dérivée seconde s'appelle l'*accélération* (du mouvement).

(b) *Cas n quelconque, $p = 1$.* Dans ce cas, f est une fonction (à n variables). Chaque dérivée partielle $\partial_i f$ est aussi une fonction. La différentielle en un point a est une *forme linéaire*, encore appelée covecteur. L'espace des formes linéaires $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ est appelé espace dual de $\mathbb{R}^n$ et est souvent noté $\mathbb{R}^{n*}$. Rappelons que la base duale $\{e_1^*, \ldots, e_n^*\}$ de $\mathbb{R}^{n*}$ est définie par les conditions $e_i^*[e_j] = \delta_{ij}$ (le symbole de Kronecker : $\delta_{ij} = 0$ si $i \neq j, \delta_{ii} = 1$). Autrement dit, e_i^* est la projection de $\mathbb{R}^n$ sur sa i-ème composante. En appliquant la définition de la différentielle, on voit que $dx_i = e_i^*$ (où, par abus de notation, on note par x_i la fonction projection). C'est la notation que nous utiliserons dorénavant pour e_i^*. Avec cette notation, on écrira (1.5) sous la forme :

$$df(a) = \sum_{i=1}^{n} \partial_i f(a) dx_i. \tag{1.6}$$

Comme $\mathbb{R}^{n*}$ est isomorphe à $\mathbb{R}^n$, la différentielle d'une fonction f est une application $x \in U \mapsto (\partial_1 f(x), \ldots, \partial_n f(x)) \in \mathbb{R}^n$.

(c) *Le cas général : n et p quelconques.*

L'application $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}^p$ s'écrit $f(x) = (f_1(x), \ldots, f_p(x))$, où les fonctions f_i sont les composantes de f (en toute rigueur, on devrait écrire $(f_1(x), \ldots, f_p(x))^T$!). Il est trivial que la différentiabilité de l'application f en a soit équivalente à la différentiabilité de toutes les fonctions composantes. En fait la différentielle $df_i(a) = \sum_{j=1}^n \partial_j f_i(a) dx_j$ forme la i-ème ligne de la matrice représentative de $df(a)$, encore appelée *matrice jacobienne* de f en a (cela suit par exemple du théorème 1.1 ci-après, en remarquant que $f_i = \pi_i \circ f$, où π_i est la projection sur la i-ème coordonnée dans $\mathbb{R}^p$). Comme il a été noté plus haut, les vecteurs colonnes de cette matrice sont les dérivées partielles (vectorielles) $\partial_j f(a) = (\partial_j f_1(a), \ldots, \partial_j f_p(a))^T$. On peut donc écrire (avec un léger abus de notation : identification des applications linéaires et de leurs matrices représentatives) :

$$df(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right) = \left(\partial_j f_i(a)\right), \quad i = 1, \ldots, p, \quad j = 1, \ldots, n \tag{1.7}$$

où i est l'indice de ligne (p lignes) et j l'indice de colonne (n colonnes).

Exemple. Soit l'application $f : (x_1, x_2) \mapsto f(x_1, x_2) = (x_1 x_2^2, x_1^2 + x_2^3)$ de $\mathbb{R}^2$ dans $\mathbb{R}^2$ (ici $n = p = 2$). Calculons les différentes dérivées partielles : $\frac{\partial f_1}{\partial x_1} = x_2^2$, $\frac{\partial f_1}{\partial x_2} = 2x_1x_2$, $\frac{\partial f_2}{\partial x_1} = 2x_1$ et finalement $\frac{\partial f_2}{\partial x_2} = 3x_2^2$. L'expression de la différentielle est donc :

$$df(x_1, x_2) = \begin{pmatrix} x_2^2 & 2x_1x_2 \\ 2x_1 & 3x_2^2 \end{pmatrix}.$$

1.1.3. Composition des différentielles

Théorème 1.1 **[6].** *Soit U un ouvert de $\mathbb{R}^n$ et V un ouvert de $\mathbb{R}^p$, une application f de U dans $\mathbb{R}^p$ telle que $f(U) \subset V$ et g une application de V dans $\mathbb{R}^q$. On suppose que f soit différentiable en $a \in U$ et que g soit différentiable en $f(a)$. Alors l'application composée $g \circ f$ est différentiable en a et sa différentielle est égale à :*

$$d(g \circ f)(a) = dg(f(a)) \circ df(a). \tag{1.8}$$

Autrement dit, à la composition des fonctions correspond la composition des différentielles.

Remarquez que la définition de la notion de dérivée directionnelle en $a \in U \subset \mathbb{R}^n$ d'une application f utilise la différentiation de l'application composée de l'application affine $\varphi : t \mapsto a + tu$ avec f. Comme $d\varphi(0)[h] = hu$, la formule (1.4) : $f'_u(a) = df(a)[u]$ de la proposition 1.1 est une conséquence directe du théorème 1.1.

Le théorème 1.1 permet d'établir par récurrence sur la classe de différentiabilité le résultat théorique suivant :

Proposition 1.2. *Soit f une application de U dans $\mathbb{R}^p$ et g une application de V dans $\mathbb{R}^q$ comme dans le théorème. Supposons que ces applications soient différentiables de classe C^k avec $0 \leq k \leq \infty$. Alors l'application composée $g \circ f$ est aussi de classe C^k.*

Il est très important pour la pratique du calcul différentiel d'expliciter la formule (1.8). Pour simplifier les notations, nous allons désigner les applications par le nom de la variable de l'espace but : si $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_p)$, $z = (z_1, \ldots, z_q)$ sont les coordonnées de $\mathbb{R}^n, \mathbb{R}^p$ et $\mathbb{R}^q$ respectivement, on désigne $f(x)$ par $y(x)$, $g(y)$ par $z(y)$ et $g \circ f(x)$ par $z(x)$. Alors, la règle de composition (1.8) s'écrit :

$$\frac{\partial z_l}{\partial x_i}(x) = \sum_{j=1}^{p} \frac{\partial z_l}{\partial y_j}(y(x)) \cdot \frac{\partial y_j}{\partial x_i}(x) \tag{1.9}$$

où $l = 1, \ldots, q$ et $i = 1, \ldots, n$.

Exemple. Soit $F(x_1,x_2,x_3)=\sqrt{x_1^2+x_2^2+x_2^2}$. On prend $U=\mathbb{R}^3-\{0\}$, $V=\mathbb{R}^+$, $y=f(x_1,x_2,x_3)=x_1^2+x_2^2+x_3^2$ et $g(y)=\sqrt{y}$. On vérifie immédiatement que $n=3$, $p=q=1$ et que $F=g\circ f$. D'après la formule (1.6) et la convention sur ∂_i, on a $dF(x_1,x_2,x_3)=\sum_{i=1}^{3}\frac{\partial F(x_1,x_2,x_3)}{\partial x_i}dx_i$. La formule (1.9) nous donne $\frac{\partial F(x_1,x_2,x_3)}{\partial x_i}=\frac{1}{2\sqrt{y}}\frac{\partial f(x_1,x_2,x_3)}{\partial x_i}=\frac{x_i}{\sqrt{y}}$. On a donc finalement $dF(x_1,x_2,x_3)=\frac{1}{\sqrt{x_1^2+x_2^2+x_2^2}}(x_1dx_1+x_2dx_2+x_3dx_3)$.

1.2. Formule des accroissements finis

La définition de la différentielle nous dit que celle-ci approxime l'accroissement de l'application d'une manière asymptotique, lorsque l'accroissement h de la variable tend vers 0. Il est possible de donner une estimation plus quantitative de l'accroissement de l'application, sous des conditions très générales sur la différentielle. Considérons des normes sur $\mathbb{R}^n$ et $\mathbb{R}^p$, notées toute deux $\|\cdot\|$. On rappelle que si $L\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)$, on appelle norme d'opérateur de L la norme définie par $||L||=\mathrm{Sup}\{||L[h]|| \mid h\in\mathbb{R}^n \text{ avec } ||h||=1\}$. C'est cette norme que l'on considère ci-dessous. On a alors le *théorème des accroissements finis* :

Théorème 1.2 **[6].** *Soit f une application différentiable de l'ouvert $U\subset\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$. Soit a,b deux points de U tels que le segment de droite $[a,b]$ soit contenu dans U. Supposons que* $\mathrm{Sup}\{||df(z)[h]|| \mid z\in[a,b]\}=M$ *soit fini. Alors*

$$||f(b)-f(a)\,||\leq M||b-a||. \tag{1.10}$$

1.3. Théorème de l'inverse, difféomorphisme

Théorème 1.3 **(Théorème de l'inverse [6]).** *Soit U un ouvert de $\mathbb{R}^n$, $a\in U$ et $f: U\mapsto\mathbb{R}^n$ une application de classe C^r, $1\leq r\leq\infty$. On suppose que la différentielle $df(a)$ est un* isomorphisme linéaire *de $\mathbb{R}^n$. Alors, il existe un ouvert $W\subset U$, contenant a, et un ouvert W' de $\mathbb{R}^n$ contenant $f(a)$ tels que f soit une bijection de W sur W', la bijection réciproque étant aussi de classe C^r.*

Remarque 1.5. La différentielle $df(a)$ est représentée par la matrice jacobienne des dérivées partielles $\left(\frac{\partial f_i}{\partial x_j}(a)\right)$. Dire que cette différentielle est un isomorphisme linéaire est équivalent à dire qu'elle est bijective et cette condition est équivalente à $\det\left(\frac{\partial f_i}{\partial x_j}(a)\right)=\left|\frac{\partial f_i}{\partial x_j}(a)\right|\neq 0$.

En appliquant le théorème de composition des différentielles à l'identité $f\circ f^{-1}=\mathrm{Id}$, on obtient $df(x)\circ d(f^{-1})(f(x))\equiv\mathrm{Id}$. La différentielle $df(x)$ est un isomorphisme linéaire pour tout $x\in W$, donc inversible. Il en résulte que $d(f^{-1})$

s'exprime en fonction de f, f^{-1} et de df : $d(f^{-1})(y) = (df)^{-1}(f^{-1}(y))$, pour $y \in W'$. Cette formule permet de montrer facilement, par récurrence sur la classe de différentiabilité, que l'application inverse d'une application de classe $\mathcal{C}^r$ est *nécessairement* de classe $\mathcal{C}^r$, comme il est affirmé dans l'énoncé du théorème.

Définition 1.5. Une bijection f d'un ouvert W de $\mathbb{R}^n$ sur un autre ouvert W' de $\mathbb{R}^n$ qui est différentiable de classe $\mathcal{C}^r$ ainsi que son inverse de W' sur W est appelée *difféomorphisme de classe $\mathcal{C}^r$ de W sur son image.*

L'application inverse d'un difféomorphisme de classe $\mathcal{C}^r$ est aussi un difféomorphisme de classe $\mathcal{C}^r$. Le théorème de l'inverse dit que l'application f est un difféomorphisme local en a, au sens où il est nécessaire de restreindre f à un voisinage W du point a : la partie linéaire en a de f (sa différentielle en a) est *globalement* inversible, mais l'application n'est en général elle-même que *localement* inversible au voisinage de a.

Remarque 1.6. Une application peut être différentiable et être un homéomorphisme sans être un difféomorphisme. Par exemple, la fonction $x \mapsto x^3$ est un homéomorphisme de classe $\mathcal{C}^\infty$ de $\mathbb{R}$ sur $\mathbb{R}$ mais n'est pas un difféomorphisme car son inverse $x \mapsto \text{signe}(x)|x|^{1/3}$ est continu mais non de classe $\mathcal{C}^1$.

Les *propriétés différentiables* de classe $\mathcal{C}^r$ sont par définition celles qui sont invariantes par les difféomorphismes de classe $\mathcal{C}^r$: par exemple, le fait pour deux courbes d'être tangentes en un point est une propriété de classe $\mathcal{C}^1$, le fait pour une courbe d'avoir une courbure finie est une propriété différentiable de classe $\mathcal{C}^2$, et ainsi de suite.

Un difféomorphisme de l'ouvert W sur l'ouvert W' est aussi ce qui est traditionnellement appelé *changement de coordonnées locales* (autrefois, on disait : *changement de coordonnées curvilignes* car un tel changement envoie en général les droites sur des courbes). Dans cette interprétation, les nouvelles coordonnées sont les coordonnées de l'ouvert W et les coordonnées anciennes (celles que l'on change) sont les coordonnées de l'ouvert W'.

Par exemple, considérons le plan usuel muni d'une origine O et d'une base orthonormale qui définit une orientation du plan. Soit la fonction $\Phi : (\theta, r) \mapsto (r\cos\theta, r\sin\theta)$ qui, à tout couple (θ, r) de nombres réels, associe le point M du plan de coordonnées $x = r\cos\theta$ et $y = r\sin\theta$. Tout couple (θ, r) tel que $\Phi(\theta, r) = M$ est un *système de coordonnées polaires* de M (figure 1.1). Ce système se construit de la façon suivante : soit Ou l'axe, passant par M, tel que $\widehat{(\vec{Ox}, \vec{Ou})}$ ait pour mesure θ (mesuré en radians). Le vecteur unitaire $\vec{u}$ de Ou a

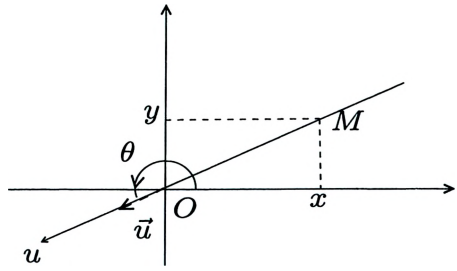


FIGURE 1.1. Coordonnées polaires.

pour coordonnées $(x = \cos\theta, y = \sin\theta)$. On a donc $\overrightarrow{OM} = r\vec{u}$, de sorte que M est le point de Ou d'abscisse r sur Ou. Remarquons que l'application Φ est surjective mais non injective. Remarquons tout d'abord que : $r^2 = x^2 + y^2$. Distinguons ensuite les deux cas suivants :

1. M est à l'origine, alors $r = 0$ et θ est arbitraire.
2. M n'est pas à l'origine, alors $r \neq 0$: $r = \epsilon\sqrt{x^2 + y^2}$ avec $\epsilon = \pm 1$. Le nombre θ doit satisfaire aux conditions

$$\cos\theta = \frac{x}{\epsilon\sqrt{x^2 + y^2}}, \quad \sin\theta = \frac{y}{\epsilon\sqrt{x^2 + y^2}}.$$

Comme la somme des carrés des valeurs imposées à $\cos\theta$ et $\sin\theta$ est égale à 1, il y a, pour chaque choix de ϵ, des solutions en θ qui forment une classe de nombres réels congrus modulo 2π. L'application Φ est donc bien surjective mais non injective globalement ; elle n'est bijective que localement en (r, θ).

On remarque que

$$d\Phi(\theta, r) = \begin{pmatrix} -r\sin\theta & r\cos\theta \\ \cos\theta & \sin\theta \end{pmatrix} = -r.$$

La différentielle est toujours inversible sauf pour $r = 0$. Donc, pour différentes raisons, l'application Φ n'est pas un difféomorphisme global. Mais si on restreint (r, θ) à un rectangle ouvert : $U =]0, +\infty) \times]\theta_1, \theta_2[$, avec $0 < \theta_2 - \theta_1 < 2\pi$, l'application Φ devient un difféomorphisme de U sur son image.

Un autre exemple est celui des coordonnées sphériques en dimension 3. Dans l'espace ordinaire, choisissons une origine O et une base orthonormale $(\vec{e_1}, \vec{e_2}, \vec{e_3})$. À tout $(\varphi, \lambda, r) \in \mathbb{R}^3$ faisons correspondre le point M de coordonnées cartésiennes $x = r\cos\lambda\cos\varphi, y = r\cos\lambda\sin\varphi, z = r\sin\lambda$: autrement dit, $M = \Phi(\varphi, \lambda, r) = (r\cos\lambda\cos\varphi, r\cos\lambda\sin\varphi, r\sin\lambda)$. On vérifie facilement que $r^2 = x^2 + y^2 + z^2$,

l'ensemble des points de l'espace pour lesquels r est égal à une constante donnée, est une sphère de centre O. D'où le nom de *coordonnées sphériques* donné aux variables (φ, λ, r) (figure 1.2).

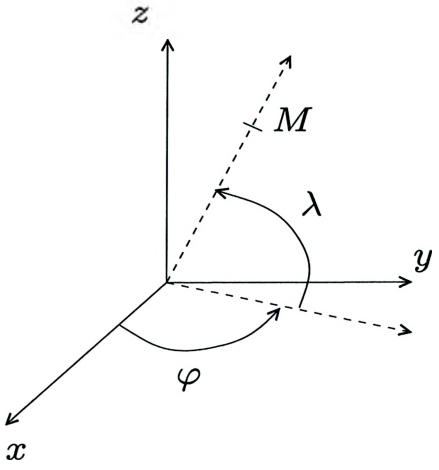


FIGURE 1.2. Coordonnées sphériques.

Comme dans le cas des coordonnées polaires, l'application $\Phi : (\varphi, \lambda, r) \in \mathbb{R}^3 \mapsto M(x, y, z) \in \mathbb{R}^3$ est surjective et non injective. Pour retrouver un difféomorphisme, il faut restreindre convenablement le domaine de Φ, par exemple à un ouvert $U = \{(r, \lambda, \varphi) \in \mathbb{R}^3 \,|\, 0 < r, \lambda_1 < \lambda < \lambda_2, \varphi_1 < \varphi < \varphi_2\}$, avec $-\frac{\pi}{2} < \lambda_1 < \lambda_2 < \frac{\pi}{2}$ et $0 < \varphi_2 - \varphi_1 < 2\pi$. L'application Φ est alors un difféomorphisme de U sur son image qui est un secteur sphérique ouvert, centré à l'origine.

Remarque 1.7.

(a) En dimension 1, une fonction qui vérifie en chaque point d'un intervalle ouvert I les hypothèses du théorème de l'inverse est un *difféomorphisme global :* la fonction est un difféomorphisme de I sur son image. En effet une fonction de dérivée non nulle en chaque point d'un intervalle ouvert est strictement monotone : elle est donc nécessairement injective, et une application qui est localement un difféomorphisme (au voisinage de chaque point de son domaine de définition) et injective est un difféomorphisme global. Cette propriété de monotonicité est bien évidemment basée sur l'existence d'un ordre sur $\mathbb{R}$, compatible avec la topologie.

(b) Il n'existe pas de tel ordre en dimension ≥ 2. En conséquence, si la dimension est ≥ 2, une application peut vérifier en tout point de son domaine de

définition U les conditions du théorème ($df(x)$ inversible en tout point) sans être pour cela un difféomorphisme global sur U, même si U est connexe. Par exemple l'application exponentielle complexe $\exp z = e^z$ a une dérivée complexe égale à e^z et elle est donc non nulle en tout $z \in \mathbb{C}$ (ce qui implique que, en tant qu'application de $\mathbb{R}^2$ dans $\mathbb{R}^2$, sa différentielle est inversible). Et pourtant cette application est périodique de période $2\pi i$ *et n'est donc pas un difféomorphisme global.* Remarquez d'ailleurs que l'image de $\mathbb{C}$ par e^z est égale à $\mathbb{C} \setminus \{0\}$ et on peut montrer que $\mathbb{C}$ n'est pas homéomorphe à $\mathbb{C} \setminus \{0\}$.

1.4. Théorème des fonctions implicites

Nous allons maintenant montrer qu'une application différentiable ayant en un point une différentielle *surjective* est localement surjective et même qu'elle est équivalente à difféomorphisme local près (c'est-à-dire à un changement de coordonnées près) à une projection linéaire. Ce résultat, qui généralise le théorème de l'inverse, en est en fait une conséquence directe.

Nous allons tout d'abord introduire quelques notations pour une application définie sur un espace produit :

Définition 1.6. Soit n, p deux entiers quelconques ≥ 1 et un ouvert U de $\mathbb{R}^{p+n} = \mathbb{R}^p \times \mathbb{R}^n$. On considère des coordonnées $x \in \mathbb{R}^p$ et $y \in \mathbb{R}^n$. Supposons que f soit une application différentiable de U à valeurs dans $\mathbb{R}^q$, pour un $q \in \mathbb{N}$. Soit $(a, b) \in U$ (cette notation (a, b) signifie que l'on prend $a \in R^p$ et $b \in \mathbb{R}^n$). On notera par $f_a(y) = f(a, y)$ et par $f_b(x) = f(x, b)$ les *applications partielles* définies respectivement sur $U \cap \{a\} \times \mathbb{R}^n$ et $U \cap \mathbb{R}^p \times \{b\}$. Ces applications sont différentiables (comme composées d'injections linéaires avec f). On dira que leurs différentielles $df_a(y), df_b(x)$ sont des *différentielles partielles de* f. Elles sont définies respectivement sur $\mathbb{R}^n$ et $\mathbb{R}^p$, espaces que l'on préfère noter $\{0\} \times \mathbb{R}^n$ et $\mathbb{R}^p \times \{0\}$, pour rappeler qu'ils sont paramétrés par les coordonnées y et x respectivement.

Remarquez que $df_a(y) = df(a, y)|_{\{0\} \times \mathbb{R}^n}$ et que $df_b(x) = df(x, b)|_{\mathbb{R}^p \times \{0\}}$. La matrice jacobienne de $df_a(y)$ est formée des n dernières colonnes de la matrice jacobienne de $df(a, y)$ et celle de $df_b(x)$ des p premières colonnes de celle de $df(x, b)$.

On a alors le théorème suivant, lorsque l'espace de définition de la fonction est donné comme un produit.

Théorème 1.4 **(Forme moderne du théorème des fonctions implicites [6]).** *Soit n, p deux entiers quelconques ≥ 1 et un ouvert U de $\mathbb{R}^{p+n} = \mathbb{R}^p \times \mathbb{R}^n$. On considère des coordonnées $x \in \mathbb{R}^p$ et $y \in \mathbb{R}^n$. Supposons que f soit une application de classe $\mathcal{C}^r$, $r \geq 1$, de U à valeurs dans $\mathbb{R}^n$. Soit un point $(a,b) \in \mathbb{R}^p \times \mathbb{R}^n$. Posons $f(a,b) = c$. On suppose que la différentielle $df_a(b) : \{0\} \times \mathbb{R}^n \mapsto \mathbb{R}^n$ soit inversible. Alors, il existe des voisinages ouverts T de $a \in \mathbb{R}^p$ et T' de $b \in R^n$, un difféomorphisme H de classe $\mathcal{C}^r$, de $T \times T'$ dans U, de la forme*

$$H(x,y) = (x, h(x,y)), \tag{1.11}$$

avec $H(a,c) = (a,b)$ (ce qui est équivalent à dire que $h(a,c) = b$), qui vérifie l'identité

$$f \circ H(x,y) = f(x, h(x,y)) = y \quad \text{sur} \quad T \times T'. \tag{1.12}$$

Ce théorème est une conséquence directe du théorème de l'inverse. En effet, il suffit d'appliquer le théorème de l'inverse à $F(x,y) = (x, f(x,y))$ qui est une application de classe $\mathcal{C}^r$ de U à valeurs dans $\mathbb{R}^p \times \mathbb{R}^n$. On vérifie sans peine que la différentielle de F est inversible en (a,b). Le difféomorphisme H inverse de F donné par le théorème des fonctions implicites peut être restreint à un voisinage du point (a,c), de la forme $W' = T \times T'$. Ce difféomorphisme H s'écrit manifestement, comme F elle-même, sous la forme $H(x,y) = (x, h(x,y))$.

La forme particulière de H fait que ce difféomorphisme préserve globalement chacun des sous-espaces $\{x\} \times \mathbb{R}^n$, et donc qu'il envoie tout facteur $\mathbb{R}^p \times \{y\}$ *sur le graphe d'une application* de classe $\mathcal{C}^r$, de T à valeurs dans $\mathbb{R}^n$. Nous allons maintenant exploiter cette remarque pour énoncer la forme classique du théorème des fonctions implicites dans lequel on fait mention explicitement d'une fonction implicite.

Théorème 1.5 **(La forme classique du théorème des fonctions implicites).** *On suppose vérifiées les conditions du théorème précédent et on désigne par T et T' les ouverts donnés par ce théorème. Alors il existe une application $\varphi : x \mapsto \varphi(x)$ de classe $\mathcal{C}^r$, de T à valeurs dans $\mathbb{R}^n$, avec $\varphi(a) = b$, un voisinage ouvert W de a, et un voisinage ouvert V de (a,b) dans U, tels que*

$$\{(x,y) \in V \mid f(x,y) = c\} = \text{Graphe}(\varphi) = \{(x, \varphi(x)) \mid x \in W\}. \tag{1.13}$$

Remarque 1.8. Le théorème précédent affirme que l'équation $\{f(x,y) = c\}$ peut se résoudre localement en la variable y pour donner le graphe d'une application $y = \varphi(x)$. Cette application φ est donc donnée *implicitement* par l'équation, d'où le nom donné classiquement au théorème (trouver l'application φ à valeurs dans $\mathbb{R}^n$ est équivalent à trouver ses n fonctions composantes, d'où le nom de *théorème des fonctions implicites*).

Il est important de remarquer que l'on peut appliquer le théorème à une application $\mathcal{C}^r$ d'un ouvert $U \subset \mathbb{R}^q$ à valeurs dans $\mathbb{R}^n$, avec $n \leq q$, en un point $z^0 \in \mathbb{R}^q$ où la différentielle est surjective. En effet, à isomorphisme linéaire près, on peut alors écrire que $\mathbb{R}^q = \mathbb{R}^{q-n} \times \mathbb{R}^n$ où le facteur $\mathbb{R}^{q-n} \times \{0\}$ est le noyau de $df(z^0)$. La différentielle est alors inversible sur le facteur $\{0\} \times \mathbb{R}^n$ et on peut appliquer le théorème 1.5.

Dans la pratique, voici comment procéder. On détermine la matrice jacobienne de l'application. La différentielle $df(z^0)$ est surjective si et seulement si sa matrice jacobienne représentative a un déterminant mineur $n \times n$ non nul. À une permutation près des coordonnées $(z_1, \ldots, z_q)$, on peut supposer que c'est le mineur sur les n dernières coordonnées qui est non nul. Renommons les coordonnées par $(x, y) = (x_1, \ldots, x_p, y_1, \ldots, y_n)$ avec $p = q - n$. Le théorème des fonctions implicites donne alors n fonctions $\varphi_1(x), \ldots, \varphi_n(x)$, composantes d'une application $\varphi(x)$ à valeurs dans $\mathbb{R}^n$, solution de l'équation $f(x, \varphi(x)) = c = f(z^0)$. Cette équation vectorielle est évidemment équivalente à un système de n équations en dimension 1 : $f_i(x, \varphi(x)) = c_i$ pour $i = 1, \ldots, n$.

Le théorème des fonctions implicites est un *théorème d'existence* qui ne permet pas de donner une expression explicite à l'application φ. Bien au contraire, ce théorème permet de construire de nouvelles fonctions à partir de fonctions déjà connues (par exemple la fonction arcsinus par inversion de la fonction sinus). Il est cependant possible de calculer explicitement la différentielle de φ au point b. Pour cela, différentions l'équation

$$f(x, \varphi(x)) \equiv c,$$

où $f : U \subset \mathbb{R}^p \times \mathbb{R}^n \mapsto \mathbb{R}^n$. On obtient :

$$d_x f(x, \varphi(x)) + d_y f(x, \varphi(x)) \circ d\varphi(x) \equiv 0,$$

et en particulier $d_x f(a, b) + d_y f(a, b) \circ d\varphi(a) = 0$ ($d_x f$ et $d_y f$ désignent les différentielles partielles dans les directions de $\mathbb{R}^n$ et $\mathbb{R}^p$). L'hypothèse faite est précisément que $d_y f(a, b)$ est inversible. On peut donc calculer $d\varphi(a)$ en fonction des dérivées partielles de f en (a, b) :

$$d\varphi(a) = -d_y f(a, b)^{-1} \circ d_x f(a, b).$$

Plus généralement on pourra calculer explicitement toutes les dérivées partielles de φ au point a.

Exemple. Soit la fonction f, définie sur $\mathbb{R}^3$ de coordonnées (x, y, z) :

$$f : (x, y, z) \mapsto f(x, y, z) = x^2 + xy^2 - z^3 y + z^4.$$

Par (1.6) on calcule que

$$df(x,y,z) = (2x + y^2)dx + (2xy - z^3)dy + (-3yz^2 + 4z^3)dz.$$

On peut appliquer le théorème des fonctions implicites aux points (x,y,z) où la différentielle $df(x,y,z)$ est surjective, ce qui est ici équivalent à dire qu'elle n'est pas nulle. On appellera ensemble singulier de f l'ensemble Σ_f des points où cette condition n'est pas satisfaite :

$$\Sigma_f = \{2x + y^2 = 2xy - z^3 = -3yz^2 + 4z^3 = 0\}.$$

Il est facile de vérifier que $\Sigma_f = \{(0,0,0)\}$. En particulier, on peut appliquer le théorème des fonctions implicites au point $(1,0,0)$. En ce point, remarquons que $\frac{\partial f}{\partial x}(1,0,0) = 2$ (les deux autres dérivées partielles sont nulles) et que $f(1,0,0) = 1$. Par le théorème des fonctions implicites, il existe donc une fonction unique $x = \varphi(y,z)$, définie au voisinage de $(0,0)$, vérifiant $f(\varphi(y,z),y,z) \equiv 1$ et $\varphi(0,0) = 1$. On a

$$d\varphi(0,0) = -\left(\frac{\partial f}{\partial x}(1,0,0)\right)^{-1}\left(\frac{\partial f}{\partial y}(1,0,0)dy + \frac{\partial f}{\partial z}(1,0,0)dz\right) = 0_{\mathbb{R}^2},$$

que l'on écrit, avec un abus de notation, $d\varphi(0,0) = 0$.

2

VARIÉTÉS ET SOUS-VARIÉTÉS

Les sous-espaces vectoriels ou bien affines $E^k \subset \mathbb{R}^n$ considérés en algèbre linéaire sont les modèles et les premiers exemples des sous-variétés de dimension k de $\mathbb{R}^n$. Ces espaces conservent les propriétés locales des espaces affines mais peuvent avoir des propriétés globales très différentes. Il est d'autre part intéressant de pouvoir considérer plus généralement des variétés, et pas seulement des sous-variétés plongées dans $\mathbb{R}^n$. En effet, beaucoup d'exemples importants sont donnés de façon naturelle par une construction indépendante de tout plongement.

Les sous-variétés, et plus généralement les variétés, sont les objets de la topologie différentielle. Ils servent de cadre au calcul différentiel moderne. Nous allons commencer en précisant la notion très utile de variété dont les sous-variétés seront des exemples plongés. Le théorème des fonctions implicites joue un rôle essentiel dans la construction des sous-variétés.

2.1. Variétés différentiables

On rappelle qu'un *homéomorphisme* d'un espace topologique sur un autre est une bijection bi-continue. Deux espaces topologiques sont dits homéomorphes s'il existe un homéomorphisme de l'un sur l'autre. Être homéomorphe est manifestement une relation d'équivalence, et deux espaces homéomorphes sont indiscernables du point de vue de leurs propriétés topologiques.

2.1.1. Définitions

Définition 2.1. Une variété différentiable V, de dimension p, de classe $\mathcal{C}^s$, avec $p \in \mathbb{N}$ et $0 \leq s \leq \infty$, est un espace topologique muni d'un recouvrement ouvert $\{U_i\}_{i\in I}$ (I, ensemble d'indices quelconques) et, pour chaque $i \in I$, d'une application $\varphi_i : U_i \mapsto \mathbb{R}^p$ tels que :

1. Chaque φ_i est un homéomorphisme de l'ouvert U_i sur son image $\Omega_i = \varphi_i(U_i)$, ouvert de $\mathbb{R}^p$.

2. Si $U_i \cap U_j \neq \emptyset$, l'application $\varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \mapsto \mathbb{R}^p$ est un difféomorphisme de classe $\mathcal{C}^s$ de l'ouvert $\varphi_i(U_i \cap U_j)$ de $\mathbb{R}^p$ sur son image $\varphi_j(U_i \cap U_j)$ (un homéomorphisme si $s = 0$) (figure 2.1).

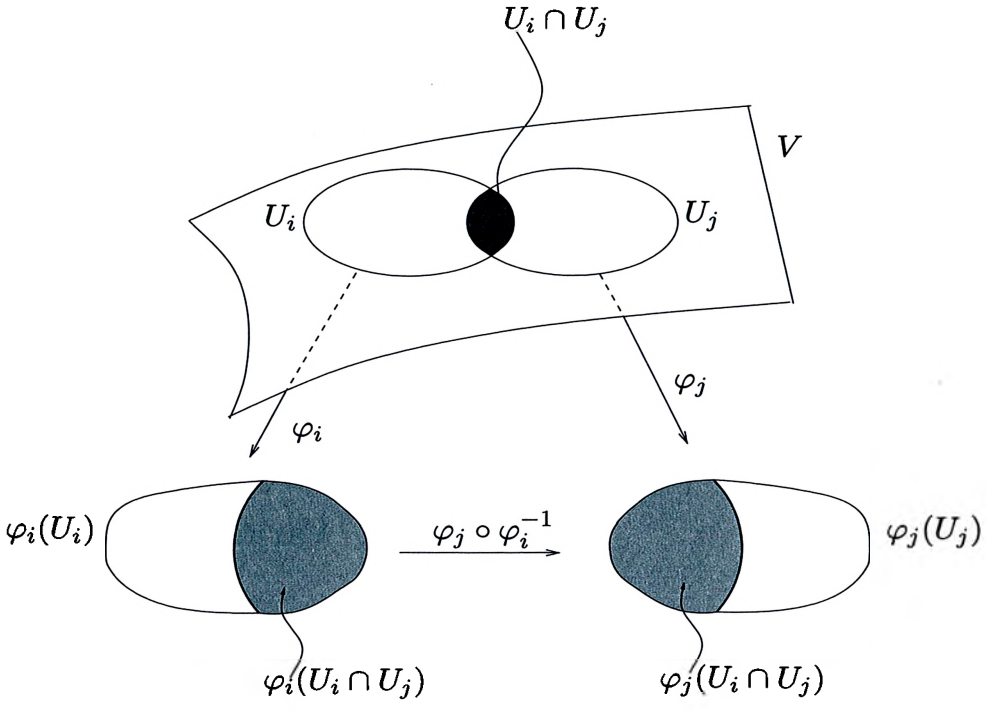


FIGURE 2.1. Changement de cartes.

Chaque paire (U_i, φ_i) est appelée *carte* de la variété et la collection des cartes $\{(U_i, \varphi_i)\}_{i\in I}$ est un *atlas*. Les applications $\varphi_j \circ \varphi_i^{-1}$ sont appelées : *changements de carte*. Cette terminologie vient de la considération des cartes d'un atlas

géographique de la Terre. En effet, le dessin de chaque carte définit implicitement une application (celle qui fait correspondre à chaque point de la région sur la Terre, définie par une carte, le point correspondant sur le dessin). Il est clair que les changements de cartes sont implicitement supposés être différentiables.

À vrai dire, on ne parle de variété différentiable que si $s \geq 1$. Pour $s = 0$, on parle de variété topologique et, dans ce cas, la condition (2) de la définition est automatiquement vérifiée, avec la convention qui consiste à dire qu'un homéomorphisme est un difféomorphisme de classe $\mathcal{C}^0$. Une variété de classe $\mathcal{C}^s$, avec $s \geq 1$, est naturellement de toute classe $\mathcal{C}^\sigma$ pour tout σ, $0 \leq \sigma \leq s$. Si les changements de cartes sont de classe $\mathcal{C}^\omega$ (analytiques réels), on dit que la variété est analytique réelle. Dans la suite, on ne considérera, sauf mention expresse du contraire, que des variétés de classe $\mathcal{C}^\infty$, variétés que l'on appellera tout simplement *variétés différentiables*. On rappelle que l'on identifie $\mathbb{R}^2$ avec $\mathbb{C}$ de la manière canonique en posant $z = x + iy \in \mathbb{C}$ pour tout $(x, y) \in \mathbb{R}^2$. Une variété à cartes à valeurs dans $\mathbb{R}^{2n}$ identifié à $\mathbb{C}^n$, et à changements de cartes holomorphes, est dite *variété complexe* ou holomorphe (de dimension n).

Une variété de dimension 1 est appelée *courbe*. Une variété de dimension 2 est appelée *surface*. Si, sous l'identification canonique de $\mathbb{R}^2$ avec $\mathbb{C}$, les changements de cartes sont holomorphes, on dit que la surface est une *surface de Riemann* (ou bien une *courbe holomorphe*).

Définition 2.2. Soit V une variété comme plus haut définie par un *atlas* $\mathcal{A} = \{(U_i, \varphi_i)\}_{i \in I}$. On dit qu'une paire (U, φ) d'un ouvert U de V et d'un homéomorphisme φ de U sur un ouvert de $\mathbb{R}^p$ est une carte compatible avec l'atlas si, pour tout U_i tel que $U \cap U_i \neq \emptyset$, le point (2) de la définition précédente est vérifié pour les deux paires $(U_i, \varphi_i), (U, \varphi)$. On dit qu'un *atlas* est *maximal* s'il contient toutes les cartes compatibles avec lui-même. Un atlas quelconque est alors inclus dans un unique atlas maximal. On identifie la notion de structure différentiable d'une variété avec la donnée d'un atlas maximal.

Remarque 2.1.

1. Si (U, φ) est une carte, avec pour image $\varphi(U) = \Omega$ ouvert de $\mathbb{R}^p$, l'application φ^{-1} définie sur Ω est appelée *paramétrisation locale* ou *système de coordonnées locales*.

2. En particulier, pour qu'un changement de cartes locales assure la même valeur à l'intégrale d'une forme différentielle sur une variété V, on est conduit à n'autoriser que les changements de cartes pour lesquels le jacobien du changement de variable, i.e. le changement de coordonnées locales par φ_i (associé

à un ouvert U_i), est strictement positif pour tout i. Si un tel atlas maximal existe, alors la *variété est orientable.* Le choix d'un tel atlas définit une orientation sur V. Le ruban de Möbius – voir l'item 5 du paragraphe 2.1.3 – est un exemple de variété non orientable.

3. Il est très commode de considérer l'atlas maximal de la structure différentiable de V, pour pouvoir choisir la bonne carte pour traiter tel ou tel problème (par exemple, en chaque point $m \in V$ et pour tout voisinage W de m dans V, on peut choisir une carte (U, φ) telle que $m \in U \subset W$). Par contre, pour des raisons pratiques, on essayera de définir la structure de la variété en exhibant un atlas qui ne contienne qu'un minimum de cartes, comme on va le voir dans les exemples qui suivent (un tel atlas sera appelé : atlas de définition de la variété).

2.1.2. Topologie quotient

Avant de donner quelques exemples de variétés, il convient peut-être de rappeler quelques notions de base concernant la notion de *topologie quotient.* Si E est un ensemble muni d'une relation d'équivalence $\mathcal{R}$, on appelle ensemble quotient, désigné par $E/\mathcal{R}$, l'ensemble des classes d'équivalence. L'application ρ qui, à chaque $x \in E$, associe sa classe d'équivalence $\dot{x}$ est appelée *projection canonique* : $\dot{x} = \rho(x)$. Si maintenant E est un espace topologique, on définit sur l'espace quotient $E/\mathcal{R}$ une unique topologie, dite *topologie quotient*, dont on supposera toujours muni l'espace quotient. Cette topologie est la plus fine topologie qui rend continue la projection canonique. Ses ouverts sont les images par ρ des *ouverts de E saturés (ou invariants) par $\mathcal{R}$*, c'est-à-dire des ouverts qui sont la réunion de classes d'équivalence.

Cette topologie est caractérisée par la propriété suivante :

Si F est un espace topologique quelconque, les applications continues de $E/\mathcal{R}$ dans F sont les applications obtenues par passage au quotient des applications de E dans F qui sont à la fois continues et $\mathcal{R}$-invariantes (une application f sur E est $\mathcal{R}$-invariante si $x\mathcal{R}y$ implique que $f(x) = f(y)$; une telle application passe au quotient, *c'est-à-dire induit une application $\dot{f}$ sur $E/\mathcal{R}$ définie par $\dot{f}(\dot{x}) = f(x)$).*

À titre d'exemple, on peut considérer $E = \mathbb{R}$ avec pour relation d'équivalence $\mathcal{R}$ la congruence modulo $\mathbb{Z}$: $t\mathcal{R}t' \iff t - t' \in \mathbb{Z}$. On note l'espace quotient $\mathbb{R}/\mathbb{Z}$. Les applications continues sur $\mathbb{R}/\mathbb{Z}$ sont obtenues par passage au quotient des applications continues et 1-périodiques.

Un *domaine fondamental* pour le quotient est une partie $D \subset E$ qui a une intersection non vide avec chaque classe d'équivalence. On peut induire sur D la topologie et la relation d'équivalence. L'inclusion de D dans E passe alors

au quotient pour définir un *homéomorphisme* de $D/\mathcal{R}$ sur $E/\mathcal{R}$. Pour traiter du quotient, on peut alors remplacer E par le domaine fondamental D. Cela sera d'autant plus avantageux que le domaine D sera petit. Par exemple, dans le cas du quotient $\mathbb{R}/\mathbb{Z}$, on peut prendre $D = [0,1]$. Les classes d'équivalence sur D sont des classes à un seul élément $\{t\}$ pour les $t \in]0,1[$ et la classe à deux éléments : $\{0,1\}$. Autrement dit, l'espace topologique $\mathbb{R}/\mathbb{Z}$ s'obtient en identifiant les extrémités du segment $[0,1]$.

2.1.3. Exemples de variétés

1. *Un ouvert U de $\mathbb{R}^n$* est trivialement une variété : la structure est définie par l'unique carte (U, Id). L'intérêt de cet exemple est d'inclure le calcul différentiel élémentaire sur les ouverts de $\mathbb{R}^n$ dans le cadre général de la topologie différentielle.

2. *Le cercle $T^1 = \mathbb{R}/\mathbb{Z}$* . On verra dans le paragraphe 2.2.3 qu'une réalisation de T^1 comme sous-variété de $\mathbb{R}^2$ est le cercle trigonométrique, ce qui explique le nom de cercle. Considéré comme sous-variété, on désignera T^1 par S^1, car le cercle trigonométrique est la sphère de dimension 1. La notation T^1 est justifiée par le fait que le cercle est aussi le tore de dimension 1. Le tore général de dimension n est introduit dans l'exemple suivant.

   Détaillons un peu la construction de T^1. On considère l'espace quotient de $\mathbb{R}$ par la relation de congruence des réels *modulo* les entiers : deux réels x, y sont identifiés si et seulement si $x - y \in \mathbb{Z}$. Soit ρ la projection canonique de $\mathbb{R}$ sur son quotient.

   Une remarque essentielle est que, si $0 < \tau < 1$, la projection canonique ρ induit une injection sur l'intervalle $]t_0, t_0 + \tau[$, pour tout $t_0 \in \mathbb{R}$. Autrement dit, la projection de cet intervalle coïncide avec la projection de son saturé $]t_0, t_0 + \tau[+\mathbb{Z}$. Il en résulte immédiatement que $\rho(]t_0, t_0 + \tau[)$ est un ouvert du quotient, et de plus que ρ est un homéomorphisme $]t_0, t_0 + \tau[$ sur son image (passer au quotient les sous-intervalles ouverts de $]t_0, t_0 + \tau[$).

   On peut alors définir une structure de variété sur S^1, de dimension 1 (et de classe C^∞), par les deux cartes suivantes (U_i, φ_i), $i = 1, 2$,

$$U_1 = \rho\left(]-\frac{1}{8}, \frac{5}{8}[\right), \; U_2 = \rho\left(]\frac{3}{8}, \frac{9}{8}[\right).$$

   On peut inverser ρ sur les deux images U_1 et U_2 pour définir les homéomorphismes $\varphi_i = (\rho|_{U_i})^{-1}$. L'ouvert $U_1 \cap U_2$ de T^1 a deux composantes connexes

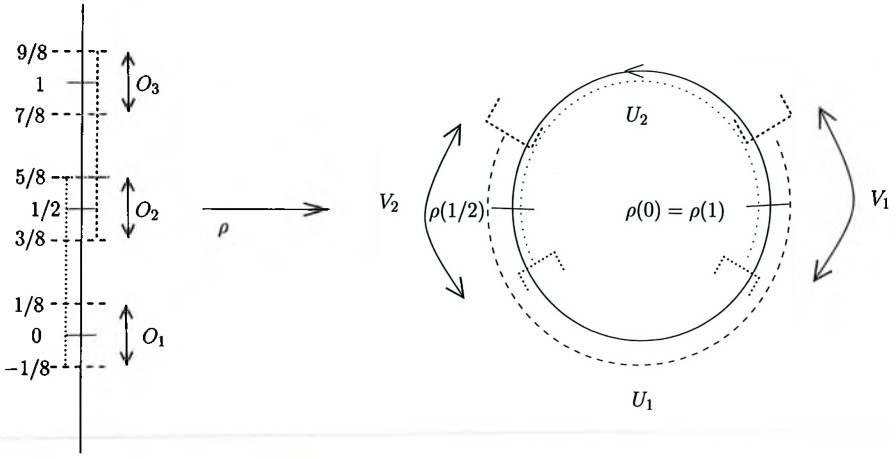


FIGURE 2.2. L'arc de cercle U_1 (resp. U_2) est représenté en tirets (resp. pointillés). Le sens de rotation sur S^1 est le sens trigonométrique.

$V_1 = \rho(]-\frac{1}{8}, \frac{1}{8}[)$ et $V_2 = \rho(]\frac{3}{8}, \frac{5}{8}[)$ (figure 2.2). Considérons

$$O_1 = \varphi_1(V_1) =]-\frac{1}{8}, \frac{1}{8}[,$$
$$O_2 = \varphi_1(V_2) = \varphi_2(V_2) =]\frac{3}{8}, \frac{5}{8}[$$

et

$$O_3 = \varphi_2(V_1) =]\frac{7}{8}, \frac{9}{8}[.$$

L'application $\varphi_2 \circ \varphi_1^{-1}$ envoie O_2 dans lui-même et O_1 sur O_3. Dans les deux cas, les valeurs t et $\varphi_2 \circ \varphi_1^{-1}(t)$ diffèrent par un élément de $\mathbb{Z}$.

Il en résulte à la fois que $\varphi_2 \circ \varphi_1^{-1}|_{O_2}(t) = t$ et $\varphi_2 \circ \varphi_1^{-1}|_{O_1}(t) = t + 1$ (notez qu'une fonction continue sur un intervalle et à valeurs dans $\mathbb{Z}$ doit être constante). Le cercle T^1 est donc une variété (topologique) avec un atlas de définition à deux cartes.

3. On peut généraliser l'exemple précédent de la façon suivante. On considère une action différentiable d'un groupe discret G agissant de façon *propre* et *libre* sur une variété V (voir [12] par exemple, p. 70-72). Une telle action est donnée par une application $\rho : (\gamma, x) \in G \times V \mapsto \rho(\gamma, x) \in V$, telle que $\rho(\gamma \cdot \gamma', x) = \rho(\gamma, \rho(\gamma' \cdot x))$ et que, pour tout $\gamma \in G$, l'application $x \mapsto \rho(\gamma, x)$ soit un difféomorphisme de V (c'est-à-dire une application différentiable ayant un inverse différentiable. Voir la définition 2.4). Dire

que l'action est libre et propre est équivalent à dire que *tout point $x \in V$, possède un voisinage ouvert U tel que les différents ouverts $\gamma \cdot U$ soient deux à deux disjoints.* Par exemple le groupe $G = \mathbb{Z}$ (groupe discret pour l'addition) agit sur $V = \mathbb{R}$ par translation : $(\gamma, x) \in \mathbb{Z} \times \mathbb{R} \mapsto x + \gamma \in \mathbb{R}$. Cette action est évidemment libre. De plus cette action est propre, car l'intervalle $U =]-\frac{1}{3} + x, \frac{1}{3} + x[$ est tel que $\gamma(U) \cap U = \emptyset$ pour tout $\gamma \in \mathbb{Z} - \{0\}$. Par contre l'action définie par $(\gamma, x) \mapsto \gamma \cdot x = 2^{\gamma} x$ pour $\gamma \in \mathbb{Z}$ n'est ni propre ni libre car $\gamma(0) = 0$ pour tout $\gamma \in \mathbb{Z}$. Dans la théorie des systèmes dynamiques, un tel ouvert U est dit erratique (pour l'action considérée).

On peut alors munir l'espace quotient V/ρ des orbites de l'action ρ d'une structure de variété en prenant l'atlas de définition suivant : les ouverts sont les images par la projection sur le quotient des ouverts erratiques et les applications de cartes sont les inverses locaux de l'application de projection. Il suffit évidemment de se limiter à un recouvrement ouvert du quotient. C'est précisément ce que l'on a fait dans l'exemple précédent de l'action de $\mathbb{Z}$ sur $\mathbb{R}$ en prenant les deux ouverts U_1 et U_2. Dans cet exemple, tout intervalle ouvert de longueur strictement plus petite que 1 est un ouvert erratique.

La variété obtenue est appelée *variété quotient* et est notée V/ρ. Si V est une variété complexe et que chaque difféomorphisme $x \mapsto \rho(\gamma, x)$ est holomorphe, la variété quotient V/ρ est une variété complexe.

4. *L'action de $\mathbb{Z}^n$ sur $\mathbb{R}^n$ par translation* (qui généralise l'exemple de T^1 : cas $n = 1$) est une application de cette construction. En effet, toute boule de $\mathbb{R}^n$ de rayon strictement plus petit que $\frac{1}{2}$ est erratique pour l'action de $\mathbb{Z}^n$ car, si deux points x et y de $\mathbb{R}^n$ sont tels que $x - y \in \mathbb{Z}^n - \{0\}$, ils ne peuvent pas appartenir à une même boule de rayon strictement plus petit que $\frac{1}{2}$. On définit de cette façon le *tore de dimension n* : $T^n = \mathbb{R}^n/\mathbb{Z}^n$.

   Dans l'identification canonique de $\mathbb{R}^2$ avec $\mathbb{C}$, les translations sont des difféomorphismes holomorphes. Il en résulte que $\mathbb{T}^2$ est naturellement une surface de Riemann (car les changements de cartes sont des translations).

5. *Le ruban de Möbius.* Considérons l'action ρ de $\mathbb{Z}$ sur $\mathbb{R}^2$, de coordonnées (x, y), donnée par $\rho : (x, y) \mapsto (x + 1, -y)$.

   Le quotient $\mathbb{R}^2/\rho$ est appelé : ruban de Möbius. Un domaine fondamental D de l'action de ρ est la bande $[0, 1] \times \mathbb{R}$, de bord $\partial D = (\{0\} \times \mathbb{R}) \bigcup (\{1\} \times \mathbb{R})$. Le ruban de Möbius peut donc se construire en identifiant les points $(0, y)$ et $(1, -y)$ pour tout $y \in \mathbb{R}$, sur le bord de D (figure 2.3).

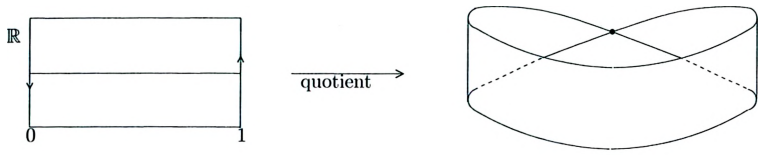


FIGURE 2.3. Ruban de Möbius, plongé dans $\mathbb{R}^3$, en dimension 2.

On peut généraliser le ruban de Möbius en toute dimension $n+1$ en remplaçant $\mathbb{R}$ par $\mathbb{R}^n$ de coordonnées $(y_1, \ldots, y_n)$ et en prenant pour ρ :

$$\rho(x, y_1, \ldots, y_n) = (x+1, -y_1, -y_2, \ldots, -y_n).$$

Les voisinages tubulaires d'une courbe simple (plongement de S^1) dans une surface sont difféomorphes soit à un ruban simple ($S^1 \times \mathbb{R}$) soit à un ruban de Möbius. Une surface est non orientable si et seulement si elle contient une courbe simple avec un voisinage tubulaire type ruban de Möbius. Cette remarque se généralise en toute dimension.

La bouteille de Klein et l'espace projectif réel (S^2/ρ où ρ est l'application $x \mapsto -x$) sont des exemples de surfaces compactes sans bord non orientables.

6. *La sphère de Riemann.* On appelle *projectif complexe de dimension* 1 *ou sphère de Riemann :* $P^1(\mathbb{C})$, l'espace des droites complexes de $\mathbb{C}^2$ passant par l'origine, autrement dit l'espace quotient

$$\left(\mathbb{C}^2 - \{0\}\right) / \{(z, Z) \sim (z', Z') \iff zZ' - z'Z = 0\}.$$

La classe d'équivalence de $(z, Z) \in \mathbb{C}^2 - \{0\}$ est notée : $[z, Z]$. On munit cet espace de la topologie quotient. On peut recouvrir $P^1(\mathbb{C})$ par les deux ouverts $U_1 = \{[z, 1] \,|\, z \in \mathbb{C}\}$ et $U_2 = \{[1, Z] \,|\, Z \in \mathbb{C}\}$. Les applications $\varphi_1 : [z, 1] \mapsto z$ et $\varphi_2 : [1, Z] \mapsto Z$ sont des homéomorphismes de U_1 et U_2 sur $\mathbb{C}$. On définit ainsi deux cartes (U_1, φ_1) et (U_2, φ_2). Le changement de cartes $\varphi_2 \circ \varphi_1^{-1}$ est le difféomorphisme holomorphe de $z \in \mathbb{C} - \{0\} \mapsto Z = \frac{1}{z} \in \mathbb{C} - \{0\}$: en effet, dire que $Z = \varphi_2 \circ \varphi_1^{-1}(z)$ revient à dire que $[z, 1] = [1, Z]$, ce qui s'écrit $zZ = 1$. Ces deux cartes font de $P^1(\mathbb{C})$ une surface de Riemann, puisque le changement de carte est holomorphe.

Il est, d'autre part, clair que topologiquement $P^1(\mathbb{C})$ peut être interprété comme l'adjonction d'un point à l'infini au plan (réel) des z (point de coordonnée complexe $Z = 0$ dans la carte (U_2, φ_2)). La surface $P^1(\mathbb{C})$ est donc homéomorphe à la sphère, d'où le nom de sphère de Riemann. On rappelle qu'adjoindre le point infini ∞ à un espace topologique E (opération connue

également sous le nom de *compactification d'Alexandroff*) signifie que l'on considère l'espace $\overline{E} = E \cup \{\infty\}$ muni de la topologie définie par les deux conditions suivantes : l'injection de E dans $\overline{E}$ est un homéomorphisme sur son image, et les complémentaires des parties compactes de E forment un système fondamental de voisinages de ∞ dans $\overline{E}$. Dans le cas particulier de la sphère, la compactification d'Alexandroff peut être réalisée pratiquement en utilisant la projection stéréographique à partir du pôle Nord, le point ∞ étant alors identifié au pôle Nord. Nous décrivons la projection stéréographique à la sous-section suivante.

7. Si M_1 et M_2 sont des variétés de classe $\mathcal{C}^s$ avec des atlas de définition $(U_i, \varphi_i)_{i\in I}$ et $(W_j, \psi_j)_{j\in J}$ respectivement, alors la collection $(U_i \times V_j, \varphi_i \times \psi_j)_{i\times j\in I\times J}$ définit, sur le produit topologique $M_1 \times M_2$, une structure de variété de classe $\mathcal{C}^s$, que l'on appellera variété produit et que l'on désignera évidemment par $M_1 \times M_2$. On a clairement que $\dim(M_1 \times M_2) = \dim(M_1) + \dim(M_2)$. L'opération de produit est associative à difféomorphisme près (voir la définition de difféomorphisme ci-dessous), ce qui permet de définir un produit quelconque de variétés $M_1 \times M_2 \ldots \times M_k$, à difféomorphisme près. Par exemple, T^n est difféomorphe à $T^1 \times \ldots \times T^1$, n fois.

2.1.4. Difféomorphisme entre variétés

Définition 2.3. Soit M et N deux variétés différentiables de classe $\mathcal{C}^s$. On dira qu'une application $f : M \mapsto N$ est une application de classe $\mathcal{C}^s$ si, pour toute paire de cartes (U, φ) de M et (V, ψ) de N telle que $f(U) \subset V$, l'application composée $\psi \circ f \circ \varphi^{-1}$ est de classe $\mathcal{C}^s$ (au sens usuel pour les applications entre ouverts d'espaces euclidiens). On définit de la même façon les applications analytiques réelles et holomorphes.

On a donné la définition de différentiabilité d'une application entre variétés, en la lisant localement dans des cartes. De la même façon, pour $s \geq 1$, on peut transposer la plupart des concepts du calcul différentiel élémentaire. Par exemple, on définira le *rang* de f au point $x \in U \subset M$ comme le rang de la différentielle $d(\psi \circ f \circ \varphi^{-1})(\varphi(x))$. Il est immédiat de voir que cette définition est indépendante du choix des cartes (U, φ) et (V, ψ) (avec $x \in U$).

Définition 2.4. On dit que $f : M \mapsto N$ est un difféomorphisme de classe $\mathcal{C}^s$ si et seulement si f est bijective et si f, ainsi que son application inverse f^{-1}, sont de classe $\mathcal{C}^s$. On dit alors que M et N sont difféomorphes en classe $\mathcal{C}^s$.

Deux variétés difféomorphes sont des espaces topologiques homéomorphes mais la réciproque est fausse. Le premier contre-exemple a été fourni dans les années 1960 par J. Milnor, qui a construit des *sphères exotiques* de dimension 7 : ces sphères exotiques sont des variétés de classe $\mathcal{C}^\infty$ qui sont homéomorphes, mais non difféomorphes, à la sphère S^7, variété que l'on définira dans le prochain paragraphe. Ceci est une affaire très délicate et il n'est pas question d'en dire plus ici.

Deux variétés difféomorphes en classe $\mathcal{C}^s$, $s \geq 1$, ont même dimension (car les différentielles $d(\psi \circ f \circ \varphi^{-1})(\varphi(x))$ sont nécessairement des isomorphismes linéaires). Ce résultat est même vrai pour $s = 0$, comme conséquence du fait que la dimension est un invariant topologique, ce qui est un théorème profond de la topologie algébrique !

Pour des variétés de classe $\mathcal{C}^s$, le fait d'être difféomorphes en classe $\mathcal{C}^s$ est une relation d'équivalence. Cette relation généralise la notion de changement de coordonnées entre ouverts euclidiens. Les *propriétés de la topologie différentielle en classe* $\mathcal{C}^s$ sont précisément les propriétés qui sont invariantes par les difféomorphismes de classe $\mathcal{C}^s$ (pour les propriétés locales, celles qui sont invariantes par choix de cartes). Par exemple : un ordre de contact k entre deux courbes tracées sur une surface de classe $\mathcal{C}^s$, si $k \leq s$, est une propriété de la topologie différentielle en classe $\mathcal{C}^s$.

Une autre construction de la sphère de dimension 2

Soit $S^2 = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 = 1\}$ la sphère unité de $\mathbb{R}^3$. Introduisons, pour couvrir S^2, les deux ouverts suivants : $U_1 = \{p \in S^2 \ : \ z(p) < 1\}$ et $U_2 = \{p \in S^2 \ : \ z(p) > -1\}$. On définit la projection stéréographique $\Pi_1 : U_1 \mapsto \mathbb{R}^2$, à partir du pôle Sud, par

$$\Pi_1(x, y, z) = (\xi_1, \eta_1) = \left(\frac{x}{1+z}, \frac{y}{1+z} \right);$$

(notez que $\frac{\xi_1}{x} = \frac{1}{1+z} = \frac{\eta_1}{y}$: voir figure 2.4).

De même, soit la projection stéréographique, à partir du pôle Nord, $\Pi_2 : \mathcal{U}_2 \mapsto \mathbb{R}^2$ définie par $\Pi_2(x, y, z) = (\xi_2, \eta_2) = (\frac{x}{1-z}, \frac{y}{1-z})$, (notez que $\frac{\xi_2}{x} = \frac{1}{1-z} = \frac{\eta_2}{y}$: voir figure 2.4). Les deux paires (U_1, Π_1) et (U_2, Π_2) définissent deux cartes sur S^2 avec pour changement de cartes : $\Pi_1 \circ \Pi_2^{-1}(z) = \frac{z}{|z|^2} = \frac{1}{\bar{z}}$ sur $\Pi_1(U_1 \cap U_2) = \mathbb{C} - \{0\}$

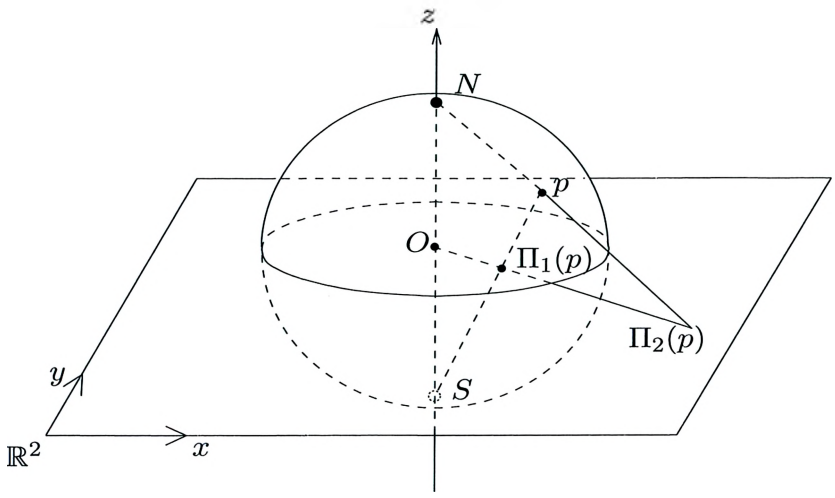


FIGURE 2.4

(on a identifié $\mathbb{R}^2$ avec $\mathbb{C}$ en posant $z = x + iy$). Géométriquement, ce changement de cartes est l'involution de $\mathbb{R}^2$ par rapport au cercle unité.

Cet atlas est à comparer à celui défini plus haut sur la sphère de Riemann. La seule différence est que le changement de cartes $z \mapsto \frac{1}{z}$ est remplacé par $z \mapsto \frac{1}{\bar{z}}$. On passe de l'un à l'autre par la composition par le difféomorphisme $z \mapsto \bar{z}$. Il en résulte immédiatement que la sphère de Riemann est difféomorphe en classe $\mathcal{C}^\infty$ à la sphère S^2 (notez que la sphère de Riemann est holomorphe, mais que S^2 est seulement définie en classe $\mathcal{C}^\infty$).

L'application de coordonnées sphériques Φ induit une application, que nous noterons encore Φ, de $\mathbb{R}^2$ sur S^2 définie par

$$x = \cos\lambda\cos\varphi,\ \ y = \cos\lambda\sin\varphi,\ \ z = \sin\lambda,$$

(on a posé $r = 1$ dans l'application de coordonnées sphériques). Cette application induit une application de classe $\mathcal{C}^\infty$ entre le tore T^2 et la sphère S^2 qui est surjective, mais loin d'être un difféomorphisme (par exemple le cercle $\gamma = \{\lambda = \pi/2\}$ est envoyé sur le pôle Nord). On peut d'ailleurs montrer facilement que la sphère et le tore ne sont pas homéomorphes : par exemple, le cercle γ a un complémentaire connexe sur T^2, alors que son image par un éventuel homéomorphisme de T^2 sur S^2 serait une courbe simple dont le complémentaire sur S^2 aurait 2 composantes connexes (ce qui est assez intuitif, mais suit en toute rigueur d'un théorème de Jordan). Si l'on ne peut pas paramétrer toute la sphère par les coordonnées sphériques, on peut utiliser ces coordonnées pour obtenir des coordonnées locales

(longitude et latitude) en dehors des deux pôles. Par exemple, si on prend un rectangle $\Omega = (\lambda, \varphi) \in]\lambda_1, \lambda_2[\times]\varphi_1, \varphi_2[$, avec $-\frac{\pi}{2} < \lambda_1 < \lambda_2 < \frac{\pi}{2}$ et $|\varphi_1 - \varphi_2| < 2\pi$, alors l'application Φ sur Ω est une paramétrisation locale de la sphère, autrement dit l'inverse d'une application de carte. On peut recouvrir par de telles cartes toute la sphère moins les deux pôles.

2.2. Sous-variété d'un ouvert de $\mathbb{R}^n$

2.2.1. Codimension. Sous-espaces vectoriels transverses

Soit E un espace vectoriel de dimension finie. Considérons S et T deux sous-espaces vectoriels de E. On note par $S + T$ le sous-espace vectoriel formé des sommes $x + y$ où x parcourt S et y parcourt T. Si $S + T = E$, on dit que S et T sont des sous-espaces *transverses* (dans E). Si de plus $S \cap T = \{0\}$, on dit que E est *somme directe* des sous-espaces S et T ou bien que S et T sont *supplémentaires* dans E. La somme directe de deux sous-espaces vectoriels S et T est notée par $S \oplus T$ (on admet le cas trivial $S = E, T = \{0\}$ pour lequel $E = S \oplus \{0\} = S$). On a trivialement que $E = S \oplus T$ est isomorphe au produit vectoriel $S \times T$. Ces notions s'étendent à un nombre fini quelconque de sous-espaces $E_1, \ldots, E_k \subset E$. On définit la somme

$$\sum_i E_i = \{x_1 + \ldots + x_k \mid x_i \in E_i,\ i = 1, \ldots, k\}.$$

Si $E = \sum_i E_i$ et $E_i \cap E_j = \{0\}$ pour $i \neq j$, on dit que les sous-espaces $E_1, \ldots, E_k$ sont supplémentaires ou bien que E est leur somme directe, que l'on note $E = \oplus_i E_i$ et qui est isomorphe au produit vectoriel $\Pi_i E_i$. La décomposition de chaque vecteur de E en somme $x = \sum_i x_i$, avec $x_i \in E_i, i = 1, \ldots, k$ est alors unique. Il s'ensuit immédiatement que :

$$\dim(\oplus_i E_i) = \sum_i \dim(E_i).$$

Il est possible maintenant de définir la codimension d'un sous-espace vectoriel.

Définition 2.5. La *codimension* d'un sous-espace S de E, notée : $\mathrm{codim}_E(S)$, est égale à $\dim(E/S)$, c'est-à-dire à la dimension de l'espace vectoriel quotient de E par S.

Si S est un sous-espace de E, il est facile de trouver un autre sous-espace T de E tel que S et T soient supplémentaires dans E. La projection canonique

$\rho : E \mapsto E/S$ induit alors un isomorphisme de T sur E/S. Il en résulte que

$$\mathrm{codim}_E(S) = \dim(E/S) = \dim(E) - \dim(S).$$

Si S, T sont des sous-espaces de E, alors $S \cap T$ est un sous-espace de S et T. On a le résultat suivant concernant les dimensions :

Lemme 2.1.

$$\dim(S) + \dim(T) = \dim(S \cap T) + \dim(S + T). \tag{2.1}$$

Démonstration. Choisissons $S' \subset S$, supplémentaire de $S \cap T$ dans S ainsi que $T' \subset T$ supplémentaire de $S \cap T$ dans T. Il est clair que S', T' et $S \cap T$ sont supplémentaires dans $S + T$, d'où il résulte que :

$$\dim(S') + \dim(T') + \dim(S \cap T) = \dim(S + T). \tag{2.2}$$

Comme

$$\dim(S') + \dim(S \cap T) = \dim(S),$$

et

$$\dim(T') + \dim(S \cap T) = \dim(T),$$

on obtient (2.1) en ajoutant $\dim(S \cap T)$ aux deux membres de (2.2). ∎

La formule (2.1) est évidemment équivalente à la formule suivante sur les codimensions :

$$\mathrm{codim}(S) + \mathrm{codim}(T) = \mathrm{codim}(S \cap T) + \mathrm{codim}(S + T). \tag{2.3}$$

On remarquera que, si l'on a $\dim(S) + \dim(T) < \dim(E)$, S et T ne peuvent pas être transverses (même si $\dim(S \cap T) = \{0\}$). D'autre part, si $\dim(S) + \dim(T) = \dim(E)$, alors S et T sont transverses si et seulement si S et T sont supplémentaires.

2.2.2. Définition d'une sous-variété d'un ouvert de $\mathbb{R}^n$

Définition 2.6. Soit $U \subset \mathbb{R}^n$ un ouvert et $p \in \mathbb{N}, 0 \leq p \leq n$. On dit que $V \subset U$ est une sous-variété de classe $\mathcal{C}^s$, de dimension p (et de codimension $n - p$) de U si, pour tout $x \in V$, il existe un voisinage ouvert W de x dans U et un difféomorphisme $\Phi : W \mapsto \mathbb{R}^p \times \mathbb{R}^{n-p} \approx \mathbb{R}^n$ de W sur son image tel que $\Phi(W \cap U) = \Phi(W) \cap (\mathbb{R}^p \times \{0\})$. On dit que la paire (W, Φ) est un difféomorphisme (local) de trivialisation de la sous-variété.

Remarque 2.2.

1. Il serait équivalent à la définition ci-dessus de considérer une collection de difféomorphismes de trivialisation $\{(W_i, \Phi_i)\}_{i\in I}$, choisie telle que les U_i recouvrent V, c'est-à-dire telle que $V \subset \cup_{i\in I} U_i$.

   En effet, si $\{(W_i, \Phi_i)\}_{i\in I}$ est une telle collection, chaque point $x \in V$ appartient à l'un des U_i et, inversement, si pour chaque point $x \in V$ on a une paire (W_x, Φ_x), la collection des $(W_x, \Phi_x)_{x\in V}$ est telle que les W_x recouvrent V. Évidemment, on aura intérêt à choisir une collection avec un minimum d'éléments.

2. Considérons une telle collection. Soit deux éléments (U_i, Φ_i) et (U_j, Φ_j) de la collection, tels que $V \cap U_i \cap U_j \neq \emptyset$. Alors $\Phi_j \circ \Phi_i^{-1}$ se restreint en un difféomorphisme d'un ouvert de $\mathbb{R}^p$ sur son image. Il en suit immédiatement que si on pose $U_i = W_i \cap V$ et $\varphi_i = \Phi_i|_{U_i} : U_i \mapsto \mathbb{R}^p$, la collection (U_i, φ_i) est un atlas de variété pour V (avec la topologie induite par $\mathbb{R}^n$). Donc *toute sous-variété est naturellement une variété* (de même classe et dimension).

3. La question inverse a été résolue par Whitney : toute variété de dimension p et *de topologie à base dénombrable* peut se réaliser comme sous-variété fermée de $\mathbb{R}^{2p}$ (on trouvera une démonstration facile pour une réalisation dans $\mathbb{R}^{2p+1}$ dans [12], p. 90-91). On dit alors que la variété est plongée dans l'espace euclidien (voir plus loin la définition de plongement).

4. Pour $p = 2$, on peut améliorer ce résultat : toute surface compacte orientable peut se réaliser facilement comme sous-variété de $\mathbb{R}^3$. En effet, pour tout $k \in \mathbb{N}$, considérons le polynôme :

$$P_k(x,y) = \Big(\Big(\frac{k+1}{2}\Big)^2 - \Big(x - \frac{k+1}{2}\Big)^2 - y^2\Big)\Pi_{i=1}^k\Big((x-i)^2 + y^2 - \frac{1}{9}\Big).$$

   La région compacte $\Omega_k = \{P_k(x,y) \geq 0\}$ est égale à $\Omega_k = \bar{D} \setminus \cup_{i=1}^k D_i$, où $\bar{D}$ est le grand disque fermé $\bar{D} = \Big\{\Big(x - \frac{k+1}{2}\Big)^2 + y^2 \leq \frac{k+1}{2}\Big)^2\Big\}$ et les D_i sont les petits disques ouverts $D_i = \Big\{(x-i)^2 + y^2 \leq \frac{1}{9}\Big\}$. Il s'ensuit que la surface Σ_k de $\mathbb{R}^3$ d'équation

$$\Sigma_k = \{z^2 = P_k(x,y)\},$$

   est une surface orientable de genre k (une sphère si $k = 0$, un tore si $k = 1$ et, plus généralement, une surface de genre k (appelée aussi « tore à k trous ») si $k \geq 2$) (figure 2.5) (voir définition 4.13 pour le genre d'une surface).

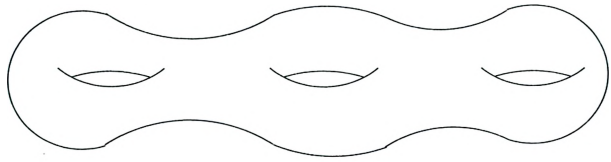


FIGURE 2.5. Tore T_k, $k = 3$.

2.2.3. Premiers exemples de sous-variétés

Nous donnerons, grâce à l'introduction de la notion de valeur régulière d'une application, dans la section 2.3, un procédé assez systématique pour la construction de sous-variétés. Cependant on peut examiner immédiatement quelques exemples.

1. Un ouvert $V \subset U$ est une sous-variété de codimension 0 de U.

2. Un ensemble de points isolés de U est une sous-variété de dimension 0 de U.

3. Tout sous-espace affine de dimension p dans $\mathbb{R}^n$ est une sous-variété de dimension p de $\mathbb{R}^n$.

4. Le cercle trigonométrique $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ est une sous-variété de $\mathbb{C} \approx \mathbb{R}^2$ (on le note S^1 car c'est la *sphère de dimension* 1, que l'on peut définir par $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$). Pour établir que S^1 est une sous-variété de $\mathbb{R}^2$, on donnera plus loin un argument général basé sur la notion de valeur régulière d'application. Remarquez que l'application $t \in \mathbb{R} \mapsto \exp(2i\pi t) \in S^1$ passe au quotient en un homéomorphisme de $T^1 = \mathbb{R}/\mathbb{Z}$ sur S^1. Il sera facile de vérifier que cette application quotient est un difféomorphisme de T^1 sur S^1, de classe $\mathcal{C}^\infty$ (et même analytique réelle). Donc S^1 est une réalisation en tant que sous-variété du cercle T^1. On confondra dorénavant S^1 et T^1 et l'on utilisera plutôt la notation S^1 qui est plus commune.

5. On appelle courbe simple toute sous-variété de dimension 1. Si elle est compacte et connexe, elle est difféomorphe à S^1. Le cercle trigonométrique est un exemple d'une telle courbe simple de $\mathbb{R}^2$. Une courbe simple non compacte et connexe est dite courbe unicursale. Elle est difféomorphe à $\mathbb{R}$ (par contraste, il existe une variété de dimension 1 *à base d'ouverts non dénombrable,* dite la *grande droite*, qui ne peut pas se réaliser comme courbe simple dans aucun espace euclidien).

6. Soit V une sous-variété de $\mathbb{R}^n$, de dimension d et de classe $\mathcal{C}^p$, et W une sous-variété de $\mathbb{R}^m$, de dimension e et de classe $\mathcal{C}^p$. Alors $V \times W$ est une sous-variété de $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$, de dimension $d+e$ et de classe $\mathcal{C}^p$.

7. Soit W un ouvert de $\mathbb{R}^p$ et φ une application de classe $\mathcal{C}^s$ de W à valeurs dans $\mathbb{R}^q$. Alors le graphe de φ :

$$\text{Graphe}(\varphi) = \{(x, \varphi(x)) | x \in W\},$$

est une sous-variété fermée de classe $\mathcal{C}^s$ de l'ouvert $U = W \times \mathbb{R}^q$ de $\mathbb{R}^{p+q}$. En effet, $\Phi(x,y) = (x, y - \varphi(x))$ est un difféomorphisme de classe $\mathcal{C}^s$, de l'ouvert U sur lui-même qui envoie Graphe(φ) sur $\mathbb{R}^p \times \{0\}$.

À titre d'exemple, considérons la fonction $f : x \in \mathbb{R} \mapsto f(x) \in \mathbb{R}$ définie par : $f(x) = x^{s+2} \sin(\frac{1}{x^2})$ pour $x \neq 0$ et $f(0) = 0$ $(s \in \mathbb{N})$. Cette fonction est de classe $\mathcal{C}^s$, mais n'est pas de classe $\mathcal{C}^{s+1}$. Son graphe est une sous-variété de $\mathbb{R}^2$ qui est aussi de classe $\mathcal{C}^s$ mais non de classe $\mathcal{C}^{s+1}$.

2.2.4. Espace tangent en un point d'une sous-variété

On sera intéressé à l'étude des courbes tracées sur une sous-variété V, par exemple aux trajectoires d'un champ de vecteurs tangent à cette sous-variété. Pour cela, il est utile de considérer en chaque point $x \in V$ l'ensemble des vecteurs tangents aux différentes courbes sur V, passant par x.

Définition 2.7. Soit x un point d'une sous-variété V d'un ouvert $U \subset \mathbb{R}^n$. Un vecteur tangent en x à V est un vecteur $u \in \mathbb{R}^n$ de la forme $u = \frac{dc}{dt}(0)$, où $c : t \in I \mapsto c(t) \in U$ est un arc de courbe défini sur un intervalle ouvert I au voisinage de $0 \in \mathbb{R}$, différentiable en $t = 0$ et contenu dans (c'est-à-dire tracé sur) V (pour tout $t \in I$, $c(t) \in V$). On appelle *espace tangent en x à V*, l'ensemble T_xV de tous les vecteurs tangents en x à V.

Proposition 2.1. *Soit V une sous-variété de l'ouvert $U \subset \mathbb{R}^n$ de classe $\mathcal{C}^1$ et de dimension p. Pour tout $x \in V$, l'espace tangent T_xV est un sous-espace vectoriel de $\mathbb{R}^n$ de dimension p.*

Démonstration. On considère un difféomorphisme de trivialisation $W \mapsto \mathbb{R}^p \times \mathbb{R}^{n-p}$ avec $x \in W \cap V$, $\Phi(x) = (0,0)$ et $\Phi(W \cap V) \subset \mathbb{R}^p \times \{0\}$.

Vérifions que $d\Phi(x)[T_xV] = \mathbb{R}^p \times \{0\}$. En effet, si $c(t)$ est une courbe passant par x comme dans la définition ci-dessus, $\Phi \circ c(t) \in \mathbb{R}^p \times \{0\}$ pour tout t assez proche de $0 \in \mathbb{R}$, d'où il résulte que

$$\frac{d}{dt}[\Phi \circ c](0) = d\Phi(x)(\frac{dc}{dt}(0)) = d\Phi(x)(u) \in \mathbb{R}^p \times \{0\}.$$

Inversement, si $w \in \mathbb{R}^p$ et si on considère la courbe $c(t) = \Phi^{-1}(tw)$, il vient : $d\Phi(x)\left(\frac{dc}{dt}(0)\right) = w$. ∎

Remarque 2.3. Pour garder une trace du point x, on a coutume de dessiner l'espace affine $\{x\} + T_xV$, espace affine qui est souvent lui-même appelé à tort « espace tangent ». L'intérêt de cet espace affine est qu'il est la meilleure approximation au premier ordre de la surface au point x.

Exemple d'espace tangent. Si V est le graphe d'une application $\varphi : W \mapsto \mathbb{R}^q$ comme plus haut, et $x \in W$, le plan tangent au point $(x, \phi(x)) \in V$ est égal à

$$T_{(x,\varphi(x))}V = \{(w, d\varphi(x)[w]) | w \in \mathbb{R}^p\}.$$

Si $\{e_i\}, i = 1, \ldots, p$, est la base canonique de $\mathbb{R}^p$, $T_{(x,\varphi(x))}V$ est engendré par le système de vecteurs : $\{(e_i, d\varphi(x)[e_i])\}_{i=1,\ldots,p}$ de $\mathbb{R}^p \times \mathbb{R}^{n-p}$. Par exemple, pour le graphe d'une fonction numérique ($p = q = 1$), la droite tangente (affine) en $(x, \varphi(x))$ est dirigée par le vecteur $(1, \varphi'(x))^T$, autrement dit, c'est la droite par $(x, \varphi(x))$ de pente $\varphi'(x)$.

2.3. Valeur régulière d'application différentiable

2.3.1. Équation cartésienne d'une sous-variété

Définition 2.8. Soit U un ouvert de $\mathbb{R}^n$ et $f : U \mapsto \mathbb{R}^{n-p}$ une application différentiable (on suppose que $0 \leq p < n$).

1. Un point $x \in U$ est un point régulier de f si et seulement si : rang $df(x) = n - p$. Le rang est alors maximum et $df(x)$ est une application linéaire surjective.

2. Un $y \in \mathbb{R}^{n-p}$ est une valeur régulière si et seulement si tout $x \in f^{-1}(y)$ (c'est-à-dire tout x tel que $f(x) = y$) est un point régulier de f.

Un point (ou bien une valeur) non régulier(ère) est dit singulier(ère).

Remarque 2.4. Remarquez que tout $y \notin f(U)$ est qualifié de valeur régulière. Autrement dit, toute valeur non atteinte est régulière ! L'image par f de l'ensemble des points singuliers s'identifie à l'ensemble des valeurs singulières. Par contre, la

contre-image de l'ensemble des valeurs singulières contient l'ensemble des points singuliers, mais peut être strictement plus large (considérez par exemple la fonction $f : x \mapsto x - x^3$) (figure 2.6).

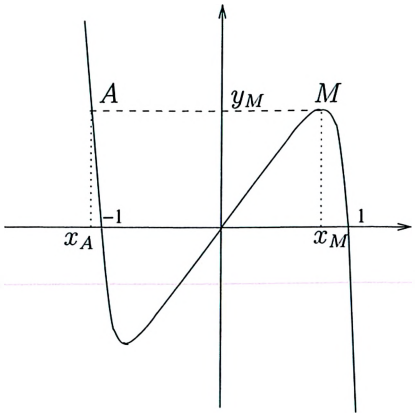


FIGURE 2.6. Graphe de f. L'ordonnée y_M du point M de coordonnées ($x_M = \sqrt{3}/3, y_M = 2\sqrt{3}/9$) est une valeur singulière de f. La contre-image de y_M est constituée, sur l'axe des x, des points d'abscisses x_M et x_A des points M et A (x_M (resp. x_A) est un point singulier (resp. régulier)).

Le théorème des fonctions implicites permet d'établir un lien entre les notions de sous-variété et de valeur régulière.

Théorème 2.1. *Soit U un ouvert de $\mathbb{R}^n$ et $y \in \mathbb{R}^{n-p}$ une valeur régulière de $f : U \mapsto \mathbb{R}^{n-p}$, application de classe C^k avec $k \geq 1$. Alors, ou bien $f^{-1}(y) = \emptyset$, ou bien $f^{-1}(y)$ est une sous-variété fermée de U de dimension p et de classe C^k.*

Démonstration. La preuve est une conséquence directe du théorème des fonctions implicites et de la définition de sous-variété. Soit $y = c$ une valeur régulière et supposons que $V = f^{-1}(c)$ soit non vide. Considérons x_0 un point quelconque dans V. Par hypothèse, la différentielle $df(x_0)$ est de rang maximum et f est de classe C^k avec $k \geq 1$. Il existe un isomorphisme linéaire de $\mathbb{R}^n$ avec $\mathbb{R}^p \times \mathbb{R}^{n-p}$ tel que, au point $x_0 = (a, b)$, l'application $df(x_0)|_{\{0\}\times\mathbb{R}^{n-p}}$ soit un isomorphisme linéaire de $\mathbb{R}^{n-p}$. On peut appliquer le théorème des fonctions implicites au point x_0 : il existe un voisinage ouvert W de (a, c), un voisinage ouvert $T = T_1 \times T_2$ de (a, b) dans $\mathbb{R}^p \times \mathbb{R}^{n-p}$, et un difféomorphisme H de classe C^k de T sur $W = H(T)$ tels que $H(a, c) = (a, b) = x_0$ et $f \circ H(u_1, u_2) = u_2$ pour tout $(u_1, u_2) \in T_1 \times T_2$. La paire $(W, \Phi = H^{-1})$ est un difféomorphisme de trivialisation pour V au voisinage du point x_0. D'après la définition 2.6, W est donc une sous-variété. Cette sous-variété est fermée car f étant de classe C^k est *a fortiori* continue. ∎

Définition 2.9. Si c est une valeur régulière atteinte de $f : U \mapsto \mathbb{R}^{n-p}$, la sous-variété $V = f^{-1}(c)$ est formée des points de U solutions de l'équation : $\{f(x) = c\}$. Cette équation est appelée équation cartésienne de V.

Exemples

1. La sphère unité S^2 dans $\mathbb{R}^3$ a pour équation cartésienne :

$$S^2 = \{(x, y, z) \in \mathbb{R}^3 | f(x, y, z) = x^2 + y^2 + z^2 = 1\}.$$

La différentielle de f en un point (x, y, z) quelconque est égale à

$$df(x, y, z) = 2(xdx + ydy + zdz).$$

L'unique point singulier de f est l'origine. Comme $f(0, 0, 0) = 0$, la valeur 1 est régulière et atteinte car $f(x, y, z) \to \infty$ quand $||(x, y, z)|| \to \infty$ et donc f doit prendre, sur tout rayon, la *valeur intermédiaire* 1. Il s'ensuit que la sphère S^2 est une surface de $\mathbb{R}^3$ (sous-variété de dimension 2). En fait, sans utiliser le théorème des fonctions implicites, il est facile de trouver un atlas de cartes de trivialisation pour cette structure différentiable et donc, d'après le premier item de la remarque 2.2, prouver ainsi que la sphère unité S^2 est une sous-variété de dimension 2. Il suffit de remarquer en toute généralité que, si un ouvert U sur une sous-variété V de classe $\mathcal{C}^\infty$ de R^n est l'image d'un graphe d'une application π de classe $\mathcal{C}^\infty$, $\pi : W \mapsto \mathbb{R}^q$ (avec W ouvert de $\mathbb{R}^p$ et $p + q = n$), alors (U, φ^{-1}) est une carte de V. Par exemple $U = \{(x, y, z) \in S^2 | z > 0\}$ est l'image du graphe $\pi(x, y) = \sqrt{1 - x^2 - y^2}$ au-dessus de $\Omega = \{x^2 + y^2 < 1\}$. En considérant l'application $-\pi$ et les applications similaires associées aux autres axes de coordonnées de $\mathbb{R}^3$, on définit un atlas sur S^2. Toutes les structures de variété $\mathcal{C}^\infty$ que l'on a construites jusqu'ici sur la sphère coïncident, comme on peut le vérifier facilement. En fait la sphère S^2 n'a qu'une seule struture de variété de classe $\mathcal{C}^\infty$.

2. On peut définir de la même manière la sphère unité S^n dans $\mathbb{R}^{n+1}$ pour tout $n \in \mathbb{N}$. La sphère S^0 est l'ensemble $\{-1, 1\}$ et la sphère S^1 est le cercle trigonométrique. Les sphères S^n et les tores T^n forment deux familles de variétés, définies en toute dimension $n \geq 1$ qui ne coïncident que pour $n = 1$: $S^1 = T^1$, mais S^n n'est pas difféomorphe à T^n (ni même homéomorphe) dès que $n \geq 2$. Ceci est un résultat classique mais non trivial de la topologie algébrique (cf. par exemple [9]).

Lorsqu'une sous-variété est définie par une équation cartésienne, il est facile de déterminer ses plans tangents par linéarisation de son équation cartésienne en chaque point :

Proposition 2.2. *Soit $f : U \mapsto \mathbb{R}^{n-p}$ une application de classe $\mathcal{C}^k$, $k \geq 1$, et α une valeur régulière atteinte. Si $x \in V = f^{-1}(\alpha)$, on a*

$$T_x V = \text{Ker } df(x) = \{u \in \mathbb{R}^n | df(x)[u] = 0\}.$$

Démonstration. On sait que $\dim(T_x V) = \dim(\text{Ker } df(x)) = p$. Il suffit donc de montrer que $T_x V \subset \text{Ker } df(x)$. Considérons pour ce faire un arc de courbe quelconque $c(t)$, comme dans la définition 2.7. On a $f \circ c(t) \equiv \alpha$, d'où il suit que $df(x)(\frac{dc}{dt}(0)) = 0$ et donc que : $\frac{dc}{dt}(0) \in \text{Ker } df(x)$. Ce raisonnement montre que $T_x V \subset \text{Ker } df(x)$. ∎

Exemples

1. Pour la sphère S^2 définie par $f(x, y, z) = x^2 + y^2 + z^2$, on a :

$$T(x, y, z)S^2 = \text{Ker } df((x, y, z)) = \{(U, V, W) \in \mathbb{R}^3 | xU + yV + zW = 0\},$$

ce qui s'interprète comme le 2-plan perpendiculaire au vecteur (x, y, z). On voit sur cet exemple que le plan tangent en $m = (x, y, z)$ est orthogonal au vecteur gradient $\nabla f(m) = \left(\frac{\partial f}{\partial x}(m), \frac{\partial f}{\partial y}(m), \frac{\partial f}{\partial z}(m)\right)$. Ce résultat est évidemment vrai pour toute fonction convenable et en toute dimension.

2. Soit $f : \mathbb{R}^3 \mapsto \mathbb{R}^2$ donnée par $f(x, y, z) = (X, Y)$ avec $X = x^2 + y^2 - z^2$ et $Y = xyz$. Les différentielles des deux composantes de f sont égales à $dX(x, y, z) = 2xdx + 2ydy - 2zdz$ et $dY(x, y, z) = yzdx + xzdy + xydz$. L'ensemble des points singuliers est formé des trois axes de coordonnées et des quatre droites $\{x = \pm y = \pm z\}$. L'ensemble des valeurs singulières est donc formé des points du demi-axe $\{X \geq 0, Y = 0\}$ et de la courbe d'équation $\{X = x^2, Y = x^3\}$. Pour toute valeur régulière (α, β), les équations $\{X(x, y, z) = \alpha, Y(x, y, z) = \beta\}$ définissent une courbe simple $C_{(\alpha,\beta)}$ dans $\mathbb{R}^3$. En chaque point $(x, y, z) \in C_{(\alpha,\beta)}$, on a un espace tangent (ici une droite tangente d'équation cartésienne $\{(U, V, W) \in \mathbb{R}^3 | xU + yV + zW = 0, yzU + xzV + xyW = 0\}$).

2.3.2. Existe-t-il beaucoup de valeurs régulières ?

Une application suffisamment différentiable possède toujours beaucoup de valeurs régulières comme l'a montré M. Sard (voir une preuve dans [1] par exemple) :

***Théorème 2.2* (Théorème de Sard).** *Soit $f : \mathbb{R}^n \mapsto \mathbb{R}^{n-p}$ une application de classe $\mathcal{C}^\infty$. Alors l'ensemble des valeurs singulières de f est de mesure de Lebesgue nulle dans $\mathbb{R}^{n-p}$ (on dit aussi que c'est un ensemble négligeable).*

Rappelons qu'un ensemble $A \subset \mathbb{R}^q$ est dit négligeable dans $\mathbb{R}^q$ si, pour tout $\varepsilon > 0$, on peut trouver un recouvrement dénombrable de A par des boules B_i, $i \in \mathbb{N}$, telles que $\Sigma_i \text{vol}(B_i) \leq \varepsilon$. (Ici, $\text{vol}(B_i)$ désigne le volume euclidien de la boule B_i, autrement dit sa mesure de Lebesgue au sens de $\mathbb{R}^q$). Un ensemble est négligeable dans $\mathbb{R}^q$ si et seulement s'il est de mesure de Lebesgue nulle dans $\mathbb{R}^q$. Par exemple : $\mathbb{R}^{q-1} \subset \mathbb{R}^q$ est de mesure nulle dans $\mathbb{R}^q$. On dit qu'une propriété est vraie *presque partout* s'il existe $A \subset \mathbb{R}^q$ avec $\text{mes}(A) = 0$ tel que la propriété soit vraie en tout point $x \in \mathbb{R}^q - A$.

Remarque 2.5. Si $A \subset \mathbb{R}^q$ est de mesure nulle, $\mathbb{R}^q - A$ est dense dans $\mathbb{R}^q$. (En effet pour tout G ouvert non vide de $\mathbb{R}^q$, $G \cap (\mathbb{R}^q - A) \neq \emptyset$; sinon il existe un tel ouvert non vide tel que $G \cap (\mathbb{R}^q - A) = \emptyset$, alors $G \subset A$ et ceci est impossible car A est de mesure nulle.)

Le théorème de Sard est difficile à démontrer en toute généralité. Pour donner une idée de sa démonstration, nous allons en prouver un cas particulier (preuve inspirée de [2]).

Proposition 2.3. *Si $f : \mathbb{R} \mapsto \mathbb{R}$ est une fonction de classe $\mathcal{C}^2$, alors l'ensemble des valeurs singulières est négligeable.*

Démonstration. Il suffit de montrer le résultat pour $f|_{[a,b]}$ où $[a, b]$ est un intervalle borné quelconque. En effet si $f : \mathbb{R} \mapsto \mathbb{R}$ est une fonction sur $\mathbb{R}$, si Σ est l'ensemble de ses points singuliers et Σ_n est l'ensemble des points singuliers de la restriction $f|_{[-n,n]}$, alors $\Sigma_n = \Sigma \cap [-n, n]$ d'où $\Sigma = \cup_n \Sigma_n$. Il en résulte que $f(\Sigma) = \cup_n f(\Sigma_n)$. Si chaque $f(\Sigma_n)$ est négligeable, $f(\Sigma)$ est aussi négligeable comme réunion dénombrable d'ensembles négligeables.

Considérons donc une fonction f de classe $\mathcal{C}^2$, définie sur un intervalle borné $[a, b]$. Comme la dérivée seconde f'' est continue, il existe un réel $M > 0$ tel que $|f''(x)| \leq M$ pour tout $x \in [a, b]$ (car l'intervalle est compact). Soit $\Sigma = \{x \in [a, b] | f'(x) = 0\}$ l'ensemble des points singuliers et $f(\Sigma)$ l'ensemble des valeurs singulières. Soit $n \geq 1$ un entier quelconque. Subdivisons l'intervalle $[a, b]$ en n sous-intervalles égaux : $I_1, \ldots, I_n$ $(I_i = [a + \frac{i-1}{n}(b - a), a + \frac{i}{n}(b - a)])$.

Pour n'importe quelle partie bornée K de $\mathbb{R}$, on appelle $\text{long}(K)$ la longueur du plus petit intervalle contenant K. Nous allons montrer que si $\Sigma \cap I_i \neq \emptyset$, alors $\text{long}(f(I_i)) \sim \frac{1}{n^2}$ (intuitivement, les intervalles I_i contenant un point singulier ont une « image pliée » qui est, de ce fait, de longueur équivalente au carré de la longueur de I_i). Pour cela, choisissons un point singulier $x_0 \in \Sigma \cap I_i$ et un x

quelconque dans I_i. La formule de Mac-Laurin à l'ordre 2 appliquée à f entre les points x_0 et x donne

$$f(x) - f(x_0) = f'(x_0)(x - x_0) + \frac{f''(\theta)}{2}(x - x_0)^2$$

pour un $\theta \in I_i$. Comme $f'(x_0) = 0$, que $|f''|$ est bornée par M et que $|x - x_0| \leq \frac{|b-a|}{n}$ il vient :

$$|f(x) - f(x_0)| \leq \frac{M|b-a|^2}{2}\frac{1}{n^2}.$$

Ceci implique évidemment que $\text{long}(f(I_i)) \leq \frac{M|b-a|^2}{2}\frac{1}{n^2}$. Comme il a été annoncé plus haut, on constate que l'image est « pliée » quadratiquement autour d'une valeur singulière (figure 2.7).

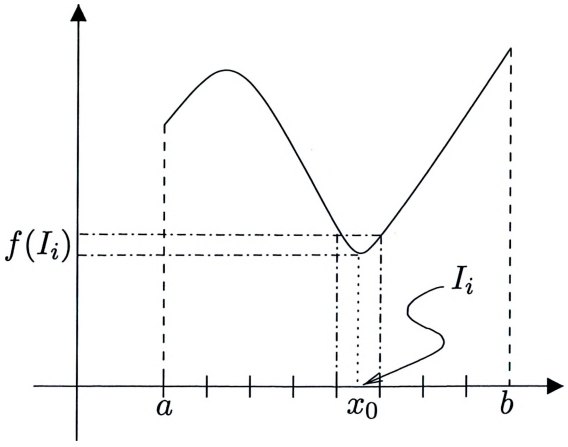


FIGURE 2.7

Soit $I_{i_1}, \ldots, I_{i_k}$ pour $k \leq n$ les sous-intervalles intersectant Σ. L'ensemble $f(\Sigma)$ des valeurs singulières est recouvert par $f(I_{i_1}), \ldots, f(I_{i_k})$, qui sont aussi des intervalles (des boules de $\mathbb{R}$!), et on a

$$\Sigma_{j=1}^{k}\text{long}(f(I_{i_j})) \leq \frac{M|b-a|^2}{2}\frac{1}{n^2} \cdot n = \frac{M|b-a|^2}{2}\frac{1}{n}.$$

En faisant tendre n vers l'infini, on obtient que $f(\Sigma)$ est négligeable. ∎

Remarque 2.6.

1. Le théorème de Sard nous permet de dire que l'ensemble des valeurs singulières est de mesure nulle, c'est-à-dire que f a toujours très peu de valeurs

singulières. Par contre, les points singuliers peuvent être très nombreux. Par exemple, si l'on considère la fonction f à valeur constante, $f(x) \equiv c \in \mathbb{R}^{n-p}$ pour tout $x \in \mathbb{R}^n$, tous les $x \in \mathbb{R}^n$ sont points singuliers alors que c est l'unique valeur singulière.

2. Le théorème de Sard n'a d'intérêt que si f a des valeurs régulières atteintes. Par exemple, si f est la fonction à valeur constante $c \in \mathbb{R}^{n-p}$, les valeurs régulières sont exactement les valeurs non atteintes, c'est-à-dire l'ensemble $\mathbb{R}^{n-p} - \{c\}$.

3. La démonstration du théorème de Sard pour toute paire de nombres (n,p), $0 \leq p < n$, est similaire à celle qui est donnée ci-dessus pour $(1,0)$. On montre que le fait que l'application f « contracte » fortement (ou « plie ») l'espace $\mathbb{R}^n$ au voisinage de ses points singuliers a pour conséquence que l'ensemble des valeurs critiques est de mesure nulle dans $\mathbb{R}^{n-p}$. La difficulté de la preuve générale tient au fait que les points critiques peuvent être plus ou moins dégénérés.

4. Il n'est pas nécessaire que l'application soit de classe $\mathcal{C}^\infty$. En fait, pour toute paire (n,p) comme ci-dessus, il existe un entier minimal $k(n,p)$ tel que le théorème de Sard soit valide pour les applications de classe $\mathcal{C}^{k(n,p)}$ (voir [1] pour une démonstration du théorème avec une évaluation de cette borne $k(n,p)$). Pour $n = p = 1$, la régularité $\mathcal{C}^2$ suffit pour le résultat.

5. On peut généraliser le théorème de Sard aux applications entre deux variétés (on dit qu'un sous-ensemble A d'une variété N de dimension q est négligeable dans N si pour toute carte (U,φ) de N l'ensemble $\varphi(A \cap U)$ est négligeable dans $\mathbb{R}^q$).

6. Si $f : \mathbb{R}^n \mapsto \mathbb{R}^{n+p}$ avec $p \geq 1$, le théorème reste valide. Dans ce cas, une valeur est régulière si et seulement si elle est non atteinte et on a précisément le résultat suivant : *si f est de classe $\mathcal{C}^1$, l'ensemble $f(\mathbb{R}^n)$ est négligeable dans $\mathbb{R}^{n+p}$.* Ce résultat est en fait beaucoup plus facile à établir que le théorème 2.2 (voir [4], corollaire p. 125 par exemple). De ce résultat suit par exemple que toute sous-variété d'un espace euclidien, de codimension ≥ 1, y est négligeable. Ce résultat est très important dans la théorie de l'intégration : il assure par exemple que le bord d'un domaine d'intégration raisonnable est de mesure nulle. Enfin il faut noter que le résultat serait faux sous la seule hypothèse de continuité : on sait construire par exemple des applications continues du cercle dans $\mathbb{R}^2$ dont l'image remplit toute la surface d'un carré ! (courbe de Peano).

7. Si f est une application polynomiale ou analytique, on a des résultats beaucoup plus précis. Par exemple si $P(x)$ est une fonction polynomiale de degré $n \geq 1$, le nombre de points singuliers (comptés avec leur multiplicité) et donc le nombre de valeurs singulières est borné par $n-1$ par le théorème de d'Alembert. Si f est une fonction holomorphe sur $\mathbb{C}$, ses points singuliers sont isolés : ils forment un ensemble au plus dénombrable, ainsi que l'ensemble des valeurs singulières, qui est donc trivialement négligeable dans $\mathbb{C}$.

2.4. Compléments sur les variétés

2.4.1. Espace tangent à une variété

Dans ce paragraphe, nous allons définir l'espace tangent en chaque point d'une variété, ainsi que la différentielle d'une application différentiable entre deux variétés. Sauf mention expresse du contraire, et cela ne sera pas explicitement précisé, toute variété et application est supposée de classe $\mathcal{C}^\infty$. Pour commencer, nous allons définir ce que l'on appelle un vecteur tangent en un point d'une variété M.

Définition 2.10. Soit $m \in M$. Un arc différentiable sur M passant par m est une courbe différentiable $c : t \in I \mapsto M$, avec I un intervalle ouvert contenant $0 \in \mathbb{R}$, et $c(0) = m$. Soit $\mathcal{A}_m$ l'ensemble de tels arcs. On introduit la relation d'équivalence suivante entre ces arcs : $c \sim c'$ si et seulement s'il existe une carte (U, φ) avec $m \in U$ et des restrictions de c, c' à des intervalles I, I' telles que $c(I), c'(I') \subset U$ et $\frac{d(\varphi \circ c)}{dt}(0) = \frac{d(\varphi \circ c')}{dt}(0)$.

Un *vecteur tangent* en m à M est une classe d'équivalence $\dot{c}$ dans $\mathcal{A}_m$ pour cette relation d'équivalence. L'ensemble des vecteurs tangents en m à M est appelé *espace tangent* en m à M et est noté $T_m M$. On notera par TM l'union ensembliste $TM = \coprod_m T_m M$, ensemble que l'on appelle *fibré tangent* de M.

La définition précédente a pour principal mérite de donner une définition intrinsèque de la notion de vecteur tangent au point m de la variété M. Pour obtenir la structure vectorielle de $T_m M$ et travailler pratiquement avec les vecteurs tangents, on choisira une carte de M contenant le point m. Dans une telle carte, l'espace tangent $T_m M$ sera identifié à l'espace euclidien $\mathbb{R}^n$ et le vecteur $v \in T_m M$ sera tout simplement représenté par un vecteur de $\mathbb{R}^n$.

Proposition 2.4. *Soit $m \in M$ et (U, φ) une carte telle que $m \in U$. L'application $\Phi_m : \dot{c} \in T_m M \mapsto \frac{d(\varphi \circ c)}{dt}(0) \in \mathbb{R}^p$ est une bijection. Cette bijection induit sur $T_m M$ une structure d'espace vectoriel de dimension p, indépendante du choix de la carte.*

Démonstration. Le fait que l'application est une bijection suit de la définition de vecteur tangent en m à M, une fois que l'on a remarqué que tout vecteur $v \in \mathbb{R}^p$ est image de la classe d'équivalence de l'arc $c(t) = \varphi^{-1}(tv)$ défini sur un intervalle assez petit.

Soit une deuxième carte (U', φ') telle que $m \in U'$. Comme l'application linéaire $d(\varphi' \circ \varphi^{-1})(\varphi(m))$ est un isomorphisme linéaire de $\mathbb{R}^p$, les deux cartes induisent la même structure vectorielle sur T_mM. ∎

Une situation plus concrète est celle d'une sous-variété dans un ouvert euclidien. Rappelons que si M est une sous-variété d'un ouvert euclidien U, elle peut être considérée comme une variété (voir l'item 2 de la remarque 2.2). On a maintenant deux définitions de l'espace tangent T_mM pour $m \in M$: la définition 2.7 et la définition 2.10 de la présente section. *Il est clair que ces deux définitions coïncident* puisque, dans la première définition, on a défini les vecteurs de T_mM comme les vecteurs dérivés en $c(0) = m$ aux courbes $c(t)$ tracées sur M. Un tel vecteur dérivé s'identifie à la classe d'équivalence $\dot{c}$ de la définition 2.10.

L'avantage du point de vue des sous-variétés M d'un ouvert de $\mathbb{R}^N$ est qu'un vecteur tangent est alors simplement un vecteur de $\mathbb{R}^N$, et non pas une classe d'équivalence comme dans la définition 2.10. De ce point de vue, l'espace tangent T_mM s'identifie à un sous-espace vectoriel de $\mathbb{R}^N$. On a vu, par exemple, que dans le cas de la sphère S^2, l'espace tangent T_mS^2 est le sous-espace vectoriel de dimension 2 de $\mathbb{R}^3$, orthogonal au vecteur $m \in S^2 \subset \mathbb{R}^3$. D'autre part, ce point de vue n'est pas restrictif, puisque toute variété (à base dénombrable) peut être réalisée comme sous-variété, grâce au théorème de Whitney. Évidemment, si l'on veut prouver que ce que l'on définit ainsi sur une variété M est intrinsèque, on devra prouver l'indépendance des raisonnements par rapport au choix de la réalisation de M comme sous-variété. Cela peut être assez laborieux. C'est pourquoi on préfèrera, lorsque cela sera possible, travailler directement avec la notion de variété, sans choisir une réalisation au caractère arbitraire, comme sous-variété.

Considérons maintenant les applications différentiables entre variétés.

Définition 2.11. Soit M et N des variétés de dimension p et q respectivement et $f : M \mapsto N$ une application différentiable. On définit la différentielle de f en $m \in M$ par $df(m) : T_mM \mapsto T_{f(m)}N$ par la formule

$$df(m)[\dot{c}] = \overline{(f \circ c)}^{\cdot},$$

où $c(t)$ est un arc différentiable quelconque de M en m.

Ici encore, le principal intérêt de cette définition est d'être intrinsèque. Cette définition a aussi le mérite de mettre en relation la différentielle $df(m)$ avec le transport par f d'arcs de courbes passant par m, relation qui nous sera très utile lorsque l'on traitera des champs de vecteurs. Pour montrer que $df(m)$ est bien défini et travailler pratiquement avec cette notion de différentielle, on considérera des cartes de M et N, choisies respectivement au voisinage des points m et $f(m)$:

Proposition 2.5. *Soit (U, φ) et (V, ψ) des cartes de M et N respectivement, telles que $m \in U$ et $f(U) \subset V$. Alors nous avons*

$$df(m) = (\overline{\Phi})^{-1}_{f(m)} \circ d(\psi \circ f \circ \varphi^{-1})(\varphi(m)) \circ \Phi_m,$$

où $\Phi_m : T_mM \mapsto R^p$ et $\overline{\Phi}_{f(m)} : T_{f(m)}N \mapsto \mathbb{R}^q$ sont les isomorphismes définis dans la proposition 2.4.

La proposition précédente signifie que, dans les cartes choisies, la différentielle $df(m)$ s'identifie avec la différentielle usuelle $d(\psi \circ f \circ \varphi^{-1})(\varphi(m))$ qui est une application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^q$, si M et N sont de dimension p et q respectivement. Autrement dit, la différentielle $df(m)$ est représentée par une matrice jacobienne.

2.4.2. Plongement, immersion, submersion

Rappelons que le rang d'une application linéaire est la dimension de son image.

Définition 2.12. Soit f une application différentiable d'une variété M de dimension m dans une autre variété N de dimension n.

1. On dit que f est une *immersion* si en chaque point $x \in M$ le rang de f est égal à m.

2. On dit que f est une *submersion* si en chaque point $x \in M$ le rang de f est égal à n.

Le rang de f en x est le rang de l'application différentielle $df(x) : T_xM \mapsto T_{f(x)}N$. Pratiquement, on calcule ce rang en prenant des cartes. Evidemment, l'existence d'une immersion suppose que $m \leq n$ et celle d'une submersion, que $m \geq n$.

Les plongements sont des cas particuliers d'immersion, d'un grand intérêt :

Définition 2.13. Une application différentiable d'une variété M dans une variété N est un *plongement* si f est une immersion injective et est un homéomorphisme de M sur son image $f(M)$.

On peut généraliser sans difficulté la notion de sous-variété aux variétés générales en adoptant la même définition que dans le cas d'un ouvert de $\mathbb{R}^n$:

Définition 2.14. Soit M une variété de dimension m. On dit que $Q \subset M$ est une sous-variété de M de dimension q si pour chaque point $a \in Q$ il existe une carte (W, φ) de M, telle que $a \in W$ et $\varphi(W \cap Q)$ soit un ouvert de $\mathbb{R}^q \times \{0\} \subset \mathbb{R}^q \times \mathbb{R}^{m-q} = \mathbb{R}^m$. On dit que (W, φ) est une *trivialisation locale* de la sous-variété M.

Un exemple trivial de sous-variété de la sphère S^2 est donné par le cercle $\gamma = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 = 1,\ z = 0\}$. Remarquez que la notion de sous-variété est héréditaire : si L est une sous-variété de Q et Q une sous-variété de M, alors L est une sous-variété de M.

La notion de plongement est équivalente à celle de sous-variété. Plus exactement on a :

Lemme 2.2. *Si f est un plongement de M dans N, alors $f(M)$ est une sous-variété de N et f est un difféomorphisme sur son image (considérée comme variété). Inversement, si $M \subset N$ est une sous-variété de N, l'inclusion de M (considérée comme variété) dans N est un plongement.*

Démonstration. Supposons que f soit un plongement de M dans N. Soit a un point quelconque de M. L'hypothèse implique que l'on peut trouver une carte de N : $(\mathcal{W}, \varphi)$, telle que $f(a) \in W$ et telle que $W \cap f(M) = f(W)$, où W est un ouvert domaine d'une carte dans M. On peut alors choisir $(\mathcal{W}, \varphi)$ comme une trivialisation locale de $V = f(M)$. Cet argument appliqué en chaque point $a \in M$, montre que V est une sous-variété. L'application f est un homéomorphisme de M sur V qui est de plus différentiable et de rang égal à m en tout point : c'est donc un difféomorphisme de M sur V.

Inversement, si M est une sous-variété, il suit directement des définitions que l'inclusion de M dans N est un plongement. ∎

Le lemme a pour signification qu'un plongement peut être vu comme une simple reparamétrisation de la sous-variété, définie par un difféomorphisme de

cette sous-variété. Dans le cas d'une sous-variété M de $\mathbb{R}^n$, l'identification de T_mM avec un sous-espace de $\mathbb{R}^n$, comme on l'a fait dans la section 2, revient à identifier T_mM avec $df(m)[T_mM]$, si f désigne l'inclusion de M dans $\mathbb{R}^n$.

Comparons maintenant les notions d'immersion et de plongement. Une immersion n'est pas nécessairement injective. Par exemple, l'immersion de S^1 dans $\mathbb{R}^2$ définie par $x = \cos 2t$, $y = \sin 3t$ dont l'image est une courbe de Lissajoux, est une immersion avec des points doubles. Mais, même lorsque l'immersion est injective, il peut se faire que l'image de M soit notablement différente de la variété M elle-même, au sens par exemple que la topologie de M puisse être *plus fine* que la topologie induite par N sur l'image $f(M)$ (c'est-à-dire, avec plus d'ouverts). Par exemple, pour l'immersion injective de $\mathbb{R}$ dans $\mathbb{R}^2$ représentée dans la figure 2.8, chaque voisinage du point a coupe $f(\mathbb{R})$ en un ensemble dont la contre-image par f contient un intervalle de la forme $]\alpha, +\infty)$. On peut même trouver des immersions de $\mathbb{R}$ dans $\mathbb{R}^3$, dont l'image est dense sur une surface. On aura par exemple de telles pathologies avec certaines orbites de champs de vecteurs de $\mathbb{R}^3$.

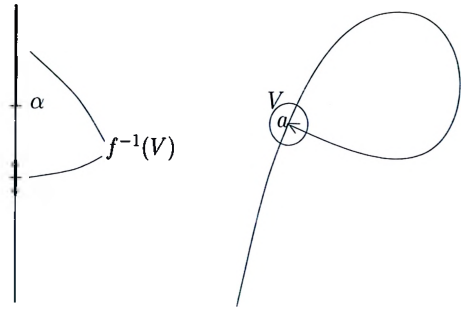


FIGURE 2.8

Pour un plongement, les pathologies signalées ci-dessus ne peuvent pas arriver. La proposition suivante, qui donne une condition suffisante pour que l'application f soit un homéomorphisme sur son image, est très utile pour vérifier qu'une immersion est un plongement :

Proposition 2.6. *Si f est une immersion* injective *d'une variété* compacte *dans une autre variété, alors f est un plongement.*

Démonstration. En effet, une application injective continue d'un compact dans un espace topologique séparé (un ouvert de $\mathbb{R}^n$ par exemple, et donc une variété en général) est un homéomorphisme. ∎

Exemple. L'application $\Phi : (\theta, \varphi) \mapsto (x, y, z)$ définie par :

$$x = (R + r\cos\theta)\cos\varphi, y = (R + r\cos\theta)\sin\varphi, z = r\sin\theta,$$

avec $R > r > 0$, est un plongement du tore $T^2 = \mathbb{R}^2/\mathbb{Z}^2$ dans $\mathbb{R}^3$. Cette application peut être vue comme une paramétrisation globale du tore de révolution $\Phi(T^2)$. C'est une surface de révolution autour de l'axe Oz.

Remarquons que le plongement peut être compliqué, par exemple un nœud de $\mathbb{R}^3$. Considérons l'application $f : t \in S^1 \mapsto f(t) \in \mathbb{R}^3$ suivante :

$$f(t) = ((2 + \cos 3t)\cos 2t, (2 + \cos 3t)\sin 2t, \sin 3t).$$

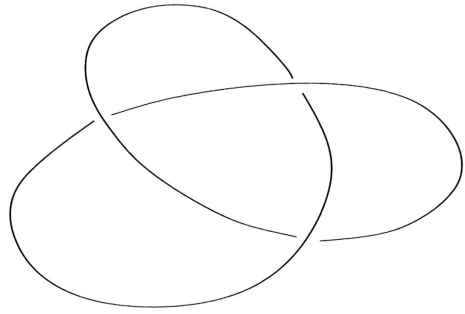


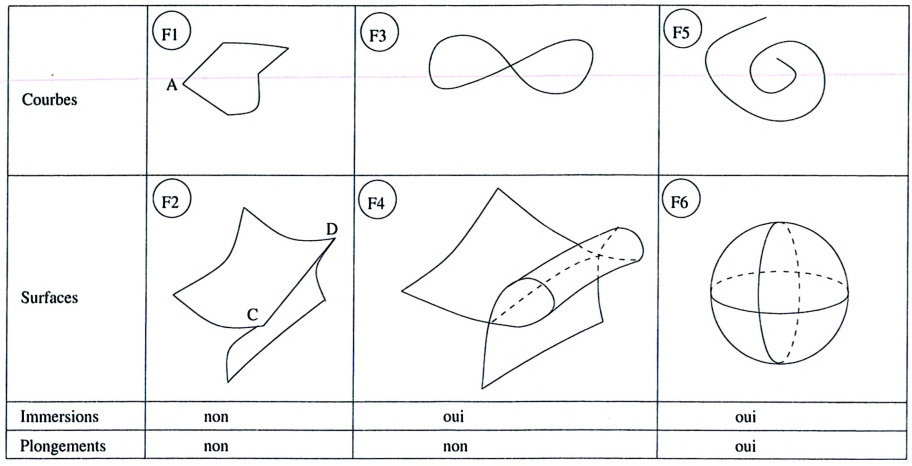
Figure 2.9. Nœud de trèfle dans $\mathbb{R}^3$.

On remarque que f est de rang un et est donc une immersion. De plus f est une application injective de S^1, et elle est donc un plongement dans $\mathbb{R}^3$ d'après la proposition 2.6. Le cercle plongé est *noué*, c'est-à-dire que l'on ne peut pas prolonger f en un plongement du disque $D = \{x^2 + y^2 = 1\}$. Ce fait semble assez intuitif en examinant la figure 2.9 : il semble que si l'on cherche à construire une immersion du disque, on doive nécessairement introduire des self-intersections et même des points singuliers. La non-existence d'un plongement est cependant très difficile à démontrer rigoureusement, et ne peut se faire que dans le cadre de la théorie des nœuds.

Remarque 2.7. Résumons les différences entre les notions d'immersion et de plongement. L'image d'une immersion peut posséder éventuellement des points de self-intersection (immersion non injective). Si la variété M que l'on immerge par une application φ n'est pas compacte, et même si l'immersion est injective, l'application φ n'est pas nécessairement un homéomorphisme entre M et son image $\varphi(M)$ (munie de la topologie induite). Heureusement, cette pathologie n'arrive pas si M est compacte et, dans ce cas, plongement est synonyme d'immersion injective.

Une sous-variété de dimension 1 est appelée tout simplement une courbe plongée, une sous-variété de dimension 2 est une surface plongée. La figure 2.10 montre

différents objets de dimension 1 ou 2. Dans l'encadré $F1$ (resp. $F2$), la courbe (resp. surface) n'est pas dérivable (resp. différentiable) au point A (resp. le long du segment CD), le rang de l'application en ce point (resp. segment CD) est égal à zéro (resp. un), nous n'avons donc pas une immersion. Dans les encadrés $F3$ et $F4$, nous avons respectivement un point double et une ligne de points doubles, nous n'avons donc pas de plongement. Dans les encadrés $F2$, $F4$ et $F6$, on montre des objets de dimension 2. Dans l'encadré $F2$, on montre une surface avec une ligne de singularité. Dans l'encadré $F4$, on montre une immersion d'une surface et dans $F6$ une sphère plongée.



Immersions	non	oui	oui
Plongements	non	non	oui

FIGURE 2.10

2.4.3. Distance sur une variété

Dans un espace topologique, il est très intéressant de pouvoir utiliser une distance définissant la topologie, lorsque cela est possible, c'est-à-dire lorsque l'espace est métrisable. C'est justement le cas pour les variétés, comme on va le voir maintenant. En fait, sur une variété M, on peut définir des distances très particulières associées aux *métriques riemanniennes*. On suppose que M est une variété de dimension n et de classe C^∞.

Définition 2.15. Une *métrique riemannienne* sur une variété M est la donnée d'un produit scalaire $\langle \cdot, \cdot \rangle_x$ sur chaque espace tangent T_xM. On suppose de plus que sur chaque carte (U, φ), le produit scalaire est une expression :

$$\langle \tilde{u}, \tilde{v} \rangle_m = \sum_{i,j=1}^{n} g_{ij}(\varphi(m)) u_i v_j,$$

où $\tilde{u} = (u_1, \ldots, u_n) = \Phi_m(u)$ et $\tilde{v} = (v_1, \ldots, v_n) = \Phi_m(v)$ respectivement, sont les représentants dans $\mathbb{R}^n$ des vecteurs tangents $u, v \in T_mM$, comme dans la proposition 2.4 ci-dessus. Dire que $\langle \cdot, \cdot \rangle_m$ est un produit scalaire est équivalent à dire que la forme quadratique $\sum g_{ij}(\varphi(m))u_iv_j$ est définie positive pour tout $m \in U$. Enfin, on suppose que les fonctions $\tilde{x} \mapsto g_{ij}(\tilde{x})$ sont $\mathcal{C}^\infty$ sur $\Omega = \varphi(U)$, pour tout $i, j \in \{1, \ldots, n\}$.

Par exemple le produit scalaire euclidien définit une métrique riemannienne constante sur $\mathbb{R}^n$ avec $g_{ij}(m) = \delta_{ij}$ (symboles de Kronecker).

Une métrique riemannienne permet de définir une norme sur chaque espace tangent, par $||u|| = \sqrt{\langle u, u \rangle_m}$ pour $u \in T_mM$. Si maintenant $\gamma : t \in [0,1] \mapsto M$ est un chemin de classe $\mathcal{C}^1$ dans M, on définit sa longueur par

$$\text{long}(\gamma) = \int_0^1 \left|\left| \frac{d\gamma}{dt}(t) \right|\right| dt.$$

Remarquez que l'hypothèse de différentiabilité faite sur le chemin γ implique que la fonction $t \mapsto ||\frac{d\gamma}{dt}(t)||$ est continue, et donc que l'intégrale ci-dessus existe au sens de Riemann.

On peut maintenant définir une distance sur M en posant, pour toute paire de points $x, y \in M$:

$$\text{dist}(x, y) = \text{Inf}_\gamma\{\text{long}(\gamma)\},$$

où γ est un chemin quelconque de classe $\mathcal{C}^1$ entre x et y, c'est-à-dire tel que $\gamma(0) = x$ et $\gamma(1) = y$. La fonction distance est continue sur $M \times M$. Il en résulte évidemment que la topologie définie par la distance est la topologie de la variété vue comme espace topologique (en fait, les boules fermées de rayon assez petit sont difféomorphes à des boules fermées euclidiennes). On utilisera une telle distance chaque fois que l'on souhaitera faire une démontration sur M à l'aide de suites et d'arguments du type « δ, ε » (plutôt qu'une démonstration utilisant les ouverts et fermés de M). Évidemment, pour les démonstrations locales (c'est-à-dire au voisinage d'un point), il sera plus commode de choisir une carte pour se ramener à travailler dans $\mathbb{R}^n$, avec la norme euclidienne de $\mathbb{R}^n$, si on le souhaite. Évidemment, l'intérêt des métriques riemanniennes va bien au-delà de la simple possibilité de définir des distances sur M (voir [4] p. 421, par exemple, pour une introduction à la géométrie riemannienne).

L'existence d'une métrique riemannienne sur une variété M quelconque peut être déduite du théorème de plongement de Whitney. En effet, rappelons que ce théorème permet de considérer M comme sous-variété d'un espace euclidien $\mathbb{R}^N$. Soit $\langle \cdot, \cdot \rangle$ le produit scalaire euclidien de $\mathbb{R}^N$. On peut alors définir une métrique

riemannienne $\langle \cdot, \cdot \rangle_x$ sur M par simple restriction de ce produit scalaire, en posant pour tout $x \in M$ et $u, v \in T_x M$:

$$\langle u, v \rangle_x = \langle u, v \rangle$$

où, dans le membre de droite, on considère les deux vecteurs u, v comme des vecteurs de $\mathbb{R}^N$. On appelle cette métrique la *métrique riemannienne induite* par le plongement de M dans $\mathbb{R}^N$. La condition de différentiabilité de la définition est très facilement vérifiée. Considérons pour ce faire une trivialisation locale de la variété. Cette trivialisation donne, par restriction à M, une application de carte dont l'inverse est une paramétrisation locale, c'est-à-dire une immersion $x = \psi(\tilde{x})$ avec $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_n)$, d'un ouvert Ω de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^N$ (avec une image dans M évidemment). Pour calculer le coefficient $g_{ij}(\tilde{x})$ de la métrique riemannienne dans la carte, il suffit de se rappeler que pour la forme quadratique $Q(A, B) = \sum g_{ij} A_i B_j$, on a $g_{ij} = Q(e_i, e_j)$, où e_i, e_j sont les vecteurs de la base canonique d'indices i et j. Ici, cela donne l'expression :

$$g_{ij}(\tilde{x}) = \left\langle \frac{\partial \psi}{\partial \tilde{x}_i}(\tilde{x}), \frac{\partial \psi}{\partial \tilde{x}_j}(\tilde{x}) \right\rangle, \tag{2.4}$$

qui est manifestement une fonction $\mathcal{C}^\infty$ de $\tilde{x} \in \Omega$.

La construction que l'on vient de faire est encore valable, plus généralement, si l'on considère une immersion φ d'une variété M dans $\mathbb{R}^N$. Il suffit de poser :

$$\langle u, v \rangle_x = \langle d\varphi(x)[u], d\varphi(x)[v] \rangle$$

pour tout $x \in M$ et $u, v \in T_x M$, pour définir une métrique riemannienne sur M (on retrouve la formule donnée pour une sous-variété M, en prenant pour φ l'inclusion de M dans $\mathbb{R}^N$).

Remarquons que l'on aurait pu obtenir plus simplement une distance sur la sous-variété M de $\mathbb{R}^N$, en restreignant simplement la distance euclidienne. Si M est compacte, cette distance est équivalente à la distance associée à la métrique riemannienne induite (deux métriques sur un ensemble sont dites équivalentes si leur rapport reste compris entre deux constantes positives). Si M n'est pas compacte, les deux distances définissent la topologie de M mais ne sont pas équivalentes en général.

Il est possible de construire des métriques riemanniennes sur une variété M sans recourir à un plongement dans un espace euclidien, en recollant des métriques riemanniennes définies localement dans des cartes, à l'aide d'une partition de l'unité sur M (voir [12] par exemple).

2.4.4. Transversalité

Nous allons maintenant introduire, et expliquer brièvement, la notion de transversalité introduite par René Thom au début des années 1960, pour généraliser la notion de valeur régulière d'une application, et donner, de ce fait, une plus grande portée au théorème des fonctions implicites. Nous reviendrons plus tard, dans le tome 2, sur cette notion de transversalité pour montrer comment elle a permis à Thom d'exploiter le théorème de Sard pour introduire la notion de généricité.

Définition 2.16. Soit P une sous-variété de $\mathbb{R}^n$ de dimension p et f une application différentiable d'un ouvert U de $\mathbb{R}^m$ à valeurs dans $\mathbb{R}^n$. On dit que f est transverse en $x \in U$ à P si et seulement si $f(x) \notin P$, ou bien si :

$$f(x) \in P \quad \text{et} \quad df(x)[\mathbb{R}^m] + T_{f(x)}P = \mathbb{R}^n, \tag{2.5}$$

ce qui suppose que $p + m \geq n$ (voir figure 2.11). On dit que f est transverse à P si f est transverse à P en tout point $x \in U$. Remarquez qu'une façon triviale pour f d'être transverse à P est que $f(U)$ soit disjoint de P. (La notation $df(x)[\mathbb{R}^m]$ désigne le sous-espace vectoriel de $\mathbb{R}^n$, image de $\mathbb{R}^m$ par l'application linéaire $df(x)$.)

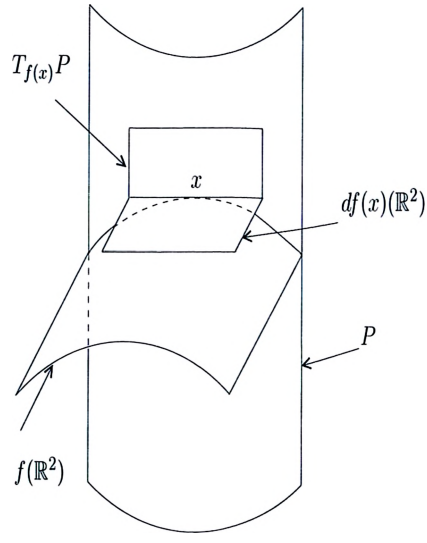


FIGURE 2.11. Application f transverse à une sous-variété P, dans le cas $p = 2$.

Remarque 2.8. Les espaces $df(x)[\mathbb{R}^m]$ et $T_{f(x)}P \subset \mathbb{R}^n$ sont transverses dans $\mathbb{R}^n$ au sens défini au début du paragraphe 2.2.1, d'où la terminologie. Par ailleurs, ces espaces ne sont pas nécessairement supplémentaires.

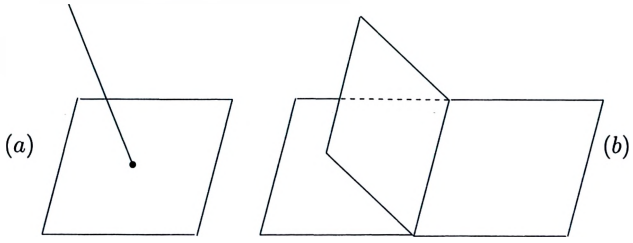
La figure 2.12 illustre cette notion de transversalité. Dans le cas de la figure 2.12 (a), où $n = 3, m = 2, p = 2$, on vérifie que :

$$\dim(T_{f(x)}P) = \dim(df(x)(\mathbb{R}^p)) = 2 \text{ et } \dim(T_{f(x)}P \cap df(x)(\mathbb{R}^p)) = 1.$$

Les espaces $T_{f(x)}P$ et $df(x)(\mathbb{R}^p)$ sont transverses mais non supplémentaires. Par contre dans le cas de la figure 2.12 (b) où $n = 3, m = 1, p = 2$, on vérifie que :

$$\dim(T_{f(x)}P) = 1,\ \dim(df(x)(\mathbb{R}^p)) = 2 \text{ et } (T_{f(x)}P \cap df(x)(\mathbb{R}^p)) = \{0\}.$$

Dans ce cas, les espaces $T_{f(x)}P$ et $df(x)(\mathbb{R}^p))$ sont supplémentaires.



Espaces supplémentaires. Espaces transverses.

FIGURE 2.12

Notation. *f transverse à P en x* sera noté : $f \overline{\pitchfork}_x P$ et *f transverse à P (en x pour tout $x \in U$)* sera noté : $f \overline{\pitchfork} P$.

Remarque 2.9. La notion de transversalité généralise la notion de valeur régulière d'une application : dire qu'une application f admet $\alpha \in \mathbb{R}^n$ comme valeur régulière est équivalent à dire que f est transverse au point $\{\alpha\}$, considéré comme sous-variété de dimension 0 de $\mathbb{R}$. En effet, pour tout point régulier $x \in f^{-1}(\alpha)$, le rang de $df(x)$ est maximum et égal à n, et donc : $df(x)[\mathbb{R}^m] = \mathbb{R}^n$. Dès lors, la relation (2.5) est satisfaite avec $P = \{\alpha\}$, sous-variété de $\mathbb{R}^n$ de dimension zéro pour laquelle $T_\alpha\{\alpha\} = \{0\}$.

Le théorème suivant généralise alors le théorème 2.1.

Théorème 2.3. *Soit P une sous-variété de $\mathbb{R}^n$ de classe $\mathcal{C}^k$, $1 \leq k \leq +\infty$, et f une application de même classe de différentiabilité, d'un ouvert U de $\mathbb{R}^m$ à valeurs dans $\mathbb{R}^n$. Supposons que f soit transverse à P et que $f(U) \cap P \neq \emptyset$. Alors $f^{-1}(P)$ est une sous-variété de U de classe $\mathcal{C}^k$. De plus la codimension de $f^{-1}(P)$ dans U est égale à celle de P dans $\mathbb{R}^n$ (autrement dit la codimension est préservée par contre-image). (Figure 2.13 avec $m = 2$, $n = 3$, $U = \mathbb{R}^2$ et $P = S^2$.)*

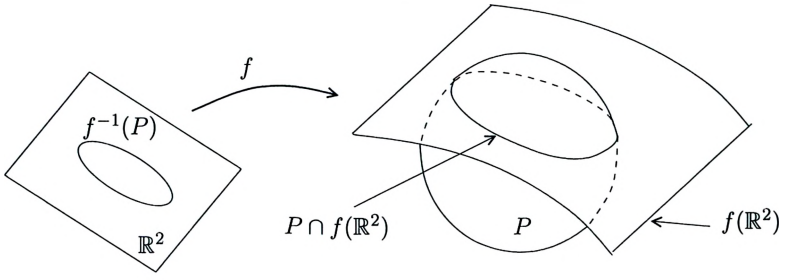


Figure 2.13

Démonstration. Soit $x \in U$ tel que $f(x) \in P$. Considérons un difféomorphisme de trivialisation (W, Φ) de P tel que $f(x) \in W$. On suppose que P est de codimension p. Par la définition 2.6, on a $\Phi(P \cap W) = \Phi(W) \cap (\mathbb{R}^{n-p} \times \{0\}) \subset \mathbb{R}^{n-p} \times \mathbb{R}^p$. Si π est la projection linéaire de $\mathbb{R}^{n-p} \times \mathbb{R}^p$ sur $\mathbb{R}^p$, l'inclusion ci-dessus est équivalente à $P \cap W = (\pi \circ \Phi)^{-1}(0)$: en effet si $\Phi(x) = (a, b) \in \mathbb{R}^{n-p} \times \{0\}$, cela signifie que $b = 0$, c'est-à-dire que $(a, b) \in \pi^{-1}(0)$ et donc que $x \in \Phi^{-1} \circ \pi^{-1}(0) = (\pi \circ \Phi)^{-1}(0)$.

En prenant la contre-image par f, on obtient que sur l'ouvert $W' = f^{-1}(W)$ de U, on a

$$W' \cap f^{-1}(P) = (\pi \circ \Phi \circ f)^{-1}(0).$$

L'hypothèse de transversalité implique que 0 est une valeur régulière de $\pi \circ \Phi \circ f$. Le théorème 2.1 implique donc que $W' \cap f^{-1}(P)$ est une sous-variété de codimension p de W'. L'ensemble $f^{-1}(P)$ étant une sous-variété de codimension p au voisinage de chacun de ses points, est globalement une sous-variété de U de codimension p. Comme toutes les applications utilisées sont de classe $\mathcal{C}^k$, la sous-variété $f^{-1}(P)$ est aussi de classe $\mathcal{C}^k$. ∎

Il sera très utile de pouvoir considérer une variété M à la place de l'ouvert U. On peut le faire aisément en travaillant dans des cartes.

Définition 2.17. Soit P une sous-variété de $\mathbb{R}^n$ et f une application différentiable d'une variété M de dimension m à valeurs dans $\mathbb{R}^n$. On dit que f est transverse en $x \in M$ à P si et seulement si $f(x) \notin P$, ou bien si

$$f(x) \in P \quad \text{et} \quad df(x)[T_xM] + T_{f(x)}P = \mathbb{R}^n. \tag{2.6}$$

On dit que f est transverse à P si f *est transverse* à P en tout point $x \in M$. Remarquez qu'une façon triviale pour f d'être transverse à P est que $f(M)$ soit disjoint de P.

Remarque 2.10. Soit (U, φ) une carte de M avec $x \in M$. La condition de transversalité (2.6) s'écrit de la façon suivante dans la carte (U, φ) :

$$f(x) \in P \quad \text{et} \quad d(f \circ \varphi^{-1}(\varphi(x))[\mathbb{R}^m] + T_{f(x)}P = \mathbb{R}^n. \tag{2.7}$$

Remarquez que l'application $f \circ \varphi^{-1}$ représente la fonction f dans la paramétrisation de U donnée par la carte.

En utilisant la notion de sous-variété d'une variété introduite dans la section précédente, le théorème 2.3 se généralise aisément :

Théorème 2.4. *Soit P une sous-variété de $\mathbb{R}^n$ de classe $\mathcal{C}^k$, $1 \leq k \leq +\infty$, et f une application de même classe de différentiabilité, d'une variété M à valeurs dans $\mathbb{R}^n$. Supposons que f soit transverse à P et que $f(M) \cap P \neq \emptyset$. Alors $f^{-1}(P)$ est une sous-variété de M de classe $\mathcal{C}^k$. De plus la codimension de $f^{-1}(P)$ dans M (c'est la différence entre la dimension de M et celle de $f^{-1}(P)$) est égale à celle de P dans $\mathbb{R}^n$.*

Comme dans le cas d'un espace euclidien considéré dans la section précédente, on définit facilement la notion de plongement d'une variété Q dans une variété M : une sous-variété peut toujours être considérée comme image d'un plongement d'une variété qui lui est difféomorphe (d'où parfois une certaine ambiguïté du langage). Soit P et Q deux sous-variétés d'un ouvert euclidien U de $\mathbb{R}^n$. On dit que P et Q sont transverses au point $x \in P \cap Q$, si

$$T_xP + T_xQ = \mathbb{R}^n.$$

Cela est équivalent à dire que l'inclusion de Q dans U est transverse au point x à P (ou inversement, que l'inclusion de P dans U est transverse à Q). On dira que les deux sous-variétés P et Q sont transverses si elles sont transverses en chacun de leur point d'intersection. Remarquez que si P et Q sont données par des équations

cartésiennes : $P = f^{-1}(\alpha)$, $Q = g^{-1}(\beta)$, avec α et β des valeurs régulières de $f : U \mapsto \mathbb{R}^s$ et $g : U \mapsto \mathbb{R}^\ell$, alors dire que P et Q sont transverses est équivalent à dire que (α, β) est valeur régulière de l'application $(f, g) : U \mapsto \mathbb{R}^s \times \mathbb{R}^\ell$. Il en résulte que s'il n'est pas vide, l'ensemble $P \cap Q = (f, g)^{-1}(\alpha, \beta)$ est aussi une sous-variété de U.

Plus généralement, on a le corollaire suivant du théorème 2.4 :

Corollaire 2.1. *Soit P et Q deux sous-variétés transverses dans un ouvert U de $\mathbb{R}^n$ d'intersection non vide. Alors leur intersection est une sous-variété de P, de Q, et aussi de U. De plus cette sous-variété $P \cap Q$ est de dimension égale à* $\dim P + \dim Q - n$.

Par exemple, soit le plan $H = \{z = 0\}$ et la sphère S^2. Ce sont deux sous-variétés de $\mathbb{R}^3$; le plan H est transverse à la sphère S^2. L'intersection de ces deux sous-variétés est le cercle $\gamma = \{x^2 + y^2 = 1, z = 0\}$. Ce cercle est de dimension $1 = \dim H + \dim S^2 - 3$.

3

POINTS SINGULIERS DE FONCTIONS

3.1. Dérivées partielles d'ordre supérieur

3.1.1. Définitions, notations et propriétés de base

Soit $f : U \mapsto \mathbb{R}$ une fonction définie sur un ouvert U de $\mathbb{R}^n$, $n \geq 1$. On a déjà défini les dérivées partielles du premier ordre, lorsqu'elles existent, par

$$\frac{\partial f}{\partial x_i}(x_0) = f'_{e_i}(x_0) = \lim_{t \to 0} \frac{f(x_0 + te_i) - f(x_0)}{t},$$

où $e_i = (0, \ldots, 1, \ldots, 0)^T$ est le i-ème vecteur de la base canonique de $\mathbb{R}^n$, dont la i-ème composante vaut 1 et les autres valent 0.

On définit, lorsque c'est possible, les dérivées partielles de tout ordre $k \in N$ par la formule de récurrence commençant avec $k = 2$:

$$\frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}(x_0) = \frac{\partial}{\partial x_{i_1}} \left[\frac{\partial^{k-1} f}{\partial x_{i_2} \ldots \partial x_{i_k}} \right] (x_0).$$

Comme conséquence du lemme de Schwarz relatif au cas particulier des dérivées secondes, on prouve, en toute généralité [6], le Théorèm 3.1.

Théorème 3.1. *Si f est une fonction de classe C^k, alors*

$$\frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{i_{\sigma(1)}} \ldots \partial x_{i_{\sigma(k)}}},$$

pour toute permutation σ de $\{1, \ldots, k\}$.

Cette symétrie permet de réordonner l'écriture des dérivées partielles. On peut commencer par les dérivées par rapport à la coordonnée x_1, lorsqu'il y en a, puis par rapport à x_2 et ainsi de suite. De cette façon, chaque dérivée partielle d'ordre k peut s'écrire :

$$\frac{\partial^k f}{\partial x_1^{j_1} \dots \partial x_n^{j_n}} \quad \text{avec} \quad j_1 + \dots + j_n = k.$$

Le symbole $\partial x_\ell^{j_\ell}$ est mis pour la répétition de j_ℓ dérivations par rapport à la coordonnée x_ℓ, avec la convention que $j_\ell = 0$ signifie qu'il n'y a pas de dérivation dans cette coordonnée. On peut écrire de façon encore plus condensée, en introduisant le multi-indice $j = (j_1, \dots, j_n) \in \mathbb{N}^n$ et sa longueur $|j| = j_1 + \dots + j_n$:

$$\frac{\partial^{|j|} f}{\partial x_1^{j_1} \dots \partial x_n^{j_n}} = \partial^j f.$$

On peut définir, lorsque c'est possible, la différentielle d'ordre k par la formule de récurrence :

$$d^k f(x) = d[d^{k-1} f](x),$$

à partir de la définition initiale $d^0 f = f$. Comme nous n'utiliserons pas cette notion dans la suite, nous n'essayerons pas de donner une forme plus explicite à la différentielle d'ordre k en général. Disons seulement que $d^k f(x)$ est une forme k-linéaire et symétrique dont les coefficients sont les dérivées partielles d'ordre k au point x.

En appliquant par récurrence la proposition 1.1, on voit que la fonction f est de classe $\mathcal{C}^k$ si et seulement si toutes ses dérivées partielles existent et sont continues jusqu'à l'ordre k, et qu'une fonction est de classe $\mathcal{C}^\infty$ si et seulement si ses dérivées partielles de tout ordre existent (dans ce dernier cas, les dérivées partielles sont différentiables en vertu de la proposition 1.1 et donc sont automatiquement continues). Évidemment, tout ceci s'étend à une application quelconque, en considérant ses fonctions composantes.

3.1.2. Approximation de f au voisinage d'un point

Définition 3.1. On dit que $f : U \mapsto \mathbb{R}$ admet en $x_0 \in U$ un développement limité d'ordre k, de partie principale le polynôme de degré $P^k(h)$ (polynôme dans les n composantes $h_1, \dots, h_n$ de $h \in \mathbb{R}^n$), si l'on peut écrire f sous la forme :

$$f(x_0 + h) = P^k(h) + o(||h||^n).$$

On peut développer le polynôme $P^k(h)$ en composantes homogènes :

$$P^k(h) = P_0 + P_1(h) + \ldots + P_k(h).$$

Évidemment $P_0 = f(0)$ et $P_1(h) = df(x_0)[h]$. L'existence d'un développement limité à l'ordre 1 en x_0 est donc équivalente à la différentiabilité de f en x_0.

Le développement limité est unique, s'il existe. Il donne l'approximation de la fonction f au point x_0, généralisant l'approximation affine fournie par la différentielle.

L'existence d'un développement limité est assurée pour les fonctions de classe de différentiabilité suffisante :

***Théorème 3.2* (Formule de Taylor avec reste de Young [6]).** *Soit f une fonction sur l'ouvert U de $\mathbb{R}^n$, de classe $\mathcal{C}^{k-1}$ telle que toutes les dérivées partielles d'ordre k existent en x_0, alors $f(x_0+h) = T_k(h,x_0) + o(||h||^k)$ où $T_k(h,x_0)$ est le polynôme de Taylor d'ordre k en x_0 :*

$$T_k(h,x_0) = f(x_0) + \sum_{|i|=1}^{k} \frac{1}{i!}\partial^i f(x_0) h^i. \tag{3.1}$$

Outre les notations $\partial^i f$ et $|i|$ introduites plus haut pour un multi-indice $i = (i_1,\ldots,i_n)$, on a posé ci-dessus : $h^i = h_1^{i_1}\ldots h_n^{i_n}$ et $i! = i_1!\ldots i_n!$ avec $0! = 1$. Chaque terme de la sommation est un monôme en $h_1,\ldots,h_n$.

Remarque 3.1. Considérons par exemple le terme homogène de degré 2 dans la formule (3.1). Dans la littérature, il apparaît souvent sous la forme

$$\frac{1}{2}d^2f(x_0)(h,h) = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f(x_0)}{\partial x_i \partial x_j} h_i h_j, \tag{3.2}$$

où $d^2f(x_0)$ désigne la différentielle de f d'ordre 2, calculée au point x_0. Cette différentielle d'ordre 2, appelée aussi *hessienne de f en x_0* lorsque ce point est singulier, est une forme bilinéaire qui s'applique à une paire de vecteurs $u, v \in \mathbb{R}^n$:

$$d^2f(x_0)(u,v) = \sum_{i,j=1}^{n} \frac{\partial^2 f(x_0)}{\partial x_i \partial x_j} u_i v_j.$$

Compte tenu du lemme de Schwarz, on peut regrouper les termes et retrouver l'écriture proposée dans la formule (3.1). Considérons par exemple le cas $n = 3$. Les choix possibles de i tels que $|i| = 2$, lorsque $n = 3$, sont

$$i = (2,0,0)\,,(0,2,0)\,,(0,0,2)\,,(1,1,0)\,,(0,1,1)\,,(0,1,1).$$

Notons, pour simplifier l'écriture, par (x, y, z) (au lieu de (x_1, x_2, x_3)) les trois composantes d'espace et par (h, k, l) les trois composantes du vecteur $h \in \mathbb{R}^3$. La formule (3.2) s'écrit (toutes les dérivées partielles étant prises au point x_0)

$$\sum_{|i|=2} \frac{1}{i!}\partial^i f(x_0)h^i = \frac{1}{2}(f''_{x^2}h^2 + f''_{y^2}k^2 + f''_{z^2}l^2) + f''_{xy}hk + f''_{yz}kl + f''_{xz}hl,$$

où l'on note $\frac{\partial^2 f}{\partial\alpha\partial\beta}$ par $f''_{\alpha\beta}$.

Plus généralement, on a les deux écritures équivalentes du terme homogène de degré l de la formule de Taylor :

$$\frac{1}{l!}\sum_{i_1,\ldots,i_l=1}^{n} \frac{\partial^l f(x_0)}{\partial x_{i_1}\ldots\partial x_{i_l}} h_{i_1}\ldots h_{i_l} = \sum_{|i|=l} \frac{1}{i!}\partial^i f(x_0)h^i, \tag{3.3}$$

où le terme de droite est obtenu à partir du terme de gauche par le regroupement des monômes qui sont égaux, compte tenu du lemme de Schwarz.

La formule de Taylor ci-dessus est celle qui est utilisée dans les calculs sur les développements limités. Sous des conditions légèrement plus générales sur f, on peut avoir une expression plus précise du reste $R(h, x_0)$, dont ci-dessus on retient seulement l'ordre : $o(||h||^k)$. Si la fonction f est de classe C^{k+1}, on a les estimations suivantes du reste [6] :

1. *Reste intégral* $R(h, x_0) = \int_0^1 \frac{(1-t)^k}{k!}(\sum_{|i|=k+1} \frac{(k+1)!}{i!}\partial^i f(x_0 + th)h^i)dt$.

2. *Reste de Lagrange* Posons : $M_{k+1}(y) = \mathrm{Sup}_{\{||h||=1\}}\left|\sum_{|i|=k+1} \frac{1}{i!}\partial^i f(y)h^i\right|$ pour tout $y \in U$. Comme f est C^{k+1}, $M_{k+1}(y)$ est continue en y, ce qui implique que $M_{k+1}(h, x_0) = \mathrm{Sup}\{M_{k+1}(y)|y \in [x_0, x_0 + h]\} < \infty$. Alors on a l'estimation suivante, dite par reste de Lagrange :

$$|R(h, x_0)| \leq M_{k+1}(h, x_0)||h||^{k+1}.$$

Le reste de Lagrange est utile pour les estimations précises des calculs numériques sur les fonctions (intégration numérique, etc.). Le reste intégral permet de relier directement le reste à la fonction f et retient les hypothèses de différentiabilité. Par exemple, il a pour conséquence immédiate que si une fonction f de classe C^∞ s'annule en $0 \in U$, alors elle s'écrit $f(x) = \sum_{i=1}^n x_i\varphi_i(x)$ pour des fonctions φ_i de classe C^∞.

3.2. Points singuliers d'une fonction sur un ouvert

Soit U un ouvert de $\mathbb{R}^n$ et une fonction $f : U \mapsto \mathbb{R}$. Lorsqu'elle est différentiable, on va maintenant étudier l'allure locale de cette fonction au voisinage de ses points singuliers.

3.2.1. Extremums

Définition 3.2. On dit que f admet un minimum en $a \in U$ si et seulement si $f(x) \geq f(a)$ pour tout $x \in U$. Le minimum est strict si $f(x) > f(a)$ pour tout $x \neq a$. Le point a est un minimum relatif (strict) si c'est un minimum (strict) sur un de ses voisinages ouverts.

Un maximum (strict)(relatif) de f est un minimum (strict)(relatif) de $-f$. Enfin un extremum (strict)(relatif) est un minimum ou un maximum (strict)(relatif).

Les extremums d'une fonction jouent un rôle très important dans beaucoup d'applications. Ainsi les minima d'une fonction potentielle, dont un champ de forces dérive, sont des positions d'équilibre stable du système. Les maxima d'une loi de probabilité correspondent aux états localement les plus probables.

Définition 3.3. Soit une fonction f comme ci-dessus, on dit que a est un point singulier de f si $df(a) = 0$.

Lemme 3.1. *Si $a \in U$ est un extremum relatif d'une fonction différentiable f sur l'ouvert U, alors a est un point singulier de f.*

Démonstration. Supposons que la fonction différentiable f admette un extremum relatif en a. Quitte à remplacer f par $-f$, on peut suposer que ce point est un minimum relatif. Soit $u \in R^n$ un vecteur quelconque et la fonction $\varphi_u : t \mapsto \varphi_u(t) = f(a + tu)$, fonction que l'on peut voir (si $u \neq 0$) comme la restriction de f à la droite passant par a et dirigée par le vecteur u. Cette fonction est définie et différentiable sur un certain intervalle ouvert $]-\eta, \eta[$. Par la définition 1.3, sa dérivée en $t = 0$ est égale à la dérivée directionnelle $f'_u(a)$. Supprimons l'indice u de φ_u pour simplifier la notation. On peut choisir η assez petit pour que 0 soit un minimum de φ, c'est-à-dire que $\varphi(t) \geq \varphi(0)$ pour $t \in]-\eta, \eta]$. Il en résulte que le

rapport d'accroissement $\Delta\varphi(t) = \frac{\varphi(t)-\varphi(0)}{t}$ est négatif pour $t < 0$ est positif pour $t > 0$. En prenant la limite de ce rapport pour $t \to 0_-$, on constate que $\varphi'(0) \leq 0$ et, inversement, que $\varphi'(0) \geq 0$ pour $t \to 0_+$. D'où il suit que $\varphi'(0) = 0$. Comme, par la proposition 1.1, on a $df(a)[u] = f'_u(a)$, il vient $df(a)[u] = 0$. Cette égalité étant vraie pour tout $u \in \mathbb{R}^n$, il s'en suit que $df(a) = 0$ et donc que a est un point singulier de f. ∎

Remarque 3.2. Il est essentiel de supposer U ouvert dans ce lemme ou du moins que l'extremum appartient à l'intérieur du domaine. Ainsi, la fonction $f(x) = x$ atteint son maximum sur le fermé $[0, 1]$ en $x = 1$ (point qui est une extrémité de l'intervalle) et l'on voit que $f'(1) \neq 0$.

La proposition précédente dit qu'il convient de rechercher les extremums d'une fonction différentiable parmi ses points singuliers. Évidemment, certains points singuliers ne sont pas des extremums : par exemple le point 0 pour la fonction $f(x) = x^3$. Nous allons maintenant donner une condition suffisante pour qu'un point singulier soit un extremum. Nous aurons besoin, pour cela, d'utiliser le développement de la fonction au second ordre en son point singulier et des propriétés des formes quadratiques, propriétés que nous allons rappeler.

3.2.2. Rappels sur les formes quadratiques

Une forme bilinéaire sur $\mathbb{R}^n$ est une application $\varphi : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ d'expression $\varphi(u, v) = \sum_{i,j=1}^n a_{ij} u_i v_j$ pour tous les $u = (u_1, \ldots, u_n), v = (v_1, \ldots, v_n) \in \mathbb{R}^n$. Si A est la matrice $(a_{ij})_{ij}$ on peut aussi écrire

$$\varphi(u, v) = \langle A(u), v\rangle,$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire euclidien de $\mathbb{R}^n$. On dit que la forme bilinéaire φ est symétrique si $\varphi(u, v) = \varphi(v, u)$ pour tous les $u, v \in \mathbb{R}^n$, ce qui est équivalent à dire que la matrice A est symétrique ($a_{ij} = a_{ji}$ pour tout i, j). À toute forme bilinéaire φ symétrique on associe une forme quadratique $Q(u) = \varphi(u, u)$. Les formes quadratiques sur $\mathbb{R}^n$ ne sont rien d'autre que les polynômes homogènes de degré 2. L'unique forme bilinéaire symétrique associée à une forme quadratique s'appelle sa forme polaire. Elle est donnée par la formule

$$\varphi(u, v) = \frac{1}{2}(Q(u + v) - Q(u) - Q(v)).$$

Inversement, on a

$$Q(u) = \sum_{i=1}^n a_{ii} u_i^2 + 2 \sum_{i<j} a_{ij} u_i u_j.$$

On appelle *rang* de la forme quadratique, le rang de la matrice symétrique associée $(a_{ij})_{ij}$. On dit que la forme quadratique est *définie* si son rang est maximal, c'est-à-dire si la matrice associée est inversible (de déterminant non nul). Un théorème standard de Gauss classifie les formes quadratiques, à changement linéaire de coordonnées près (voir [8, 19]) :

Théorème 3.3. *Soit Q une forme quadratique de rang r, $0 \leq r \leq n$. Alors il existe un entier s, $0 \leq s \leq r$ et un isomorphisme linéaire L de R^n tels que*

$$Q \circ L(h_1, \ldots, h_n) = \sum_{i=1}^{s} h_i^2 - \sum_{i=s+1}^{r} h_i^2. \tag{3.4}$$

On convient que la première somme est absente si $s = 0$ et que la seconde somme est absente si $s = r$.

Remarque 3.3.

1. La paire (r, s), appelée *signature (de Sylvester)* de la forme quadratique Q, est un invariant de Q pour l'action du groupe linéaire. Autrement dit, une forme quadratique Q admet une réduction (3.4) unique en combinaison de carrés. Par contre, l'isomorphisme linéaire L n'est pas unique. Par exemple, dans le cas d'une forme quadratique définie qui se réduit à $\pm < u, u >$, on peut composer L à droite par un élément quelconque du groupe orthogonal.

2. La réduction d'une forme quadratique dans le groupe orthogonal est une question beaucoup plus difficile. Elle est équivalente à la réduction des matrices symétriques dans ce groupe. On rappelle que toute matrice symétrique est diagonalisable dans le groupe orthogonal, ce qui revient à dire que toute forme quadratique est réductible, dans un système orthonormé de coordonnées, à une combinaison de carrés avec des coefficients réels.

3. L'ensemble des formes quadratiques non dégénérées, ayant une signature donnée $(r, n - r)$, est un ouvert de l'ensemble de toutes les formes quadratiques non dégénérées, qui est lui-même un ouvert dense de l'ensemble de toutes les formes quadratiques. Donc, si une forme quadratique Q' est assez proche d'une forme quadratique non dégénérée, il suit, du théorème 3.3, que Q' est conjuguée à Q, c'est-à-dire qu'il existe un isomorphisme linéaire L tel que $Q' = Q \circ L$. On peut voir ce résultat comme une propriété de stabilité des formes quadratiques non dégénérées. Cette remarque nous sera utile dans la démonstration du lemme de Morse.

Définition 3.4. On dit qu'une forme quadratique définie est *positive* si $s = n$, c'est-à-dire si elle se réduit à une somme de n carrés. Elle est dite définie *négative* si $s = 0$ et $r = n$, c'est-à-dire si $-Q$ est définie positive.

Lemme 3.2. *Une forme quadratique est définie positive si et seulement si $Q(u) > 0$ pour tout $u \in \mathbb{R}^n - \{0\}$, et si et seulement s'il existe une constante $M > 0$ telle que $Q(u) \geq M||u||^2$.*

Démonstration. Il est clair que s'il existe $M > 0$ comme dans l'énoncé, alors $Q(u) > 0$ pour tout $u \in \mathbb{R}^n - \{0\}$ et par le théorème 3.3 la forme quadratique doit être définie positive.

Supposons inversement que Q soit une forme quadratique définie positive. Le théorème 3.3 implique que $Q(u) > 0$ pour tout $u \neq 0$. La restriction de Q à la sphère unité S^{n-1} est une fonction continue et strictement positive. Comme la sphère est compacte, il existe $M > 0$ telle que $Q(u) \geq M$ pour tout $u \in S^{n-1}$. La forme Q étant une fonction homogène de degré 2, on a $Q(\lambda u) = \lambda^2 Q(u)$ pour tout $\lambda \in \mathbb{R}$ et tout $u \in \mathbb{R}^n$. Si on applique cette formule à $u \neq 0$ et $\lambda = \frac{1}{||u||}$, il vient que $Q(u) = ||u||^2 Q(\frac{1}{||u||}u)$. Comme $\frac{1}{||u||}u \in S^{n-1}$, on a $Q(\frac{1}{||u||}u) \geq M$ et donc $Q(u) \geq M||u||^2$ (le vecteur $0 \in \mathbb{R}^n$ vérifie trivialement l'inégalité). ∎

3.2.3. Condition suffisante d'extrémalité

Revenons maintenant à l'étude des fonctions. On va montrer comment repérer les extremums parmi les points singuliers et plus généralement comment étudier une fonction au voisinage d'un point singulier (rappelons qu'au voisinage d'un point régulier, la situation est réglée par le théorème des fonctions implicites).

On suppose donc que $a \in U$ est un point singulier de f (c'est-à-dire que df(a)=0). On va supposer de plus que f est (au moins) de classe $\mathcal{C}^2$. On peut alors écrire la formule de Taylor à l'ordre 2 avec reste de Young :

$$f(a+h) - f(a) = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j + o(||h||^2). \tag{3.5}$$

La partie principale du terme de droite est une forme quadratique

$$Q(h) = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j,$$

appelée *hessienne* de f au point singulier a. Rappelons que, au coefficient $\frac{1}{2}$ près, sa forme bilinéaire associée est la différentielle d'ordre deux de f au point singulier a.

Cette différentielle d'ordre deux a pour expression

$$d^2f(a)(u,v) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) u_i v_j,$$

pour des vecteurs quelconques $u, v \in \mathbb{R}^n$. La forme quadratique (hessienne) $Q(h)$ s'écrit donc $Q(h) = \frac{1}{2} d^2f(a)(h,h)$.

On peut maintenant définir un *point singulier non dégénéré.*

Définition 3.5. Un point singulier $a \in U$ d'une fonction f est non dégénéré, si la hessienne de f en ce point est définie (de déterminant différent de zéro).

L'étude locale au voisinage du point singulier va se faire par comparaison entre l'accroissement de f et sa hessienne, en appliquant les propriétés des formes quadratiques rappelées plus haut. On a le résultat suivant :

Théorème 3.4. *Soit $f : U \mapsto \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ sur l'ouvert U de $\mathbb{R}^n$. Soit $a \in U$ un point singulier et $Q(h)$ la hessienne en a. On a les propriétés suivantes :*

1. *Si a est un point singulier non dégénéré, alors a est isolé parmi les points singuliers de f (il existe un voisinage W de a dans U tel que tous les points de $W - \{a\}$ soient réguliers).*
2. *Si $Q(h)$ est définie positive, alors a est un minimum relatif strict de f.*
3. *Si $Q(h)$ est définie négative, alors a est un maximum relatif strict de f.*
4. *Si $Q(h)$ est définie, ni négative ni positive, alors a n'est pas un extremum de f.*
5. *Si $Q(h)$ n'est pas définie, on ne peut pas conclure quant à la nature du point singulier qui peut être, ou bien ne pas être, un extremum.*

Démonstration.

1. On considère la fonction $F : x \in \mathbb{R}^n \mapsto F(x) = (\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)) \in \mathbb{R}^n$. Cette fonction est de classe $\mathcal{C}^1$ et sa différentielle est la matrice $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{ij}(x)$. Cette matrice est inversible au point a. On peut donc appliquer le théorème de l'inverse à F : il existe un voisinage W de a dans U tel que F soit une bijection sur son image. Comme $F(a) = 0$, on en déduit que $F(x) \neq 0$ si $x \in W - \{a\}$: en conséquence, tout point $x \in W - \{a\}$ est point régulier de f.

2. Grâce au lemme 3.2, on sait qu'il existe une constante positive $M > 0$ telle que $Q(h) \geq M||h||^2$. La formule de Taylor à l'ordre 2 en a implique que

$$f(a+h) - f(a) \geq (M + \varepsilon(h))||h||^2,$$

pour une fonction continue $\varepsilon(h)$ qui tend vers 0 pour $h \to 0 \in \mathbb{R}^n$. Choisissons un voisinage ouvert W de a dans U tel que $|\varepsilon(h)| \leq \frac{M}{2}$ pour tout $a + h \in W$. Il en résulte que pour tout $a + h \in W$, on a : $f(a+h) - f(a) \geq \frac{M}{2}||h||^2$: le point a est donc un minimum strict sur W.

3. Il suffit d'appliquer le point précédent à la fonction $-f$.

4. Grâce au théorème 3.3, on peut choisir des coordonnées de $\mathbb{R}^n$ pour lesquelles la hessienne Q a pour expression

$$Q(h) = h_1^2 + \ldots + h_s^2 - h_{s+1}^2 - \ldots - h_n^2,$$

avec $1 \leq s \leq n-1$. Posons

$$f_1 = f|_{\{a\}+(\mathbb{R}^s \times \{0\})} \quad \text{et} \quad f_2 = f|_{\{a\}+(\{0\}\times\mathbb{R}^{n-s})}.$$

En utilisant les deux points précédents avec $x_s \in \{a\} + (\mathbb{R}^s \times \{0\})$ et $y_{n-s} \in \{a\} + (\{0\} \times \mathbb{R}^{n-s})$, on obtient que a est un minimum strict pour f_1 (car $f_1(a) = f(a) < f(x_s) = f_1(x_s)$) et un maximum strict pour f_2 (car $f_2(a) = f(a) > f(y_{n-s}) = f_2(y_{n-s})$). Dans tout voisinage W de a, on peut donc trouver un point x et un point y tels que $f(y) < f(a) < f(x)$, ce qui implique que a n'est pas un extremum de f.

5. Considérons les deux fonctions $f(x) = x^4$ et $g(x) = x^3$. Ces deux fonctions ont à l'origine un point singulier avec une dérivée seconde nulle. Or la fonction f a un minimum strict en 0 alors que g est strictement monotone. On ne peut donc rien dire lorsque $Q(h)$ (qui correspond ici à la dérivée seconde) n'est pas définie. ■

La conclusion du point 4 du théorème ci-dessus est un peu décevante. En fait, on peut préciser le comportement d'une fonction de classe C^∞ au voisinage de tout point singulier non dégénéré. Commençons par un résultat général et facile en dimension 1 :

Proposition 3.1. *Soit $f : I \mapsto \mathbb{R}$ une fonction de classe C^∞, définie sur un intervalle ouvert I, voisinage de 0. Supposons que f ait une de ses dérivées successives en 0 qui soit non nulle. Soit $k \geq 1$ le plus petit entier tel que $\frac{d^k f}{dx^k}(0) \neq 0$. Alors il existe*

un difféomorphisme Φ *de classe* $\mathcal{C}^\infty$, *d'un intervalle ouvert* J *de* 0 *sur* I, *tel que* $\Phi(0) = 0$ *et*

$$f \circ \Phi(X) = f(0) \pm X^k,$$

où $\pm$ *est le signe de* $\frac{d^k f}{dx^k}(0)$.

Démonstration. Posons pour simplifier, $g(x) = f(x) - f(0)$. On a $g(0) = 0$ et k est le plus petit entier tel que $\frac{d^k g}{dx^k}(0) \neq 0$. Appliquons au point 0 la formule de Taylor avec reste intégral à l'ordre k. La formule se résume au seul reste et l'on obtient que $g(x) = x^k \varphi(x)$ pour une certaine fonction φ de classe $\mathcal{C}^\infty$ sur I, avec pour expression explicite :

$$\varphi(x) = \frac{1}{(k-1)!} \int_0^1 (1-t)^{k-1} \frac{d^k f}{dx^k}(tx) dt.$$

La fonction φ est de classe $\mathcal{C}^\infty$ car l'intégrant $(1-t)^{k-1} \frac{d^k f}{dx^k}(tx)$ est lui-même de classe $\mathcal{C}^\infty$ en (t, x). La formule de Taylor (avec reste de Young) de f en $x = 0$ s'écrit

$$f(x) = \varphi(0)x^k + o(x^k) = \frac{1}{k!} \frac{d^k g}{dx^k}(0) x^k + o(x^k),$$

d'où il suit $\varphi(0) = \frac{1}{k!} \frac{d^k g}{dx^k}(0)$ par unicité de la formule. On en tire que $\varphi(0) \neq 0$. Soit $\varepsilon = \pm$ le signe de $\varphi(0)$. On peut écrire que

$$\varphi(x) = \varepsilon A(1 + \eta(x)),$$

où $A = \frac{1}{k!} \left| \frac{d^k g}{dx^k}(0) \right|$ est une constante strictement positive et $\eta(x) = \frac{\varphi(x) - \varepsilon A}{\varepsilon A}$ est une fonction de classe $\mathcal{C}^\infty$ sur I, telle que $\eta(0) = 0$. Il existe alors un voisinage ouvert W de 0 sur lequel la fonction $(1+\eta(x))^{\frac{1}{k}}$ est définie et de classe $\mathcal{C}^\infty$. Posons $H(x) = A^{\frac{1}{k}} x (1 + \eta(x))^{\frac{1}{k}}$. Dès lors $g(x) = x^k \varphi(x) = \varepsilon (H(x))^k$. H est de classe $\mathcal{C}^\infty$ et $H'(0) = A^{\frac{1}{k}} > 0$. Par le théorème de l'inverse, H est un difféomorphisme sur un intervalle ouvert contenant 0 et l'on a sur cet intervalle :

$$f(x) - f(0) = g(x) = \varepsilon (H(x))^k.$$

Si Φ est le difféomorphisme inverse de H, on a, sur son domaine de définition J, que $f \circ \Phi(X) = f(0) + \varepsilon X^k$ où $\varepsilon = \pm$ est le signe de $\frac{d^k f}{dx^k}(0)$. ∎

Remarque 3.4. Soit f une fonction de classe $\mathcal{C}^\infty$ définie sur un ouvert U de $\mathbb{R}^n$. Si toutes les dérivées partielles de f en x_0 sont nulles, on dit que la fonction $f - f(x_0)$ est *plate* en x_0. Supposons que l'on soit en dimension $n = 1$ et que la fonction $f - f(x_0)$ soit non plate en x_0. Cela est équivalent à dire que l'une des dérivées successives $\frac{d^k f}{dx^k}(x_0)$ n'est pas nulle. On appelle *ordre* de f en x_0 le plus petit entier $k \geq 1$ avec la propriété que $\frac{d^k f}{dx^k}(x_0) \neq 0$.

Le résultat précédent permet de dire, en particulier, que si 0 est un point singulier non dégénéré ($f'(0) = 0$, $f''(0) \neq 0$), l'application f s'écrit comme une fonction quadratique dans des coordonnées bien choisies. Ce résultat se généralise en dimension quelconque par le lemme de Morse suivant. Notons par $|\frac{\partial^2 f}{\partial x_i \partial x_j}(a)|$ le déterminant de la hessienne de f au point a. Nous avons :

***Théorème 3.5* (Lemme de Morse).** *Soit $f : U \mapsto \mathbb{R}$ une fonction sur un ouvert U de $\mathbb{R}^n$ de classe C^∞. On suppose que $a \in U$ est un point singulier non dégénéré ($\frac{\partial f}{\partial x_i}(a) = 0$, $i = 1, \ldots, n$ et $|\frac{\partial^2 f}{\partial x_i \partial x_j}(a)| \neq 0$). Alors il existe un voisinage ouvert B de $0 \in \mathbb{R}^n$ et un difféomorphisme Φ de B sur son image $\Phi(B) = W$ dans U, de classe C^∞, tel que $\Phi(0) = a$ et*

$$f \circ \Phi(X_1, \ldots, X_n) = f(a) + \sum_{i=1}^{s} X_i^2 - \sum_{i=s+1}^{n} X_i^2. \tag{3.6}$$

$(s, n-s)$ est la signature de la hessienne de f au point a).

Démonstration. En dimension $n = 1$, c'est la proposition 3.1 avec $k = 2$. Avant de passer au cas général, nous allons donner une preuve détaillée en dimension 2.

On peut supposer sans inconvénient que le point a est l'origine. L'hypothèse est alors que l'origine est un point singulier et que la forme bilinéaire $d^2 f(0,0)$ est non dégénérée. La formule de Taylor à l'ordre un avec reste intégral s'écrit

$$f(x,y) - f(0,0) = \int_0^1 (1-t) d^2 f(tx, ty)((x,y),(x,y)) dt,$$

soit

$$f(x,y) - f(0,0) = \alpha(x,y)x^2 + 2\beta(x,y)xy + \gamma(x,y)y^2, \tag{3.7}$$

avec $\alpha(x,y) = \int_0^1 (1-t) f''_{x^2}(tx,ty)dt$, $\beta(x,y) = \int_0^1 (1-t) f''_{xy}(tx,ty)dt$ et $\gamma(x,y) = \int_0^1 (1-t) f''_{y^2}(tx,ty)dt$, ces trois fonctions étant de classe C^∞ (par dérivation sous le signe somme) sur un ouvert U de l'origine $(0,0)$.

Remarquons tout d'abord que l'on peut toujours supposer que $\alpha(0,0) > 0$. En effet :

1. Si $\alpha(0,0) < 0$, on remplace f par $-f$.

2. Si $\alpha(0,0) = 0$ et que $\gamma(0,0) \neq 0$, on permute les coordonnées, c'est-à-dire on remplace $f(x,y)$ par $f(y,x)$.

3. Si $\alpha(0,0) = \gamma(0,0) = 0$, remarquons que $\det(d^2 f(0,0)) = -4(\beta(0,0))^2 \neq 0$.
On fait le changement de variables linéaire : $x = X, y = Y + X$.
La fonction $F(X,Y) = f(X, Y+X) - f(0,0)$ s'écrit :

$$F(X,Y) = \tilde{\alpha}(X,Y)(X)^2 + 2\tilde{\beta}(X,Y)X(Y+X) + \tilde{\gamma}(X,Y)(Y+X)^2$$

avec $\tilde{\alpha}(X,Y) = \alpha(X, Y+X)$, $\tilde{\beta}(X,Y) = \beta(X, Y+X)$ et $\tilde{\gamma}(X,Y) = \gamma(X, Y+X)$. Il est clair que $\tilde{\alpha}(0,0) = \tilde{\gamma}(0,0) = 0$ et $\tilde{\beta}(0,0) \neq 0$.
En développant, on obtient :

$$F(X,Y) = A(X,Y)X^2 + B(X,Y)XY + C(X,Y)Y^2$$

avec $A = \tilde{\alpha} + 2\tilde{\beta} + \tilde{\gamma}$, $B = 2\tilde{\beta} + 2\tilde{\gamma}$ et $C = \tilde{\gamma}$. Maintenant on a $A(0,0) = 2\tilde{\beta}(0,0)$ et l'on est ramené au premier cas.

On est donc ramené au seul cas où, dans l'expression (3.7), le coefficient $\alpha(0,0)$ est strictement positif, ce que l'on va supposer dorénavant.

La fonction $\alpha(x,y)$ étant continue, elle reste strictement positive dans un voisinage de l'origine. En ne mentionnant plus la dépendance en (x,y) pour alléger l'écriture, on peut écrire localement au voisinage de $(0,0)$:

$$\alpha x^2 + 2\beta xy + \gamma y^2 = \alpha\left(x + \frac{\beta}{\alpha}y\right)^2 + \frac{\alpha\gamma - \beta^2}{\alpha}y^2.$$

Comme $d^2 f(0,0)$ est non dégénérée par hypothèse, on a $(\alpha\gamma - \beta^2)(0,0) \neq 0$. Soit B un voisinage ouvert de l'origine dans $\mathbb{R}^2$, dans lequel les deux fonctions α et $\alpha\gamma - \beta^2$ n'ont pas de zéros, et soit $\varepsilon = \pm 1$ le signe de la fonction $\alpha\gamma - \beta^2$. Définissons alors les fonctions u et v par : $u(x,y) = \sqrt{\alpha(x,y)}\left(x + \frac{\beta(x,y)}{\alpha(x,y)}y\right)$ et $v(x,y) = \sqrt{\varepsilon\frac{\alpha(x,y)\gamma(x,y) - \beta^2(x,y)}{\alpha(x,y)}}y$, elles sont de classe $\mathcal{C}^\infty$ et d'après (3.7)

$$f(x,y) - f(0,0) = u(x,y)^2 + \varepsilon v(x,y)^2.$$

Enfin l'application $\Phi : (x,y) \mapsto (u,v)$ est un changement de coordonnées au voisinage de $(0,0)$, le déterminant du jacobien à l'origine s'écrit

$$\det(D\Phi(0,0)) = \det\begin{pmatrix} \sqrt{\alpha} & \frac{\beta}{\sqrt{\alpha}} \\ 0 & \sqrt{\varepsilon\frac{\alpha\gamma - \beta^2}{\alpha}} \end{pmatrix}(0,0) = \left(\sqrt{\varepsilon(\alpha\gamma - \beta^2)}\right)(0,0) > 0.$$

La matrice jacobienne étant inversible, le théorème de l'inverse permet de conclure, c'est-à-dire d'affirmer que Φ est un difféomorphisme de classe $\mathcal{C}^\infty$ de B sur son image $\Phi(B)$.

Pour démontrer le lemme de Morse en dimension n, on considère comme précédemment la formule de Taylor à l'ordre un avec reste intégral

$$f(x) - f(0) = x^T Q(x) x,$$

où $Q(x)$ est la matrice symétrique

$$Q(x) = \int_0^1 (1-t) d^2 f(tx) dt,$$

qui est une fonction de classe C^∞ de x. Soit B un voisinage de $0 \in \mathbb{R}^n$ dans lequel la matrice $Q(x)$ soit non dégénérée et de signature constante, égale à celle de $Q(0)$. En utilisant le théorème 3.3 (voir aussi la remarque 3.3), on trouve une fonction $M : x \mapsto M(x)$, de classe C^∞, à valeurs dans les matrices inversibles, définie au voisinage de $0 \in \mathbb{R}^n$, telle que $Q(x) = M(x)^T Q(0) M(x)$.

Posons maintenant : $y = M(x)x$. On a

$$f(x) - f(0) = x^T M(x)^T Q(0) M(x) x = y^T Q(0) y.$$

Comme $Q(0) = (1/2) d^2 f(0)$ est de signature $(s, n-s)$, il existe, d'après le théorème 3.3, un changement linéaire de coordonnées (c'est-à-dire un changement de base) $y = Su$ où S est une matrice inversible, conservant la signature, tel que

$$y^T Q(0) y = u^T S^T Q(0) S u = \sum_{i=1}^{r} u_i^2 - \sum_{j=r+1}^{n} u_j^2.$$

Enfin l'application $x \mapsto y = M(x)x$ a pour différentielle à l'origine la matrice $M(0)$ qui est inversible. C'est donc, toujours par le théorème de l'inverse, un difféomorphisme entre deux voisinages de l'origine dans $\mathbb{R}^n$. Comme S est inversible, on peut définir l'application $x \mapsto u = S^{-1} \circ M(x)x$ qui est aussi un difféomorphisme local. Le difféomorphisme inverse Φ vérifie la conclusion souhaitée :

$$f(\Phi(u)) - f(0) = \sum_{i=1}^{r} u_i^2 - \sum_{j=r+1}^{n} u_j^2. \qquad \blacksquare$$

Remarque 3.5.

1. On peut voir dans le lemme de Morse une généralisation non linéaire de la réduction de Gauss des formes quadratiques.

2. On peut démontrer le lemme de Morse pour une fonction de classe C^3. Compte tenu de l'expression du reste intégral à l'aide des dérivées secondes dans la formule de Taylor, le difféomorphisme Φ de conjugaison construit dans la démonstration, sera seulement de classe C^1.

Le théorème 3.5 nous permet d'affirmer que les formes quadratiques non-dégénérées, en écriture réduite comme dans (3.6), sont des modèles pour le comportement d'une fonction au voisinage d'un point singulier non dégénéré. Il est facile de dessiner le dessin, des lignes de niveaux de ces modèles en dimension 2 (figure 3.1), et des surfaces de niveaux en dimension 3 (figure 3.2). Dans les coordonnées initiales, le dessin doit être modifié par un difféomorphisme local (figure 3.3).

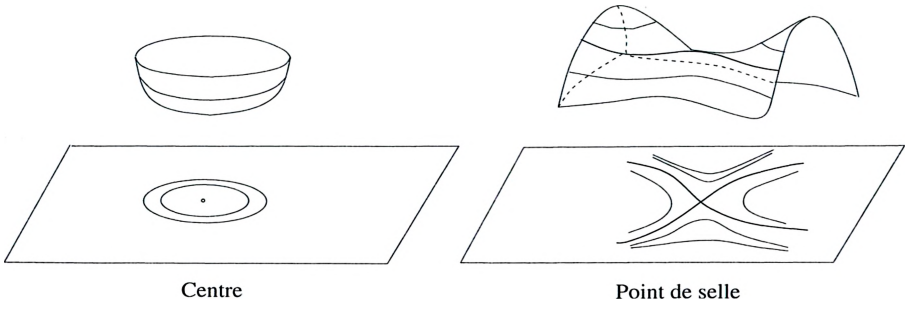


FIGURE 3.1

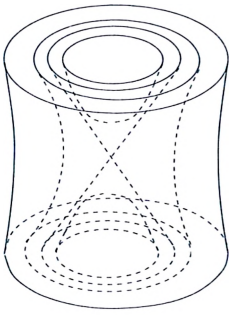


FIGURE 3.2

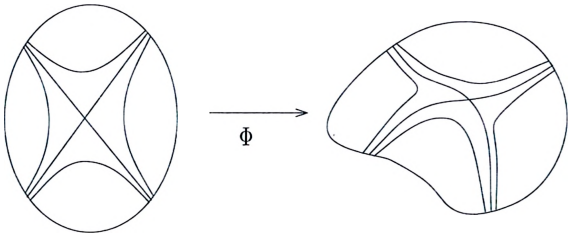


FIGURE 3.3

En dimension 2, un point singulier non dégénéré de hessienne définie positive ou négative est appelé *centre* (figure 3.1). Un point singulier avec une hessienne mixte (se réduisant à une différence de deux termes carrés) est appelée *point de selle* ou *col* (allusion à la forme dans $\mathbb{R}^3$ du graphe de la fonction modèle $Q(x,y) = x^2 - y^2$; figure 3.1).

En dimension $n \geq 3$, on appelle encore centres les points singuliers de hessienne définie, positive ou négative. Le point singulier, correspondant à la valeur singulière, est entouré par des hypersurfaces de niveau, qui sont des sphères concentriques au point singulier. La configuration locale pour les points singuliers de signature intermédiaire va dépendre de la dimension n et de cette signature. Considérons par exemple le cas $n = 3$. On a dans ce cas deux possibilités correspondant aux signatures $\{+,+,-\}$ et $\{+,-,-\}$. Évidemment, on passe d'un cas à l'autre en changeant le signe de la fonction, ce qui ne change pas globalement les surfaces de niveau. On peut donc considérer le cas de la forme quadratique $Q(x,y,z) = x^2 + y^2 - z^2$, dont les surfaces de niveau sont représentées dans la figure 3.2.

La surface singulière $Q^{-1}(0)$ est un cône. On a deux types possibles de surfaces de niveau régulières : des hyperboloïdes à une nappe, difféomorphes à $S^1 \times \mathbb{R}$ en considérant $Q^{-1}(\alpha)$ avec $\alpha < 0$, et des hyperboloïdes à deux nappes (union de deux surfaces difféomorphes à $\mathbb{R}^2$) en considérant $Q^{-1}(\alpha)$ avec $\alpha > 0$.

Étude globale d'une fonction de Morse

Pour une fonction de classe $\mathcal{C}^2$ la hessienne est bien définie, cela autorise la définition suivante :

Définition 3.6. Une fonction de Morse est une fonction de classe $\mathcal{C}^2$ dont tous les points singuliers sont non dégénérés (on dit encore qu'ils ont le type de Morse).

Grâce au lemme de Morse, on connaît le comportement local d'une fonction de Morse au voisinage de chacun de ses points singuliers. Entre deux valeurs singulières, il est facile de montrer que l'on peut trivialiser la fonction, du moins si l'on reste sur une région compacte. Plus précisément, on a le résultat suivant [12] :

Proposition 3.2. *Soit $f : U \mapsto \mathbb{R}$ une fonction sur un ouvert U de $\mathbb{R}^n$ de classe $\mathcal{C}^\infty$. Considérons un intervalle $[\alpha,\beta]$ inclus dans les valeurs régulières atteintes de la fonction et supposons que $f^{-1}([\alpha,\beta])$ soit une partie compacte de U. Alors, $V = f^{-1}(\alpha)$ est une sous-variété compacte de codimension 1 de U (hypersurface compacte) et $f^{-1}([\alpha,\beta])$ est difféomorphe à $V \times [\alpha,\beta]$, par un difféomorphisme envoyant $V \times \{y\}$ sur $f^{-1}(y)$ pour tout $y \in [\alpha,\beta]$.*

Il est alors possible de faire des études globales qualitatives de fonctions en combinant la proposition 3.2 avec le lemme de Morse. Considérons par exemple la fonction $f : (x,y) \in \mathbb{R}^2 \mapsto f(x,y) = xy(x+y-1)$. Calculons la différentielle de f :

$$df(x,y) = y(2x+y-1)dx + x(x+2y-1)dy.$$

On trouve quatre points singuliers qui sont tous non dégénérés : f est une fonction de Morse. Trois d'entre eux sont des points de selles : $O = (0,0)$, $A = (1,0)$, $B = (0,1)$. Le quatrième point $C = (\frac{1}{3}, \frac{1}{3})$, situé dans l'intérieur du triangle $[O,A,B]$, est un centre pour lequel f admet un minimum relatif (figure 3.4). Nous allons montrer ce résultat.

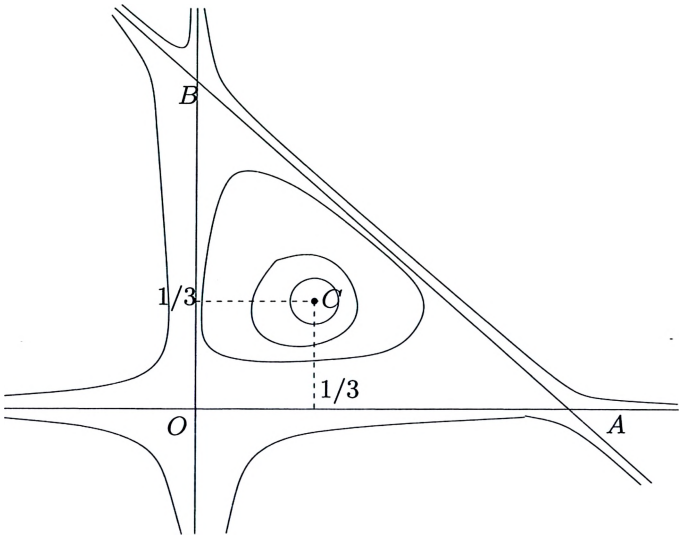


Figure 3.4

On détermine facilement cette nature des points singuliers en calculant le développement limité de f à l'ordre 2 en chacun de ces points. Considérons par exemple le point C. On introduit des coordonnées locales (X,Y) au voisinage du point C en posant : $x = \frac{1}{3} + X$, $y = \frac{1}{3} + Y$. Le développement limité de f en (X,Y) s'écrit :

$$\tilde{f}(X,Y) = f\left(\frac{1}{3}+X, \frac{1}{3}+Y\right) = -\frac{1}{27} + \frac{1}{3}(X^2+Y^2+XY) + o(||(X,Y)||^2).$$

Le discriminant de la forme quadratique X^2+Y^2+XY est égal à -3. Cette forme quadratique est donc définie positive et C est un centre (en fait $X^2+Y^2+XY = (X+\frac{Y}{2})^2 + \frac{3}{4}Y^2$ est somme des carrés des coordonnées $X+\frac{Y}{2}$ et $\sqrt{\frac{3}{4}}\,Y$). Les autres points singuliers peuvent se traiter de la même façon. Remarquez que

leur nature est imposée par le fait que la ligne de niveau $f^{-1}(0)$ est formée de l'union de trois droites transverses se coupant aux points O, A, B. On peut obtenir maintenant une représentation globale des lignes de niveaux. On considère tout d'abord la fonction dans l'intérieur U du triangle $[OAB]$ et pour les valeurs prises dans ce triangle, c'est-à-dire pour $\alpha \in [-\frac{1}{27}, 0[$. Pour $\alpha > -\frac{1}{27}$ mais proche de cette dernière valeur, les lignes de niveaux sont des cercles plongés entourant le centre C, comme il suit du théorème 3.5 : ces lignes de niveau sont des images difféomorphes de cercles $x^2 + y^2 = \text{Const}$. On peut appliquer ensuite la proposition 3.2 à une région de la forme $(f|_U)^{-1}([\alpha, \beta])$ avec $\alpha > -\frac{1}{27}$ et proche de $-\frac{1}{27}$ et $\beta < 0$ et proche de 0. En faisant tendre β vers 0, on obtient finalement un difféomorphisme, entre un disque ouvert (centré en $0 \in \mathbb{R}^2$) et l'intérieur du triangle, qui envoie les cercles concentriques sur les lignes de niveau de f dans U. Ce difféomorphisme se prolonge en un homéomorphisme envoyant le bord du disque sur le triangle. On peut faire de la même façon l'étude des lignes de niveau à l'extérieur du triangle. Une petite difficulté tient au fait que cet extérieur n'est pas compact. Une façon de procéder est de faire l'étude dans un disque dont on fait ensuite tendre le rayon vers l'infini. Le résultat final est que les six régions complémentaires, dans l'extérieur des trois droites formant $f^{-1}(0)$, sont remplies de courbes unicursales disjointes, qui sont composantes connexes de lignes de niveau.

3.3. Point singulier d'une fonction sur une sous-variété

3.3.1. Définitions et exemples

Soit $V \subset U$ une sous-variété d'un ouvert de $\mathbb{R}^n$ et f une fonction différentiable sur U. On veut maintenant étudier, par exemple, les extremums de la fonction restreinte : $f|_V$. Cette question apparaît dans beaucoup d'applications : pensez à la recherche des équilibres possibles d'un système mécanique sujet à des contraintes. Ces contraintes sont modélisées par l'appartenance du point représentatif du système à une sous-variété de l'espace total du système sans contraintes (cet espace total est en général un ouvert dense d'un espace euclidien (cf. [3] par exemple)).

On peut généraliser sans difficulté à $f|_V$ les définitions 3.2 (en fait, ces définitions n'utilisent que l'ordre de $\mathbb{R}$ et l'espace de définition de la fonction peut être quelconque). Par contre, il convient de donner un sens à la notion de point singulier.

Définition 3.7. Soit V et f comme ci-dessus. On dit que $a \in V$ est un point singulier de $f|_V$ (on disait autrefois : un point singulier lié) si et seulement s'il existe une carte (W, φ) telle que $a \in W$ et $\varphi(a)$ soit un point singulier de la fonction $f \circ \varphi^{-1}$ (au sens usuel d'une fonction définie sur un ouvert euclidien).

Remarque 3.6.

1. Il est clair que la définition ci-dessus peut être donnée directement sur une variété. Cela sera le cas pour toutes les notions que nous allons introduire dans ce paragraphe : on pourrait considérer directement une fonction sur une variété et non la restriction d'une fonction à une sous-variété. La seule raison pour se limiter à ce cas est que les exemples de fonctions seront plus explicites.

2. La définition donnée ci-dessus est cohérente dans le sens qu'elle ne dépend pas du choix de la carte (W, φ). En effet, si on considère une deuxième carte (W', φ') on voit que l'application $f \circ \varphi^{-1}$ diffère localement de l'application $f \circ \varphi'^{-1}$ par composition à droite du difféomorphisme $\varphi \circ \varphi'^{-1}$ d'ouverts euclidiens. Or, et c'est une remarque très générale, on peut aisément vérifier que toutes les définitions locales introduites dans le cadre du calcul différentiel sont indépendantes du choix des coordonnées locales. Il s'en suit qu'une notion introduite dans une carte d'une sous-variété est indépendante du choix de la carte, par exemple la notion de point singulier (cela suit immédiatement de la règle de composition des différentielles appliquée aux applications $\varphi \circ \varphi'^{-1}$ et $f \circ \varphi^{-1}$). Cela va être aussi le cas des notions que nous allons introduire dans la suite de ce paragraphe, et on se dispensera en général de répéter cette remarque à chaque occasion.

3. Il est clair que si $a \in V$ est un point singulier de f en tant que fonction sur U, ce point est *a fortiori* point singulier de $f|_V$. La réciproque est évidemment fausse. Par exemple sur le premier exemple ci-dessous, la fonction hauteur $z : \mathbb{R}^3 \mapsto \mathbb{R}$, *restreinte à la sous-variété* S^2, a le point a comme point singulier, alors qu'elle n'a pas de points singuliers *en tant que fonction de* $\mathbb{R}^3$.

On peut définir le rang et la signature de la hessienne en un point singulier a de $f|_V$ en utilisant une carte comme plus haut. L'expression de la hessienne dépend évidemment du choix de la carte, mais pas le rang et la signature. On dira que le point singulier est non dégénéré si la hessienne est une forme quadratique de rang maximal. Une fonction $f|_V$ dont tous les points singuliers sont non dégénérés est appelée fonction de Morse de la sous-variété.

Exemples. Considérons la fonction hauteur $f : (x, y, z) \mapsto f(x, y, z) = z$ et la sphère S^2 (on suppose que l'axe Oz est vertical !). La fonction restreinte $f|_{S^2}$ admet deux point singuliers : $(0, 0, 1)$ et $(0, 0, -1)$. Pour étudier la nature du point $a = (0, 0, 1)$ par exemple, on peut représenter la sphère localement au voisinage du point a comme le graphe de l'application $\pi : (x, y) \mapsto z = \pi(x, y) = \sqrt{1 - x^2 - y^2}$.

On a une carte locale (W, φ) au voisinage de a, en inversant cette application de graphe : $\varphi^{-1} = \pi$. Il en résulte que :

$$f \circ \varphi^{-1}(x, y) = f \circ \pi(x, y) = \sqrt{1 - x^2 - y^2}.$$

Cette fonction a pour développement limité :

$$\sqrt{1 - x^2 - y^2} = 1 - \frac{1}{2}(x^2 + y^2) + o(||(x, y)||^2).$$

Le point a est donc un maximum non dégénéré. En effet, que le point a soit un point singulier est trivial car la partie linéaire du développement limité est nulle (elle est absente dans ce développement !). Que ce soit un maximum est évident car la hessienne est définie négative comme étant réduite, au coefficient $\frac{1}{2}$ près, à *moins la somme des carrés des coordonnées* (hessienne de signature $\{-, -\}$) : on peut alors appliquer le point (3) du théorème 3.4.

Un exemple moins trivial est obtenu en considérant à nouveau la fonction hauteur et un tore plongé de façon verticale dans $\mathbb{R}^3$ (à la manière d'une chambre à air posée verticalement sur un plan horizontal (figure 3.5)). Par exemple, on peut choisir le plongement de paramétrisation globale : $\Phi(\theta, \varphi)$ donnée par

$$x = r \sin\theta \; ; \quad y = (R + r\cos\theta)\sin\varphi \; ; \quad z = (R + r\cos\theta)\cos\varphi + R + r,$$

avec $R > r > 0$. Des applications de paramétrisation locale (inverses d'applications de carte) en sont déduites en restreignant convenablement θ et φ. La fonction hauteur en restriction $f|_V$ est donnée dans chaque carte par la fonction coordonnée $z = (R + r\cos\theta)\cos\varphi + R + r$.

On peut observer que $f|_V$ a quatre points singuliers en $\theta = 0, \pi$ et $\varphi = 0, \pi$. Cela se déduit immédiatement de l'expression des dérivées partielles : $\frac{\partial z}{\partial \theta} = -R\cos\varphi\sin\theta$ et $\frac{\partial z}{\partial \varphi} = (R + \cos\theta)(-\sin\varphi)$. Ces points sont non dégénérés : un maximum, un minimum et deux points de selle. Par exemple, pour $\theta = \pi, \varphi = 0$, on a le point singulier A de coordonnées $(0, 0, 2R)$. En ce point, on a le développement limité

$$z = 2R + \frac{r}{2}u^2 - \frac{R - r}{2}\varphi^2 + o(u^2 + \varphi^2),$$

par rapport aux coordonnées locales $u = \theta - \pi, \varphi$. Comme $R > r > 0$, la partie quadratique du développement est une différence de deux carrés : la signature de la hessienne est donc égale à $(-1, 1)$, d'où il suit que le point A est un point singulier non dégénéré de type selle (figure 3.5).

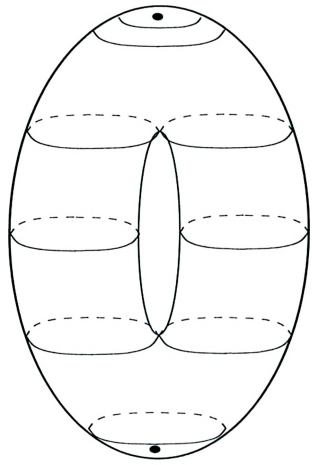


Figure 3.5

3.3.2. Multiplicateurs de Lagrange

La localisation des points singuliers d'une fonction restreinte $f|_V$ est facilitée par la caractérisation suivante :

Proposition 3.3. *Soit f et V comme ci-dessus. Alors $a \in V$ est un point singulier de $f|_V$ si et seulement si*

$$T_aV \subset \text{Ker } df(a). \tag{3.8}$$

Démonstration. On choisit une carte (W, φ) telle que $a \in W$. il est clair qu'un point a est singulier pour $f|_V$ si et seulement si, pour tout chemin différentiable $c(t)$ passant par a ($c(0) = a$), et tracé sur V (comme dans la définition 2.7 de l'espace tangent), la fonction composée $f \circ c(t)$ admet 0 comme valeur singulière. Cela est équivalent à $df(a)[\frac{dc}{dt}(0)] = 0$. On en conclut que $T_aV \subset \text{Ker } df(a)$ si a est un point singulier de $f|_V$. Inversement, si cette inclusion est vérifiée, on a $d(f \circ \varphi^{-1})(0) = 0$ pour tout chemin tracé sur V comme plus haut, et on en conclut que a est un point singulier de $f|_V$. ∎

Nous allons donner quelques formulations analytiques de la condition donnée dans la proposition 3.3. Supposons que V soit définie localement par une équation cartésienne au voisinage d'un point singulier a. Explicitement, supposons que V soit de dimension $p < n$ et que sur l'ouvert U', voisinage de a dans U, on se donne une application $g : U' \mapsto \mathbb{R}^{n-p}$ de classe $\mathcal{C}^\infty$, telle que $V \cap U' = g^{-1}(\alpha)$ pour une valeur régulière $\alpha \in \mathbb{R}^{n-p}$. Introduisons des composantes : $\alpha = (\alpha_1, \ldots, \alpha_{n-p})$ et

$g(x) = (g_1, \ldots, g_{n-p})$. Soit $x \in V \cap U'$. Comme le rang de $dg(x)$ est maximum, on a

$$\text{Ker } dg(x) = \text{Ker } dg_1(x) \cap \ldots \cap \text{Ker } dg_{n-p}(x),$$

où les noyaux Ker $dg_i(x)$ sont des hyperplans 2 à 2 transverses. Rappelons que $T_xV = \text{Ker } dg(x)$. La condition (3.8) se traduit donc par la condition d'inclusion

$$\text{Ker } dg_1(a) \cap \ldots \cap \text{Ker } dg_{n-p}(a) \subset \text{Ker } df(a).$$

Par dualité, on obtient une formule équivalente pour les formes linéaires

$$df(a) \in \mathbb{R}\{dg_1(a), \ldots, dg_{n-p}(a)\},$$

ce qui signifie que la forme $df(a)$ appartient à l'espace vectoriel engendré par les formes $dg_i(a)$. Cette relation d'appartenance se traduit par l'existence de coefficients $\lambda_1, \ldots, \lambda_{n-p}$ tels que

$$df(a) = \sum_{i=1}^{n-p} \lambda_i dg_i(a)\,; \tag{3.9}$$

composante par composante, cette égalité s'écrit

$$\frac{\partial f}{\partial x_j}(a) = \sum_{i=1}^{n-p} \lambda_i \frac{\partial g_i}{\partial x_j}(a), \quad \text{pour } \ j = 1, \ldots, n. \tag{3.10}$$

Les coefficients $\lambda_1, \ldots, \lambda_{n-p} \in \mathbb{R}$ sont appelés *multiplicateurs de Lagrange.*

3.3.3. Le cas de la codimension 1

Considérons deux fonctions f et g définies sur un même ouvert U. Soit α une valeur régulière de f et β une valeur régulière de g. Posons $V_\alpha = f^{-1}(\alpha)$ et $W_\beta = g^{-1}(\beta)$. Supposons que $a \in V_\alpha \cap W_\beta$. On voit que $f|_W$ admet a comme point singulier si et seulement s'il existe un λ tel que $df(a) = \lambda dg(a)$. Mais par hypothèse, $df(a)$ et $dg(a)$ sont des formes non nulles. Donc $\lambda \neq 0$, et l'on peut écrire que $dg(a) = \frac{1}{\lambda} df(a)$, autrement dit a est aussi un point singulier pour $g|_{V_\alpha}$. On a ainsi une relation symétrique : a est un point singulier de $g|_{V_\alpha}$ si et seulement si a est un point singulier de $f|_{W_\beta}$. Cette symétrie est très claire au niveau des plans tangents :

$$\text{Ker } df(a) = T_aV_\alpha = T_aW_\beta = \text{Ker } dg(a). \tag{3.11}$$

On dit, dans ce cas, que les deux hypersurfaces V_α et W_β sont tangentes, ou bien ont un contact au point a, ce qui est bien évidemment une relation d'équivalence.

C'est la situation qui se présente dans les exemples donnés plus haut où la fonction g est la fonction hauteur. Un point a sur la surface V_α (la sphère ou bien un tore dans les exemples proposés) est un point singulier pour la fonction hauteur, si et seulement si la surface est tangente en a au plan horizontal passant par a. Il est possible de dire plus : a est un point singulier non dégénéré pour $f|_{W_\beta}$ si et seulement si a est un point non dégénéré pour $g|_{V_\alpha}$. De plus la nature des points singuliers est identique : a est un extremum pour $f|_{W_\beta}$ si et seulement si a est un extremum pour $g|_{V_\alpha}$ et, pour des surfaces en dimension 3, le point a est un point de selle pour $f|_{W_\beta}$ si et seulement si c'est le cas pour $g|_{V_\alpha}$.

Remarquons de plus que l'on peut regarder les propriétés de contact, non plus entre deux hypersurfaces particulières V_α et W_β, mais entre deux familles d'hypersurfaces V_α et W_β, lorsque α et β varient. En effet, si les surfaces V_{α_0} et W_{β_0} ont un point de contact non dégénéré en a_0, alors le théorème des fonctions implicites va nous fournir deux fonctions $\beta(\alpha)$ et $a(\alpha)$, définies pour α voisin de α_0, avec $\beta(\alpha_0) = \beta_0$ et $a(\alpha_0) = a_0$, telles que V_α ait un contact avec $W_{\beta(\alpha)}$ au point $a(\alpha)$. Le point de contact (ou de tangence) entre les hypersurfaces décrit la courbe γ définie par la fonction $a : \alpha \mapsto a(\alpha)$ transverse aux deux familles d'hypersurfaces.

Chacune de ces familles définit ce qui est appelé un *feuilletage*, et cette courbe γ est le lieu des contacts entre les deux feuilletages (figure 3.6). Pour les différents β on a une série d'intersections avec V_α pour α fixé. La figure 3.6 montre des intersections du type centre, mais on peut aussi avoir des intersections du type selle.

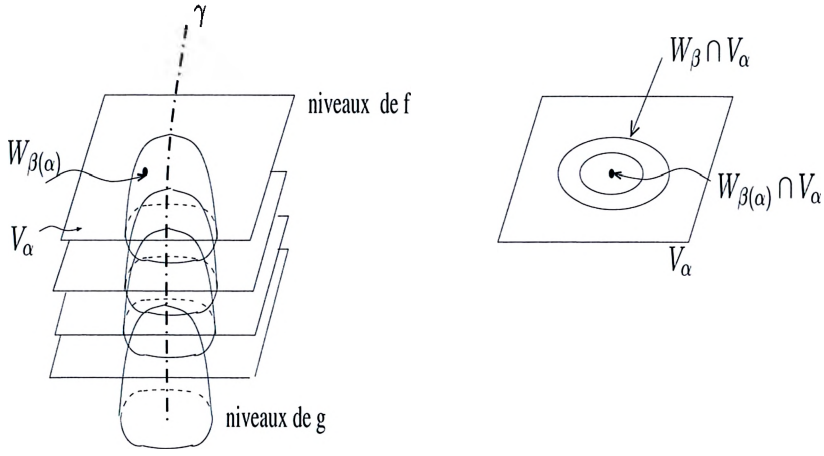


Figure 3.6. Dans la partie gauche de la figure, α et β sont variables. La partie droite de la figure montre les intersections (du type centre) des surfaces W_β (β variable) avec une même surface V_α (α fixé).

Deuxième partie

Théorie élémentaire des équations différentielles

1

GÉNÉRALITÉS

Dans toute cette partie II, comme le suggère le qualificatif élémentaire, on ne considérera, sauf mention expresse du contraire, que des champs de vecteurs définis sur des ouverts euclidiens. Une théorie plus générale sera abordée dans le tome 2.

1.1. Définition des champs de vecteurs

Soit U un ouvert de $\mathbb{R}^n$. On désigne par $\mathcal{C}^\infty(U)$ l'espace des fonctions de classe $\mathcal{C}^\infty$ et à valeurs réelles.

Définition 1.1. Soit $(X_1(x),\dots,X_n(x)) : U \mapsto \mathbb{R}^n$, une application de classe $\mathcal{C}^\infty$. On associe à cette application un *opérateur de dérivation* sur l'espace $\mathcal{C}^\infty(U)$:

$$f \in \mathcal{C}^\infty(U) \longrightarrow X \cdot f \in \mathcal{C}^\infty(U)$$

défini par

$$X \cdot f(x) = \sum_{i=1}^{n} X_i(x) \frac{\partial f}{\partial x_i}(x).$$

Le symbole X désigne donc l'opérateur de dérivation : $X = \sum_{i=1}^n X_i \frac{\partial}{\partial x_i}$. On dit que X est un *champ de vecteurs* sur U. Cette terminologie traduit le fait que X est la donnée d'un vecteur $X(x) = (X_1(x),\dots,X_n(x))^T \in \mathbb{R}^n$ en chaque point $x \in U$. On notera plutôt ce vecteur $X(x) = X_i(x)\frac{\partial}{\partial x_i}$ pour rappeler la définition de X comme opérateur de dérivation. Cette notation revient à identifier $\frac{\partial}{\partial x_i}$ avec le i-ème vecteur de la base canonique de $\mathbb{R}^n$.

Remarque 1.1.

1. La dérivation $X \cdot f$ est aussi appelée *dérivée de Lie de f par le champ de vecteurs X*. Elle est alors notée $L_X f$. C'est cette terminologie de dérivée de Lie qui est utilisée pour l'extension de l'opérateur de dérivation aux champs différentiels plus généraux que l'on peut définir sur U : formes différentielles, champs de tenseurs, métriques riemanniennes, etc.

2. On dit que X est de classe $\mathcal{C}^r$ si chaque composante X_i du champ est de classe $\mathcal{C}^r$. Dans ce cas, l'opérateur X envoie l'espace $\mathcal{C}^r(U)$ des fonctions de classe $\mathcal{C}^r$ dans l'espace $\mathcal{C}^{r-1}(U)$. On a donc un changement d'espace, que l'on évitera en considérant les champs de classe $\mathcal{C}^\infty$, ce que l'on fera, sauf mention expresse du contraire.

3. Il est important, pour traiter les applications, de considérer des ouverts $U \neq \mathbb{R}^n$, et ceci pour pouvoir écarter les points où X devient infini (figure 1.1). Par exemple, considérons le potentiel newtonien relatif à k particules. Chacune des particules est repérée par six variables : les coordonnées de sa position et les composantes de sa vitesse. Donc, le système des k particules est représenté par un point de $\mathbb{R}^{6k}$. Cependant on doit écarter l'ensemble fermé $\Delta \subset \mathbb{R}^{6k}$ où 2 particules coïncident et prendre $U = \mathbb{R}^{6k} - \Delta$ (figure 1.1).

FIGURE 1.1. Collision de deux particules x_1 et x_2.

4. On supposera toujours que le domaine de définition U d'un champ de vecteurs est un ensemble *ouvert* d'un espace euclidien $\mathbb{R}^n$. Cette condition est indispensable pour pouvoir parler aisément de propriétés liées à la différentiabilité.

Le lemme suivant va nous permettre d'introduire une vision beaucoup plus géométrique des champs de vecteurs, justifiant d'ailleurs la terminologie utilisée.

Lemme 1.1. *On considère un champ de vecteurs X sur un ouvert $U \subset \mathbb{R}^n$. Soit $x \in U$ et $\psi(t)$ un arc de courbe différentiable, $\psi : t \in I \mapsto U$ défini sur un intervalle ouvert I contenant 0, tel que $\psi(0) = x$ et $\frac{d\psi}{dt}(0) = \psi'(0) = X(x) \in \mathbb{R}^n$ (on peut dire qu'au temps $t = 0$ l'arc réalise le vecteur $X(x)$, valeur du champ au point x). Alors*

$$X \cdot f(x) = \sum_{i=1}^{n} X_i(x) \frac{\partial f}{\partial x_i}(x) = df(x)[X(x)] = \frac{d}{dt}(f \circ \psi)(0). \qquad (1.1)$$

Démonstration. La formule $\sum_{i=1}^n X_i(x)\frac{\partial f}{\partial x_i}(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x)X_i(x)$ est justement l'expression de $df(x)[X(x)]$ donnée dans l'équation (I-1.5). D'autre part, par le théorème I-1.1 de différentiation des applications composées, on obtient que

$$\frac{d}{dt}(f \circ \psi)(t) = df(\psi(t)) \circ d\psi(t)[1] = df(\psi(t))[\frac{d\psi}{dt}(t)],$$

ce qui donne le résultat en prenant $t = 0$. ∎

Remarque 1.2. Il suit du lemme que la dérivée de Lie $X \cdot f(x)$ peut se voir en chaque point x comme la dérivation ponctuelle le long du vecteur $X(x)$ (plus exactement le long d'un arc réalisant ce vecteur). Cette observation justifie la terminologie de champ de vecteurs pour désigner l'opérateur. Si l'on introduit le champ gradient de la fonction f défini par $\nabla f(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x)\frac{\partial}{\partial x_i}$, on voit que $X \cdot f(x)$ s'interprète comme le produit scalaire $< X(x), \nabla f(x) >$.

1.2. Image d'un champ de vecteurs par un difféomorphisme

Soit $y = G(x)$ un difféomorphisme de classe $\mathcal{C}^{k+1}, k \geq 1$, de l'ouvert $U \subset \mathbb{R}^n$, de coordonnée x, sur l'ouvert $V \subset \mathbb{R}^n$, de coordonnée y. Un tel difféomorphisme permet de transporter de façon isomorphe l'espace $\mathcal{C}^{k+1}(V)$ sur l'espace $\mathcal{C}^{k+1}(U)$ par l'application

$$G^* : \mathcal{C}^{k+1}(V) \longmapsto \mathcal{C}^{k+1}(U)$$

définie par la composition $G^*(f)(x) = f \circ G(x)$. Cette application G^* est tout simplement l'action du changement de coordonnées G sur l'espace des fonctions $\mathcal{C}^{k+1}(V)$.

On va maintenant définir le transport d'un champ de vecteurs X par le difféomorphisme G :

Définition 1.2. Soit $G : U \mapsto V$ un difféomorphisme de classe $\mathcal{C}^{k+1}$ et $X = \sum_{i=1}^n X_i \frac{\partial}{\partial x_i}$ un champ de vecteurs de classe $\mathcal{C}^k$ sur U. Alors le champ de vecteurs Y de classe $\mathcal{C}^k$, défini pour tout $y \in V$, par

$$Y(y) = dG(G^{-1}(y))[X(G^{-1}(y))], \tag{1.2}$$

est le *champ transporté du champ X par le difféomorphisme G*. On le note $Y = G_*(X)$.

Proposition 1.1. *Soit le champ Y transporté du champ X par le difféomorphisme G défini par la définition 1.2. Alors le champ Y vérifie les deux propriétés suivantes.*

1. *Pour tout $f \in C^{k+1}(V)$, $Y \cdot f \ (y) = X \cdot (f \circ G) \ (G^{-1}(y))$.*
2. *Soit ψ un arc de courbe différentiable réalisant le vecteur $X(x)$, comme dans le lemme 1.1. Alors le vecteur $Y(G(x))$ est réalisé par l'arc transporté $\widetilde{\psi} = G \circ \psi$. Autrement dit :*

$$Y(G(x)) = \frac{d\widetilde{\psi}}{dt}(0). \tag{1.3}$$

Démonstration. Démontrons tout d'abord le point (2). Si $Y(G(x)) = dG(x)[X(x)]$ et ψ est un arc réalisant $X(x)$, on a

$$Y(G(x)) = dG(x)\big[\frac{d\psi}{dt}(0)\big] = \frac{d\widetilde{\psi}}{dt}(0),$$

où $\widetilde{\psi} = G \circ \psi$, ce qui est la formule souhaitée. Comme évidemment $\widetilde{\psi}(0) = G(\psi(0)) = G(x)$, le vecteur $Y(G(x))$ est réalisé par l'arc transporté $\widetilde{\psi} = G \circ \psi$. Le point (1) est alors une conséquence du lemme 1.1, car la formule (1.1) est licite pour le champ Y avec $\widetilde{\psi}$ dans le rôle de ψ :

$$Y \cdot f(y) = \frac{d}{dt}(f \circ \widetilde{\psi})(0) = \frac{d}{dt}(f \circ G \circ \psi)(0) = X \cdot (f \circ G)(x),$$

ce qui démontre le résultat puisque $x = G^{-1}(y)$. ∎

Remarque 1.3. Le point (1) de la proposition énonce que

$$G_*(X) \cdot f = X \cdot G^*(f).$$

Le transport des champs de vecteurs par G_* est donc l'opération duale du transport des fonctions par G^*. L'opérateur G_* est une bijection et : $(G_*)^{-1} = (G^{-1})_*$.

Exemple. Considérons l'application de coordonnées polaires

$$\Phi : (r, \theta) \mapsto (x, y) = (r\cos\theta, r\sin\theta),$$

restreinte à l'ouvert $U = \{0 < r, \ \theta \in]0, 2\pi[\}$ pour en faire un difféomorphisme de U sur son image $V = \mathbb{R}^2 \setminus Ox$ (Ox désigne le demi-axe des x positifs).

Soit $Y(x,y) = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$, on veut trouver le champ

$$X(r,\theta) = \alpha(r,\theta)\frac{\partial}{\partial\theta} + \beta(r,\theta)\frac{\partial}{\partial r} \tag{1.4}$$

sur U tel que $\Phi_*(X) = Y$. Pour ce faire, on introduit un arc : $\psi(t) = (r(t),\theta(t))$ réalisant $X(r,\theta)$ en un point (r,θ) quelconque de U. En écrivant $\dot{g}$ la dérivée $\frac{dg}{dt}$ d'une fonction g de t et en utilisant le lemme 1.1, on a :

$$\dot{r}(0) = \beta(r,\theta), \quad \dot{\theta}(0) = \alpha(r,\theta).$$

L'arc image par Φ est $\widetilde{\psi}(t) = (r(t)\cos\theta(t), r(t)\sin\theta(t))$. Par dérivation, on obtient (on omet l'indication de la variable t pour alléger l'écriture) :

$$\dot{x} = \dot{r}\cos\theta - \dot{\theta}r\sin\theta, \quad \dot{y} = \dot{r}\sin\theta + \dot{\theta}r\cos\theta.$$

En prenant $t = 0$, grâce à la formule (1.3) (le rôle de G est joué par Φ, la première (resp. seconde) composante du champ Y est x (resp. y)) on obtient, par identification, un système d'équations pour $\alpha = \alpha(r,\theta), \beta = \beta(r,\theta)$:

$$x = \beta\cos\theta - r\alpha\sin\theta, \quad y = \beta\sin\theta + r\alpha\cos\theta.$$

Ce système se résout en $\alpha = 0, \beta = r$, ce qui donne par (1.4) $X(r,\theta) = r\frac{\partial}{\partial r}$.

Remarque 1.4. Il est peu satisfaisant de devoir se restreindre à un ouvert V comme on l'a fait dans l'exemple ci-dessus : cet ouvert V est le plan $\mathbb{R}^2$ avec une coupure le long de l'axe Ox, et, de plus, cette coupure est arbitraire. En fait, Φ est un difféomorphisme entre la surface $\mathbb{R}^+ \times S^1$ sur laquelle vit la variable (r,θ) et $\mathbb{R}^2 - \{0\}$ ($\mathbb{R}^+$ est difféomorphe à $\mathbb{R}$ et on identifie $S^1 = \mathbb{R}/\mathbb{Z}$ avec $\mathbb{R}/2\pi\mathbb{Z}$ par l'application $t \mapsto \theta(t) = 2\pi t$). Le champ de vecteurs $X(r,\theta) = r\frac{\partial}{\partial r}$ est invariant par rotation et, en conséquence, il définit un champ de vecteurs sur $\mathbb{R}^+ \times S^1$, ce qui permet de considérer la formule $\Phi_*(X) = Y$ sur cette surface toute entière. On doit évidemment définir ce que l'on appelle un champ de vecteurs sur une variété. Nous reviendrons sur cette question dans le tome 2. Ici, les champs de vecteurs sur $\mathbb{R}^+ \times S^1$ s'identifient aux champs $\alpha(r,\theta)\frac{\partial}{\partial\theta} + \beta(r,\theta)\frac{\partial}{\partial r}$ invariants par la translation $(r,\theta) \mapsto (r,\theta+2\pi)$, c'est-à-dire tels que $\alpha(r,\theta) = \alpha(r,\theta+2\pi)$ et $\beta(r,\theta) = \beta(r,\theta+2\pi)$.

1.3. Équation différentielle d'un champ de vecteurs

Le lemme 1.1 montre l'intérêt que l'on a à considérer un arc de courbe réalisant, en $t = 0$ et en un point x, la valeur du vecteur $X(x)$: un tel arc permet de calculer la dérivation par ce vecteur. Nous allons maintenant introduire des courbes $\varphi(t)$ réalisant le champ de vecteurs *pour tout* t.

Définition 1.3. Soit X un champ de vecteurs défini sur un ouvert $U \subset \mathbb{R}^n$. L'équation différentielle de X est l'équation

$$\frac{d\varphi}{dt}(t) = X(\varphi(t)). \tag{1.5}$$

Ici $\varphi(t)$ est une courbe différentiable, définie sur un intervalle ouvert $I \subset \mathbb{R}$. On l'appelle *courbe intégrale* de l'équation différentielle (ou du champ de vecteurs). La variable t est en général appelée temps (allusion aux équations de la mécanique). On appelle *trajectoire* du champ, toute courbe intégrale telle que $0 \in I$; la valeur $\varphi(0)$ est appelée *condition initiale* de la trajectoire.

Un champ de vecteurs peut modéliser le champ des vitesses d'évolution d'un système dont les points représentatifs appartiennent à U. Dans ce cas, chaque trajectoire du champ X est une évolution à partir de la condition initiale $x = \varphi(0)$. Cette interprétation, en particulier en mécanique, explique la terminologie que nous utilisons : *temps*, pour désigner la variable des courbes intégrales appelées *trajectoires*.

Remarque 1.5.

1. Pour une trajectoire paramétrée par le temps t, le vecteur dérivé $\frac{d\varphi}{dt}(t)$ (vitesse) est en général noté $\dot{\varphi}(t)$. De même, l'accélération $\frac{d^2\varphi}{dt^2}(t)$ est notée $\ddot{\varphi}(t)$ (notations de Leibniz).

2. Soit $\varphi : t \in I \mapsto U$ une courbe intégrale du champ de vecteurs X et $\tau \in I$ une valeur quelconque du temps. On pose $\varphi(\tau) = x_0$. Alors, τ étant fixée, la courbe $\psi : t \in J = I \setminus \{\tau\} \mapsto \psi(t) = \varphi(\tau + t) \in U$ est une trajectoire par x_0. Pour le vérifier, on introduit sur J la fonction $u(t) = t + \tau$, ce qui permet d'écrire $\psi = \varphi \circ u$. On a évidemment $\frac{\partial u}{dt}(t) \equiv 1$ et donc, en utilisant le théorème de composition des différentielles :

$$\frac{d\psi}{dt}(t) = \frac{d\varphi}{dt}\big(u(t)\big) \cdot \frac{du}{dt}(t) = \frac{d\varphi}{dt}\big(u(t)\big). \tag{1.6}$$

Comme φ est une courbe intégrale, on a :

$$\frac{d\varphi}{dt}\big(u(t)\big) = X\big(\varphi \circ u(t)\big) = X\big(\psi(t)\big). \tag{1.7}$$

En comparant (1.6) et (1.7), on obtient que $\frac{d\psi}{dt}(t) = X\big(\psi(t)\big)$ pour tout $t \in J$. Comme $\psi(0) = \varphi(\tau) = x_0$, la courbe ψ est donc bien une trajectoire de X par le point x_0.

On peut donc se limiter aux trajectoires du champ : toute courbe intégrale se déduit d'une trajectoire *par translation dans le temps.* C'est ce que nous ferons dorénavant.

3. Le point (2) de la proposition 1.1 nous dit que si G est un difféomorphisme de U sur V envoyant le champ de vecteurs X sur le champ $Y = G_*(X)$, alors G envoie de façon bijective les trajectoires de X sur celles de Y.

1.4. Équations différentielles générales

Considérons un ouvert U de $\mathbb{R}^{1+n\ell+m} = \mathbb{R} \times \mathbb{R}^n \times \ldots \times \mathbb{R}^n \times \mathbb{R}^m$ et une application F de classe C^k de U à valeurs dans $\mathbb{R}^n$. On peut associer à cette application l'équation différentielle suivante :

$$\varphi^{(\ell)} = F(t, \varphi(t), \ldots, \varphi^{(\ell-1)}(t), \lambda), \tag{1.8}$$

où $\varphi(t)$, appelée solution de (1.8), est une fonction vectorielle définie sur un intervalle ouvert I, à valeur dans $\mathbb{R}^n$, ℓ fois différentiable et λ est un paramètre dans $\mathbb{R}^m$. On a noté $\frac{d^s\varphi}{dt^s}$ par $\varphi^{(s)}$. Si on considère les composantes de F sur $\mathbb{R}^n$, l'équation (vectorielle) (1.8) est équivalente à un *système différentiel* de n équations en dimension 1 (c'est une terminologie ancienne ; nous parlerons plutôt d'équation différentielle, omettant le qualificatif de vectorielle).

Définition 1.4. Le nombre n est la *dimension* de l'équation (1.8). On dit que celle-ci est d'*ordre* l (l'équation différentielle d'un champ de vecteurs est d'ordre 1). Si F ne dépend pas du temps t (dépendance triviale !), on dit que l'équation est *autonome.* Une équation non autonome est donc une équation dépendant non trivialement du temps. Enfin λ est le *paramètre.* Pour chaque valeur de λ on a donc une équation différentielle, et lorsque l'équation dépend non trivialement d'un paramètre, on dit souvent que l'on a affaire à une *famille* d'équations différentielles.

Nous aurons le loisir de revenir sur ces notions à de multiples occasions. Ici, dans cette introduction, nous allons nous contenter de montrer qu'une équation différentielle générale se ramène toujours formellement à l'équation différentielle d'un champ de vecteurs, c'est-à-dire à une équation d'ordre 1, autonome, sans paramètre. Il suffit pour cela d'ajouter des variables supplémentaires dans l'espace d'arrivée de la fonction inconnue, ainsi que des lignes supplémentaires d'équation. Précisément, on introduit une nouvelle équation où le temps

est désigné par τ, t étant maintenant considéré comme une fonction de τ. Soit : $\tau \mapsto (t(\tau), \varphi_0(\tau), \ldots, \varphi_{\ell-1}(\tau), \lambda(\tau))$, solution de l'équation différentielle :

$$\begin{cases} \dfrac{dt}{d\tau} = 1 \\ \dfrac{d\lambda}{d\tau} = 0 \\ \dfrac{d\varphi_0}{d\tau} = \varphi_1 \\ \dfrac{d\varphi_1}{d\tau} = \varphi_2 \\ \vdots \\ \dfrac{d\varphi_{\ell-1}}{d\tau} = F(t, \varphi_0, \varphi_1, \ldots, \varphi_{\ell-1}, \lambda) \end{cases} \tag{1.9}$$

Cette équation différentielle (1.9) est celle d'un champ de vecteurs sur l'ouvert U. Il y a une correspondance bi-univoque entre les courbes intégrales de (1.9) et les solutions de (1.8). En effet, si $\varphi(t)$ est une solution de (1.8), $\psi(\tau) = (\tau, \varphi(\tau), \varphi^{(1)}(\tau), \ldots, \varphi^{(\ell-1)}(\tau), \lambda)$ est une courbe intégrale de (1.9). Inversement, si $\psi(\tau)$ est une courbe intégrale de (1.9), $\lambda(\tau)$ est constant, $t = \tau$ à une constante additive près, et la composante $\varphi_0(t)$ est solution de (1.8), car les composantes φ_i sont les dérivées successives de φ_0, comme l'imposent les premières lignes de (1.9), et la substitution de ces fonctions dans la dernière ligne donne précisément l'équation (1.8).

Exemple (équations de la mécanique). Les équations de la mécanique classique sont des équations différentielles d'ordre 2 pouvant être non autonomes (par exemple dans le cas d'un mouvement représenté dans un repère lui-même en mouvement), et pouvant dépendre de paramètres (masses...). Supposons pour simplifier que l'équation soit autonome et sans paramètre. Elle prend la forme

$$\frac{d^2\varphi}{dt^2}(t) = F(\varphi(t), \frac{d\varphi}{dt}(t)).$$

Le second membre de l'équation est le champ des forces (aux masses près). Ce champ des forces dépend ici de la position φ et de la vitesse $\frac{d\varphi}{dt}$ du système, et l'équation nous dit que l'accélération $\frac{d^2\varphi}{dt^2}$ est déterminée par le champ des forces. Le passage de l'équation (1.8) à l'équation (1.9) est ce qu'on appelle en mécanique le passage (ou l'écriture) du système dans l'*espace des positions-vitesses,* espace que nous appellerons improprement l'*espace de phase*, comme il est classique en mécanique. Le passage dans l'espace de phase consiste donc à introduire la vitesse comme variable indépendante à côté de la position.

Remarque 1.6. Cette dénomination d'*espace de phase* vient du fait que, dans le formalisme hamiltonien de la mécanique, on remplace les variables de positions-vitesses par des variables généralisées qui sont dans les cas classiques les variables angles-actions (par exemple l'angle d'un pendule ; l'action est liée à l'énergie). Le terme *phase* fait allusion à ces variables angles.

2

CHAMPS DE VECTEURS LINÉAIRES

L'intégration et l'étude des propriétés des champs de vecteurs linéaires se ramènent à des questions d'algèbre linéaire. Ces champs sont d'une grande importance pratique car ils sont utilisés pour la modélisation de beaucoup de systèmes simples : oscillations linéaires de la mécanique, circuits électriques, etc. Les champs de vecteurs linéaires forment une des seules classes de champs que l'on peut intégrer explicitement. Ils nous serviront aussi à introduire et à illustrer les propriétés et concepts des systèmes non linéaires qui seront étudiés dans la suite : existence et unicité des trajectoires, flot, points singuliers, diagramme de phase, etc.

2.1. Étude théorique

Définition 2.1. Soit $A = (a_{ij})_{ij}$ une matrice carrée de dimension n (on écrira $A \in \mathrm{Mat}(n,n)$, espace des matrices réelles carrées de dimension n). On lui associe un champ de vecteurs

$$X(x_1,\ldots,x_n) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} x_j \right) \frac{\partial}{\partial x_i}. \tag{2.1}$$

L'équation des trajectoires de ce champ est

$$\dot{\varphi}(t) = A\varphi(t). \tag{2.2}$$

Ici, $A\varphi(t)$ désigne l'application de la matrice $A \in \mathrm{Mat}(n,n)$ sur le vecteur $\varphi(t) \in \mathbb{R}^n$, ou bien si l'on préfère, le produit de la matrice par le vecteur

colonne. Cette équation vectorielle est équivalente aux n équations

$$\dot{\varphi}_i = \sum_{j=1}^{n} a_{ij}\varphi_j, \quad \text{pour} \quad i = 1, \ldots, n.$$

Avant de donner l'expression du flot d'un champ linéaire, nous allons rappeler quelques faits basiques concernant les fonctions analytiques à valeur matricielle et en particulier l'exponentielle complexe.

Définition 2.2. Soit $A \in \mathrm{Mat}(n,n)$. On appelle norme d'opérateur de A (associée à la norme euclidienne) la norme définie par

$$||A|| = \mathrm{Sup}\{||A(x)|| \mid x \in \mathbb{R}^n, ||x|| = 1\},$$

où $||x||$ désigne la norme euclidienne de $\mathbb{R}^n$.

On montre facilement que $||A||$ est bien une norme sur $\mathrm{Mat}(n,n)$. On dit que cette norme est la *norme matricielle induite* par la norme vectorielle $\|.\|$. L'intérêt de cette norme par rapport à d'autres normes classiques que l'on peut définir sur $\mathrm{Mat}(n,n)$ est qu'elle vérifie trivialement

$$||A(x)|| \leq ||A|| \cdot ||x|| \quad \text{et} \quad ||A \cdot B|| \leq ||A|| \cdot ||B||,$$

pour tout $x \in \mathbb{R}^n$, quelles que soient $A, B \in \mathrm{Mat}(n,n)$. On peut montrer que $||A||$ est égale à la borne supérieure des modules des valeurs propres de A, si A est une matrice normale ($AA^T = A^TA$ dans le cas réel). On ne prouvera pas ce résultat non trivial car on n'en fera pas usage dans cet ouvrage.

Soit $f(z)$ la somme d'une série de puissances de la variable complexe z,

$$f(z) = a_0 + a_1 z + a_2 z^2 + \ldots \qquad (|z| < r),$$

dont le rayon de convergence est $r > 0$. Soit $B \in \mathrm{Mat}(n,n)$ telle que sa norme d'opérateur vérifie $||B|| < r$, alors la série

$$a_0 I + a_1 B + a_2 B^2 + \ldots$$

est absolument convergente. En effet, le terme général $a_n B^n$ a une norme (d'opérateur) majorée par $|a_n|||B||^n$ qui est le terme général d'une série numérique absolument convergente. On considérera ici la série exponentielle $e^z = e^z = 1 + z + \frac{z^2}{2!} + \ldots + \frac{z^i}{i!} + \ldots$ dont le rayon de convergence est infini. On peut donc définir pour tout $B \in \mathrm{Mat}(n,n)$, l'exponentielle matricielle $e^B = I + B + \frac{B^2}{2!} + \ldots + \frac{B^i}{i!} + \ldots$

On peut étendre la propriété fonctionelle de l'exponentielle numérique : $e^{a+b} = e^a \cdot e^b$, à l'exponentielle matricielle à la condition que les matrices $A, B \in \mathrm{Mat}(n,n)$ commutent :

$$\text{si } [A,B] = A \cdot B - B \cdot A = 0, \quad \text{alors } e^{A+B} = e^A \cdot e^B. \tag{2.3}$$

Ce résultat peut s'établir directement par regroupement de termes dans les séries. La formule de Campbell-Hausdorff est l'expression générale que l'on peut écrire si A et B ne commutent pas. Donc, en général, on a $e^{A+B} \neq e^A \cdot e^B$. Une conséquence importante de cette propriété fonctionnelle de l'exponentielle est que, pour tout $B \in \mathrm{Mat}(n,n)$, on a $e^B \in \mathrm{Gl}(n,\mathbb{R})$ (le groupe linéaire des automorphismes de $\mathbb{R}^n$, c'est-à-dire le groupe pour la composition des matrices de $\mathrm{Mat}(n,n)$ inversibles). En effet $I = e^0 = e^{B-B} = e^B \cdot e^{-B}$. L'inverse de la matrice e^B est donc la matrice e^{-B}.

Pour tout $A \in \mathrm{Mat}(n,n)$, on va considérer la fonction de la variable réelle t à valeur matricielle : $t \mapsto e^{tA}$. Cette série converge normalement sur tout intervalle $[-T,T]$. Sur un tel intervalle, cette série est dérivable terme à terme, ainsi que toutes ses dérivées successives. On dit que cette fonction est analytique. Comme on peut prendre T quelconque, $t \mapsto e^{tA}$ est une fonction analytique sur $\mathbb{R}$ tout entier.

Il est possible de *calculer* les trajectoires d'un champ de vecteurs linéaire à l'aide de cette exponentielle matricielle.

Théorème 2.1. *Soit X le champ de vecteurs linéaire associé à la matrice $A \in \mathrm{Mat}(n,n)$. Alors, par tout $x \in \mathbb{R}^n$, il passe une et une seule trajectoire $\varphi(t,x)$, définie pour tout $t \in \mathbb{R}$. Cette trajectoire est donnée par la formule*

$$\varphi(t,x) = e^{tA}x. \tag{2.4}$$

Démonstration. Fixons un $T > 0$. La série $\sum \frac{t^i A^i}{i!}$ est dérivable terme à terme (car $\frac{d}{dt}(\frac{t^i A^i}{i!}) = \frac{t^{i-1}A^i}{(i-1)!}$ est terme général d'une série normalement convergente). On a donc

$$\frac{d}{dt}(e^{tA}) = \sum_{i=1}^{\infty} \frac{t^{i-1}A^i}{(i-1)!} = A \cdot e^{tA} = e^{tA} \cdot A.$$

Soit maintenant $x \in \mathbb{R}^n$ quelconque. Considérons la courbe $\varphi(t,x) = e^{tA}x$. On a

$$\frac{\partial \varphi}{\partial t}(t,x) = \frac{\partial}{\partial t}(e^{tA}x) = \frac{d}{dt}(e^{tA})x = A \circ e^{tA}x = A\varphi(t,x) \quad \text{et} \quad \varphi(0,x) = x,$$

ce qui signifie que $\varphi(t,x)$ est une trajectoire du champ par x, définie pour tout $t \in \mathbb{R}$. Montrons maintenant que $\varphi(t) = \varphi(t,x)$ est unique avec cette propriété.

Pour ce faire, considérons une deuxième trajectoire $\psi(t)$ par x, définie pour tout t. Comme la matrice e^{tA} est inversible pour tout t, d'inverse e^{-tA}, on peut écrire ψ sous la forme $\psi(t) = e^{tA}c(t)$ où $c(t) = e^{-tA}\psi(t)$ est une courbe différentiable définie sur $\mathbb{R}$. Écrivons que ψ est trajectoire par x

$$\frac{d\psi}{dt} = A \circ e^{tA}c + e^{tA}\frac{dc}{dt} \Longrightarrow \frac{dc}{dt} \equiv 0 \Longrightarrow c(t) \equiv c(0) = x.$$

On a donc $\psi(t) \equiv \varphi(t, x)$. ∎

Quelques propriétés utiles de l'exponentielle e^{tA}

Dans la démonstration ci-dessus, on a utilisé quelques propriétés de l'application exponentielle $t \mapsto e^{tA}$, propriétés que nous allons rappeler :

1. A commute avec e^{tA}, ce qui est une conséquence directe de la commutation de A avec A^i pour tout i et du passage à la limite dans la sommation de la série.

2. $\frac{d}{dt}(e^{tA}) = A \cdot e^{tA}$, formule établie ci-dessus par dérivation sous le signe somme. Il en résulte que la fonction matricielle $R(t) = e^{tA}$ est solution de l'équation différentielle matricielle

$$\frac{dR}{dt} = A \cdot R, \tag{2.5}$$

avec la condition initiale $R(0) = I$. Cette fonction matricielle est appelée *résolvante* de l'équation différentielle (2.2) car la solution de l'équation différentielle (2.2) passant par x s'écrit $\varphi(t, x) = e^{tA}x = R(t)x$, ceci quel que soit x.

3. $e^{(t+\tau)A} = e^{tA} \circ e^{\tau A}$. Cette formule peut être vue comme une conséquence de (2.3) puisque les matrices tA et τA commutent. On peut l'établir plus astucieusement sans utiliser l'argument de commutativité, en remarquant que les deux courbes $t \mapsto e^{(t+\tau)A}x$ et $t \mapsto e^{tA} \circ e^{\tau A}x$ sont trajectoires de (2.2) par x, pour tout x, puis en utilisant le résultat d'unicité établi dans le théorème. Il en résulte que pour tout t et τ, les matrices e^{tA} et $e^{\tau A}$ commutent.

4. e^{tA} est une matrice inversible pour tout $t \in \mathbb{R}$. En effet, en appliquant le point précédent, on a : $e^{tA} \circ e^{-tA} = e^{0A} = I$: la matrice e^{tA} est inversible, d'inverse e^{-tA}. Ces matrices e^{tA} forment un sous-groupe du groupe $\mathrm{Gl}(n, \mathbb{R})$ des matrices inversibles, appelé groupe à un paramètre.

Remarque 2.1. L'équation (2.5) peut paraître plus compliquée que l'équation (2.2) puisque l'on passe de n équations dans $\mathbb{R}^n$ à n^2 équations dans $\mathrm{Mat}(n,n)$. Cependant, il est utile de considérer (2.5) parce que tous ses éléments sont dans $\mathrm{Mat}(n,n)$ et qu'on peut maintenant exploiter toute la structure d'algèbre de cet espace.

Lemme 2.1. *Soit $R(t)$ une solution quelconque de l'équation différentielle matricielle (2.5). Soit $W(t) = \det(R(t))$, la fonction déterminant, encore appelée* Wronskien *de l'équation différentielle. Alors $W(t)$ est solution de l'équation différentielle $\frac{dW}{dt} = (\mathrm{Trace}A) \cdot W$, et est donc égale à*

$$W(t) = W(0)e^{(\mathrm{Trace}A)t}. \tag{2.6}$$

Cette formule s'appelle formule de Jacobi-Liouville.

Démonstration. L'équation différentielle (2.5) s'écrit

$$\frac{dR_{ij}}{dt} = \sum_k a_{ik}R_{kj} \quad \text{pour} \quad i,j = 1,\dots,n.$$

Désignons par $\Psi_i = (R_{i1},\dots,R_{in})$ la i-ème ligne de la matrice R. On peut réécrire l'équation (2.2) en lignes :

$$\frac{d\Psi_i}{dt} = \sum_k a_{ik}\Psi_k \quad \text{pour} \quad i = 1,\dots,n. \tag{2.7}$$

Maintenant, $W(t) = \det(\Psi_1,\dots,\Psi_n)$ est une forme n-linéaire par rapport aux lignes $\Psi_1,\dots,\Psi_n$. Il en résulte que

$$\frac{dW}{dt} = \det(\frac{d\Psi_1}{dt},\Psi_2,\dots,\Psi_n) + \det(\Psi_1,\frac{d\Psi_2}{dt},\dots,\Psi_n) + \dots$$
$$\dots + \det(\Psi_1,\Psi_2,\dots,\frac{d\Psi_n}{dt}),$$

soit en remplaçant par (2.7) :

$$\frac{dW}{dt} = \det(\sum_k a_{1k}\Psi_k,\Psi_2,\dots,\Psi_n) + \det(\Psi_1,\sum_k a_{2k}\Psi_k,\dots,\Psi_n) + \dots$$

En développant cette formule, et en utilisant le fait que le déterminant est une forme alternée sur ses lignes, ce qui entraîne que le déterminant est nul dès que deux de ses lignes sont égales, on obtient finalement

$$\frac{dW}{dt} = \Big(\sum_{i=1}^{n} a_{ii}\Big) \cdot W = (\mathrm{Trace}A) \cdot W.$$

La formule (2.6) s'obtient par intégration, compte tenu du théorème 2.1. ∎

Le théorème 2.1 nous permet d'affirmer que les solutions définies sur $\mathbb{R}$ (ou si l'on préfère, les trajectoires définies sur $\mathbb{R}$) forment l'ensemble $\mathcal{S} = \{e^{tA}x | x \in \mathbb{R}^n\}$. On peut faire les remarques suivantes :

1. L'ensemble $\mathcal{S}$ s'identifie au noyau de l'application linéaire

$$\varphi \mapsto \dot{\varphi} - A\varphi.$$

C'est donc un espace vectoriel sur $\mathbb{R}$. Cela peut aussi se vérifier directement sur la définition de $\mathcal{S}$.

2. L'application $x \in \mathbb{R}^n \mapsto \varphi(t,x) = e^{tA}x$ est un isomorphisme linéaire car e^{tA} est inversible. C'est un exemple du théorème d'existence et d'unicité de Cauchy que l'on démontrera en toute généralité dans le prochain chapitre.

3. Soit $R(t)$ la matrice résolvante de l'équation (2.2) et $\varphi_1(t), \ldots, \varphi_n(t)$ ses colonnes. De l'équation $\frac{dR}{dt} = A \circ R$, il suit que $\varphi_i \in \mathcal{S}$ pour tout $i = 1, \ldots, n$. Donc, pour tout $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$, la combinaison linéaire $\varphi = \sum_i c_i\varphi_i$ appartient à $\mathcal{S}$. Inversement, comme $R(0) = I$, les vecteurs $\{\varphi_1(0), \ldots, \varphi_n(0)\}$ forment une base de $\mathbb{R}^n$. Donc, pour tout $\varphi \in \mathcal{S}$, on peut résoudre par rapport à c l'équation : $x = \varphi(0) = \sum_i c_i\varphi_i(0)$ (c'est en fait l'équation $x = R(0)c$). Cela signifie que $\{\varphi_1, \ldots, \varphi_n\}$ est une base vectorielle de $\mathcal{S} = \mathbb{R}\{\varphi_1, \ldots, \varphi_n\}$.

4. Plus généralement, considérons n solutions $\psi_1, \ldots, \psi_n \in \mathcal{S}$ telles que l'ensemble $\{\psi_1(0), \ldots, \psi_n(0)\}$ forme une base de $\mathbb{R}^n$ (il faut et il suffit pour cela que la matrice $R(t)$ ayant pour colonnes les ψ_i, soit solution de l'équation différentielle matricielle (2.5) et que $R(0)$ soit inversible). Alors $\{\psi_1, \ldots, \psi_n\}$ est une base de $\mathcal{S}$. Un tel système de fonctions $\{\psi_1, \ldots, \psi_n\}$ est appelé *système fondamental de solutions* de l'équation différentielle (2.2). Chaque solution s'écrit $\varphi = \sum_i c_i\psi_i$ pour un certain choix des c_i qui s'appellent les *constantes d'intégration* de φ. Ces constantes d'intégration $(c_1, \ldots, c_n)$ sont à distinguer des *conditions initiales* $(x_1, \ldots, x_n)$. On passe du vecteur c des constantes d'intégration au vecteur x des conditions initiales par la formule $x = R(0)c$, formule que l'on doit inverser pour résoudre le problème inverse.

5. Pour chaque $t \in \mathbb{R}$, l'application $x \mapsto \varphi_t(x) = \varphi(t,x) = e^{tA}x$ est un isomorphisme linéaire de $\mathbb{R}^n$. De plus, on a $\varphi_t \circ \varphi_\tau = \varphi_{t+\tau}$ pour tous $t, \tau \in \mathbb{R}$. On a ici une illustration de ce que l'on appelle le flot : $t \mapsto \varphi_t$ d'un champ de vecteurs, notion qui sera introduite dans le chapitre 3, pour les champs de vecteurs quelconques.

Équation linéaire non autonome

On peut considérer plus généralement des champs de vecteurs linéaires non autonomes, c'est-à-dire avec des coefficients $a_{ij}(t)$ dépendant du temps. Une partie des remarques faites plus haut restera valable, en particulier le fait que l'espace de toutes les solutions définies sur $\mathbb{R}$ est un espace vectoriel de dimension n. Il n'est, par contre, pas possible de trouver – en général – une expression explicite pour ces solutions. On se reportera aux traités sur les équations différentielles ordinaires ([23] par exemple) pour en savoir plus. On peut également sortir du cadre des équations linéaires en considérant des équations affines, autonomes ou non. Un cas très proche de celui des équations différentielles linéaires à coefficients constants, sera celui d'une équation différentielle affine de la forme

$$\dot{\varphi}(t) = A\varphi + B(t), \tag{2.8}$$

où $A \in \mathrm{Mat}(n,n)$ et $B(t)$ est une fonction vectorielle, $B : t \in \mathbb{R} \mapsto B(t) \in \mathbb{R}^n$. La remarque essentielle est que si φ_1 et φ_2 sont deux solutions définies sur $\mathbb{R}$, alors $\varphi_2 - \varphi_1$ est solution sur $\mathbb{R}$ de l'équation linéaire associée : $\dot{\varphi} = A\varphi$. Il s'en suit immédiatement le résultat suivant :

Proposition 2.1. *L'espace des solutions de (2.8) est l'espace affine égal à*

$$\{\phi_0\} + \mathcal{S},$$

où ϕ_0 est une solution particulière de (2.8) et $\mathcal{S}$ est l'espace vectoriel des solutions sur $\mathbb{R}$ de l'équation linéaire associée.

Ainsi, pour résoudre (2.8), il suffit de lui trouver une solution particulière ϕ_0. C'est l'occasion pour nous d'introduire la très utile *méthode de variation des constantes* : on recherche ϕ_0 sous la forme $\phi_0(t) = e^{tA}x(t)$ (d'où la terminologie : on rend variable la constante x de l'équation autonome). Écrivons que $e^{tA}x(t)$ est solution de (2.8) :

$$\frac{d}{dt}(e^{tA}x(t)) = A \circ e^{tA}x(t) + e^{tA}\frac{dx(t)}{dt} = A \circ e^{tA}x(t) + B(t).$$

Soit

$$\frac{dx(t)}{dt} = e^{-tA}B(t),$$

qui se résout par une simple quadrature. On obtient finalement une solution particulière égale par exemple à

$$\Phi_0(t) = e^{tA}\int_0^t e^{-sA}B(s)ds. \tag{2.9}$$

Cette solution particulière vérifie $\Phi_0(0) = 0$, mais cela n'est pas une propriété essentielle : on peut appliquer la proposition 2.1 avec une solution Φ_0 quelconque. Un changement du choix de solution particulière va simplement se traduire par un changement de l'application entre constantes d'intégration et conditions initiales.

En pratique, il arrive de trouver directement une solution particulière sans avoir nécessairement recours à la formule intégrale (2.9). Par exemple, si l'équation différentielle vectorielle (2.8) provient d'une équation différentielle d'ordre n en dimension 1 de la forme :

$$\frac{d^n f}{dt^n}(t) + \sum_{i=1}^{n} a_{n-i} \frac{d^{n-i} f}{dt^{n-i}}(t) = b(t), \tag{2.10}$$

avec $b(t)$ une fonction polynomiale de degré k, on va chercher directement comme solution particulière, une fonction polynomiale de degré k (sans d'ailleurs avoir à passer par l'équation différentielle vectorielle d'ordre 1 associée).

Passage dans le complexe

Il est toujours loisible de considérer A comme une matrice opérant dans $\mathbb{C}^n$ et de rechercher les solutions $\varphi : t \in \mathbb{R} \mapsto \varphi(t) \in \mathbb{C}^n$; on dit que *l'on passe dans le complexe* (on pourrait même complexifier le temps, mais on entrerait alors dans un cadre théorique très différent). L'équation passée dans le complexe se résout exactement comme dans le réel. La solution générale est donnée par $e^{tA}x$ avec $x \in \mathbb{C}^n$ et toutes les autres considérations faites dans le cadre réel restent valables en remplaçant $\mathbb{R}$ par $\mathbb{C}$. L'intérêt de ce passage dans le complexe, même si l'on est seulement intéressé par les solutions réelles, est que le traitement gagne en généralité comme on va le voir dans le prochain paragraphe (en raison essentiellement du théorème fondamental de l'algèbre : *tout polynôme non constant a au moins une racine dans* $\mathbb{C}$). Une fois trouvées les solutions à valeurs complexes, il nous restera évidemment à rechercher parmi elles les solutions à valeurs réelles. Pour ce faire, on utilisera essentiellement la remarque suivante : si φ est une solution à valeurs complexes, alors $\varphi + \bar{\varphi}$ (où $\bar{z}$ désigne le complexe conjugué de z) est une solution réelle.

2.2. Résolution explicite

Dans cette section, nous allons utiliser l'algèbre linéaire pour rendre plus explicite l'expression théorique $e^{tA}x$ des solutions de l'équation différentielle (2.2), obtenue dans le théorème 2.1. Pour ce faire, nous allons examiner l'effet sur l'équation d'un changement linéaire des coordonnées de $\mathbb{R}^n$.

Lemme 2.2. *Soit $y = Px$ un changement linéaire des coordonnées de $\mathbb{R}^n$ (avec $P \in \mathrm{Gl}(n, \mathbb{R})$). Soit $x(t)$ une solution de l'équation (2.2) : $\dot{x}(t) = Ax(t)$. Alors $y(t) = Px(t)$ est solution de l'équation différentielle linéaire : $\dot{y}(t) = \bar{A}y(t)$ avec $\bar{A} = P \circ A \circ P^{-1}$ et inversement.*

Démonstration. On a $x(t) = P^{-1}y(t)$. Il s'en suit que

$$A \circ P^{-1}y(t) = \frac{dx}{dt} = P^{-1}\frac{dy}{dt},$$

et donc que

$$\frac{dy}{dt} = P \circ A \circ P^{-1}y(t).$$

On a utilisé ci-dessus la linéarité de la dérivation et de l'application $u \mapsto Au$. ∎

Les champs de vecteurs linéaires se comportent donc, par rapport aux changements linéaires des coordonnées, comme des endomorphismes linéaires (intuitivement, cela correspond au fait que la valeur $\varphi(t)$ d'une solution, ainsi que sa dérivée $\dot{\varphi}(t)$, sont des vecteurs de $\mathbb{R}^n$, entre lesquels l'équation différentielle établit une correspondance linéaire). On peut donc appliquer à l'équation linéaire la théorie de la *réduction des endomorphismes linéaires dans le groupe linéaire* (diagonalisation, mise sous forme de Jordan, etc.) et les notions associées (valeurs et vecteurs propres, etc.).

Voici tout d'abord un résultat général de réduction à la dimension 1.

Lemme 2.3. *Toute équation différentielle linéaire en dimension n se ramène, après passage dans le complexe, à la résolution de n équations affines non autonomes de dimension 1 de la forme $\dot{x}(t) = ax(t) + b(t)$, où les coefficients a sont les valeurs propres de la matrice A et les fonctions $b : t \in \mathbb{R} \mapsto b(t) \in \mathbb{C}$ se calculent par récurrence.*

Démonstration. Par un changement linéaire de coordonnées complexes, on peut réduire la matrice A à une matrice triangulaire inférieure, c'est-à-dire que l'on peut supposer que $a_{ij} = 0$ si $i < j$ (on peut utiliser la réduction standard par la méthode de Gauss). Les termes diagonaux a_{ii} sont alors nécessairement les valeurs propres de la matrice. Dans ces coordonnées de triangulation, l'équation différentielle prend la forme

$$\begin{cases} \dot{x}_1(t) = a_{11}x_1(t) \\ \dot{x}_2(t) = a_{21}x_1(t) + a_{22}x_2(t) \\ \vdots \\ \dot{x}_n(t) = a_{n1}x_1(t) + a_{n2}x_2(t) + \ldots + a_{nn}x_n(t) \end{cases} \tag{2.11}$$

Soit $(x_1, \ldots, x_n)$ la condition initiale de la trajectoire cherchée. On résout tout d'abord la première équation linéaire en $x_1(t)$. On obtient $x_1(t) = x_1 e^{a_{11}t}$. En reportant cette fonction $x_1(t)$ dans la deuxième équation, on obtient une équation affine sur la fonction $x_2 : t \in \mathbb{R} \mapsto x_2(t)$:

$$\dot{x}_2(t) = a_{22}x_2(t) + a_{21}x_1 e^{a_{11}t}.$$

On peut résoudre cette équation par la méthode de variation de la constante en dimension 1 (voir équation (2.9)), en définissant $b(t)$ par $a_{21}x_1 e^{a_{11}t}$. Le scalaire a_{22} jouant le rôle de la matrice A, on obtient

$$x_2(t) = e^{a_{22}t}(x_2 + a_{21}x_1 \int_0^t e^{(a_{11}-a_{22})s} ds),$$

c'est-à-dire

$$x_2(t) = e^{a_{22}t}(x_2 + a_{21}x_1 t)$$

si $a_{11} = a_{22}$, et

$$x_2(t) = e^{a_{22}t}\left(x_2 + \frac{a_{21}x_1}{a_{11} - a_{22}}(e^{(a_{11}-a_{22})t} - 1)\right)$$

si $a_{11} \neq a_{22}$. Par récurrence, si on suppose calculées, pour $\ell < n$, les composantes $x_1(t), \ldots, x_\ell(t)$ de la solution, on a une équation affine pour $x_{\ell+1}$:

$$\dot{x}_{\ell+1}(t) = a_{\ell+1,\ell+1}x_{\ell+1}(t) + b_{\ell+1}(t),$$

où $b_{\ell+1}(t) = \sum_{i=1}^{\ell} a_{\ell+1,i}x_i(t)$ est une fonction connue. On obtient $x_{\ell+1}(t)$ par la méthode de la variation de la constante. ∎

La résolution est particulièrement simple si la matrice A se diagonalise. Nous allons examiner successivement le cas d'une diagonalisation sur $\mathbb{R}$ puis sur $\mathbb{C}$. L'intérêt de ce dernier cas est qu'il existe un ensemble ouvert et dense de matrices qui se diagonalisent sur $\mathbb{C}$. Autrement dit, on peut toujours perturber une matrice aussi peu que voulu pour la rendre diagonalisable sur $\mathbb{C}$. En effet, une matrice $A \in \mathrm{M}(n, n)$ est trivialement diagonalisable sur $\mathbb{C}$ si toutes ses valeurs propres sont racines simples de son polynôme caractéristique car, dans ce cas, il y a n espaces propres de dimension 1, indépendants deux à deux.

Le fait que l'ensemble des matrices A réelles à valeurs propres simples (dans $\mathbb{C}$) est un ensemble ouvert et dense de l'espace des matrices réelles $\mathrm{Mat}(n, n)$ est un fait assez banal et connu, mais nous allons tout de même en donner une preuve dans la petite digression suivante.

Petite digression algébrique

Le nombre complexe $z_0 \in \mathbb{C}$ est une racine non simple d'un polynôme complexe $P_a(z) = \sum_{i=0}^{n} a_i z^i$, où $a = (a_0, \ldots, a_n) \in \mathbb{C}^{n+1}$, si et seulement si il est aussi racine de son polynôme dérivé $P'_a(z)$. L'ensemble $D \subset \mathbb{C}^{n+1}$ des $a = (a_0, \ldots, a_n)$ pour lesquels $P_a(z)$ a une racine non simple, est donc la projection de l'ensemble algébrique $S = \{P_a(z) = P'_a(z) = 0\} \subset \mathbb{C}^{n+2}$, par la projection $\pi(z, a) = a$, de $\mathbb{C}^{n+2}$ dans $\mathbb{C}^{n+1}$ (on dit que l'on élimine z entre les deux polynômes).

Rappelons qu'un *ensemble algébrique* S (complexe) dans $\mathbb{C}^p$ est un ensemble défini par un système de k équations polynomiales

$$\{P_1(u) = P_2(u) = \ldots = P_k(u) = 0\}$$

pour $u \in \mathbb{C}^p$ et $P_1, \ldots, P_k$ des polynômes complexes. On définit de manière analogue les ensembles algébriques réels.

L'ensemble $D = \pi(S)$ est lui-même un ensemble algébrique de $\mathbb{C}^{n+1}$. Cela suit par exemple du théorème de Tarski-Seidenberg [5], [26] :

Si $S \subset \mathbb{C}^{p+q}$ est un ensemble algébrique complexe de $\mathbb{C}^{p+q}$, sa projection $D = \pi(S)$, où π est la projection canonique de $\mathbb{C}^{p+q} \approx \mathbb{C}^p \times \mathbb{C}^q$ sur $\mathbb{C}^q$, est un ensemble algébrique complexe de $\mathbb{C}^q$.

Remarque 2.2. Le théorème de Tarski-Seidenberg est un résultat fondamental des mathématiques. Il va nous servir ici à établir le fait que l'ensemble des matrices diagonalisables sur $\mathbb{C}$ forme un ouvert dense de $\mathrm{Mat}(n, n)$. Ce type de propriété sera généralisé plus loin par la notion de généricité. Le théorème de Tarski-Seidenberg donne une notion de généricité dans le cadre polynomial.

Le théorème de Tarski-Seidenberg est partiellement un énoncé de logique : en effet la projection d'un ensemble algébrique S de $\mathbb{C}^p \times \mathbb{C}^q$, avec $(u, v) \in \mathbb{C}^p \times \mathbb{C}^q$, peut se voir comme *l'élimination du quantificateur « $\exists$ » devant la variable u* dans les équations de S.

Par contre, la projection d'un ensemble algébrique *réel* peut très bien ne pas être un ensemble algébrique réel. Par exemple la courbe algébrique $\{xy = 1\} \subset \mathbb{R}^2$ a pour projection : $\mathbb{R} - \{0\}$ sur l'axe des x. Or, un ensemble algébrique de $\mathbb{R}$ est soit un ensemble fini de points, soit $\mathbb{R}$ tout entier.

L'ensemble algébrique D, dit *(ensemble) discriminant,* est donné par une seule équation algébrique : $\{D(a_0, \ldots, a_n) = 0\}$. Le polynôme D (polynôme sur les variables $a_0, \ldots, a_n$) est appelé *(polynôme) discriminant* et sa valeur $D(a_0, \ldots, a_n)$ s'appelle le *discriminant du polynôme* P_a. Il est assez facile et élémentaire de trouver l'expression de ce polynôme discriminant D qui est égal au

déterminant d'une matrice $(2n-1, 2n-1)$ dont les entrées sont données en fonction de $a_0, \ldots, a_n$ (voir [11] par exemple ; on trouvera facilement son expression par le net, entre autres voir les sites : http ://wikipedia.org ou bien www.bibmath.net/dico/). Pour le polynôme $a_0 + a_1 z + a_2 z^2$, de degré 2, on a évidemment $D(a_0, a_1, a_2) = a_1^2 - 4a_0 a_2$. Ce polynôme D est à coefficients rationnels (donc réels !). Si on identifie le polynôme P à son vecteur de coefficients $(a_0, \ldots, a_n)$, on écrira indifféremment son discriminant : $D(P)$ ou bien $D(a_0, \ldots, a_n)$. La condition $D(P) \neq 0$, pour un polynôme P réel ou complexe, signifie que toutes ses racines *sur* $\mathbb{C}$ sont simples (et pas seulement les racines sur $\mathbb{R}$ si le polynôme est réel).

Revenons maintenant à une matrice réelle $A \in \mathrm{Mat}(n, n)$. La condition de racine simple revient à écrire que le discriminant du polynôme caractéristique $P(A)$ est non nul, ce qui s'exprime par l'inéquation algébrique

$$\{\tilde{D}(A) = D(P(A)) \neq 0\}$$

sur les coefficients de la matrice A. Comme les coefficients du polynôme caractéristique $P(A)$ sont des polynômes des coefficients de la matrice A, le polynôme $\tilde{D}$ est un polynôme à coefficients rationnels sur l'espace des coefficients de la matrice A ($\tilde{D}$ est un polynôme défini sur l'espace : $\mathrm{Mat}(n, n) \approx \mathbb{R}^{n^2}$). Nous sommes maintenant à même de montrer que :

Proposition 2.2. *L'ensemble des matrices ayant toutes leurs valeurs propres simples sur* $\mathbb{C}$ *(ou bien, ce qui est la même chose, ayant* n *valeurs propres sur* $\mathbb{C}$*, deux à deux distinctes) est un ouvert dense de* $\mathrm{Mat}(n, n)$.

D'après ce qui précède, l'ensemble des matrices à valeurs propres simples est le sous-ensemble de $\mathrm{M}(n, n)$ défini par l'inéquation réelle $\{\tilde{D}(A) \neq 0\}$. Remarquons que ce polynôme $\tilde{D}$ n'est pas le polynôme nul. En effet, considérons une matrice diagonale A_0 dont les coefficients diagonaux sont réels et 2 à 2 distincts. Les valeurs propres de A_0 sont précisément ces coefficients diagonaux et donc, par définition, $\tilde{D}(A_0) \neq 0$. Le résultat élémentaire suivant, sur les ensembles algébriques réels, implique alors la proposition 2.3 :

Proposition 2.3. *Soit* $\Sigma = \{u | Q_1(u) = \ldots = Q_k(u) = 0\}$ *un ensemble algébrique réel dans* $\mathbb{R}^p$*, avec :* $u = (u_1, \ldots, u_p)$ *la variable de* $\mathbb{R}^p$*. Supposons que l'un au moins des polynômes* Q_i *n'est pas le polynôme nul. Alors* $U = \mathbb{R}^p - \Sigma$ *est un ouvert dense de* $\mathbb{R}^p$.

Démonstration. Tout d'abord, remarquons qu'il suffit de prouver le résultat pour $k = 1$. En effet, si $k > 1$, supposons par exemple que Q_1 n'est pas le polynôme nul. Comme : $\mathbb{R}^p \setminus \{u | Q_1(u) = 0\} \subset \mathbb{R}^p \setminus \Sigma$, le fait que $\mathbb{R}^p \setminus \{u | Q_1(u) = 0\}$ est un ouvert dense implique qu'il en est de même pour $\mathbb{R}^p \setminus \Sigma$.

Soit donc $Q(u)$ un polynôme non nul. L'ensemble $U = \mathbb{R}^p \setminus \{u | Q(u) = 0\}$ est un ouvert de $\mathbb{R}^p$ car l'application $Q : u \mapsto Q(u)$ est continue, et cet ouvert est non vide car Q n'est pas le polynôme nul. On démontre ensuite que U est dense par récurrence sur *la dimension* p, c'est-à-dire sur le nombre des variables. Si $p = 1$, le polynôme Q étant non nul a un degré $q \in \mathbb{N}$, et il a au plus q racines. L'ensemble U est dense car il est l'ensemble complémentaire de cet ensemble fini des racines. Supposons maintenant que l'affirmation soit démontrée pour toute dimension $\sigma < p$ et considérons un polynôme quelconque non nul des variables $(u_1, \ldots, u_p)$.

Supposons par l'absurde que $U = \mathbb{R}^p \setminus \{u | Q(u) = 0\}$ ne soit pas dense. Cela est équivalent à dire que l'ensemble $\{u | Q(u) = 0\}$ possède un point intérieur, donc qu'il existe p intervalles ouverts *non vides* $V_1, \ldots, V_p$ tels que, pour tout $(u_1, \ldots, u_p) \in V_1 \times \ldots \times V_p$, on ait $Q(u_1, \ldots, u_p) = 0$.

Pour tout $(u_2, \ldots, u_p) \in V_2 \times \ldots \times V_p$ la fonction $u_1 \mapsto Q(u_1, u_2, \ldots, u_p)$ est alors identiquement nulle sur l'intervalle V_1. Comme on peut écrire le polynôme Q comme un polynôme à une variable u_1 avec pour coefficients des polynômes sur les variables $u_2, \ldots, u_p$, soit

$$Q(u_1, \ldots, u_p) = \sum_{i=0}^{n} Q_i(u_2, \ldots, u_p) u_1^i,$$

il suit de la première étape de la récurrence ($p = 1$) que les polynômes $Q_2, \ldots, Q_p$ prennent la valeur 0 sur $V_2 \times \ldots \times V_p$. Par hypothèse de récurrence, ces polynômes sont donc nuls et donc le polynôme Q aussi, ce qui est en contradiction avec l'hypothèse. ∎

Remarque 2.3. Comme tout ensemble algébrique complexe sur $\mathbb{C}^p$ est aussi un ensemble algébrique réel sur $\mathbb{R}^p$, la proposition 2.3 s'applique évidemment aux ensembles algébriques complexes (avec un des polynômes de définition non nul) et, par exemple, à une équation polynomiale complexe $Q(u) = 0$, avec Q polynôme complexe non nul.

Dans la terminologie de René Thom que nous introduirons dans le chapitre III-3, on dit que la propriété de diagonalisation sur $\mathbb{C}$ pour les matrices réelles est une *propriété générique.*

Fin de la digression.

1. *Supposons que A soit diagonalisable sur $\mathbb{R}$.*

   Il suffit pour cela, mais il n'est pas nécessaire, que les valeurs propres soient toutes réelles et deux à deux distinctes. Cette condition est ouverte mais non générique. Soit $y = (y_1, \ldots, y_n)$ le vecteur des coordonnées dans la base de diagonalisation de A et $\lambda_1, \ldots, \lambda_n$ les valeurs propres comptées avec leur multiplicité. Alors l'équation différentielle se réduit à n équations en dimension 1, une pour chaque coordonnée :

   $$\dot{\varphi}_i(t) = \lambda_i \varphi_i(t) \quad \text{pour} \quad i = 1, \ldots, n.$$

   Ces équations s'intègrent en $\varphi_i(t) = y_i e^{\lambda_i t}$ pour une condition initiale $y = (y_1, \ldots, y_n)$. Pour obtenir l'expression des solutions dans les coordonnées initiales, il suffit de se donner l'expression des vecteurs propres : si v_i est le vecteur propre de la valeur propre λ_i, alors $\{v_1, \ldots, v_n\}$ est la base de diagonalisation et la solution avec condition initiale $x = \sum_{i=1}^n y_i$ s'écrit $\varphi(t, x) = \sum_{i=1}^n y_i e^{\lambda_i t} v_i$. Autrement dit, les fonctions vectorielles $e^{\lambda_i t} v_i, i = 1, \ldots, n$ forment une base de l'espace vectoriel $\mathcal{S}$ des solutions sur $\mathbb{R}$.

2. *Supposons que A soit diagonalisable sur $\mathbb{C}$*

   Une condition suffisante mais non nécessaire est que les valeurs propres de A soient deux à deux distinctes. Cette condition est générique dans $\mathrm{Mat}(n, n)$, c'est-à-dire vérifiée pour un ensemble ouvert et dense dans $\mathrm{Mat}(n, n)$.

   Pour exploiter cette condition, on va passer dans le complexe, c'est-à-dire considérer l'équation différentielle dans $\mathbb{C}^n$ et rechercher dans un premier temps toutes les solutions à valeurs dans $\mathbb{C}^n$. Soit

   $$\mu_1, \ldots, \mu_k, \lambda_1, \bar{\lambda}_1, \ldots, \lambda_\ell, \bar{\lambda}_\ell, \quad \text{avec} \quad k + 2\ell = n,$$

   la liste des valeurs propres comptées avec leur multiplicité, où l'on a indiqué d'abord les k valeurs propres réelles, puis les ℓ paires de valeurs propres complexes conjuguées (le fait que les valeurs propres non réelles apparaissent par paires conjuguées vient du fait que le polynôme caractéristique de la matrice est réel, puisque la matrice elle-même est réelle). Soit

   $$L_1, \ldots, L_k, H_1, \bar{H}_1, \ldots, H_\ell, \bar{H}_\ell,$$

   des vecteurs propres correspondant aux valeurs propres ci-dessus, avec $L_i \in \mathbb{R}^n$. Le vecteur $\bar{H}$ est le vecteur dans les composantes sont conjuguées de celles de H. On peut choisir les vecteurs L_i, associés aux valeurs propres réelles μ_i, d'être réels. En effet, de $AL_i = \mu_i L_i$ on tire que $A\bar{L}_i = \mu_i \bar{L}_i$; alors, le vecteur $L_i + \bar{L}_i$ est à la fois réel et vecteur propre de la valeur propre

réelle μ_i. De même, si H_j est un vecteur propre de la valeur propre non réelle λ_j, le vecteur conjugué $\bar{H}_j$ est vecteur propre de la valeur propre $\bar{\lambda}_j$.

On associe à ces vecteurs propres une base des solutions sur $\mathbb{R}$ à valeurs dans $\mathbb{C}^n$

$$e^{\mu_1 t}L_1, \ldots, e^{\mu_k t}L_k, e^{\lambda_1 t}H_1, e^{\bar{\lambda}_1 t}\bar{H}_1, \ldots, e^{\lambda_\ell t}H_\ell, e^{\bar{\lambda}_\ell t}\bar{H}_\ell.$$

Les k premières fonctions sont réelles et les autres sont conjuguées deux à deux. Il est très facile de remplacer ces paires par des paires de fonctions réelles engendrant le même sous-espace $\mathbb{C}$-vectoriel :

$$\begin{cases} \mathrm{Re}(e^{\lambda_j t}H_j) = \frac{e^{\lambda_j t}H_j + e^{\bar{\lambda}_j t}\bar{H}_j}{2} \\ \mathrm{Im}(e^{\lambda_j t}H_j) = \frac{e^{\lambda_j t}H_j - e^{\bar{\lambda}_j t}\bar{H}_j}{2i} \end{cases}.$$

Explicitement, si $H_j = A_j + iB_j$, avec $A_j, B_j \in \mathbb{R}^n$ et $\lambda_j = \alpha_j + i\beta_j$ avec $\alpha_j, \beta_j \in \mathbb{R}$, on a

$$\begin{cases} \mathrm{Re}(e^{\lambda_j t}H_j) = e^{\alpha_j t}(\cos(\beta_j t)A_j - \sin(\beta_j t)B_j) \\ \mathrm{Im}(e^{\lambda_j t}H_j) = e^{\alpha_j t}(\cos(\beta_j t)B_j + \sin(\beta_j t)A_j) \end{cases}.$$

Les fonctions vectorielles

$$e^{\mu_1 t}L_1, \ldots, e^{\mu_k t}L_k, \mathrm{Re}(e^{\lambda_1 t}H_1), \mathrm{Im}(e^{\lambda_1 t}H_1), \ldots, \mathrm{Re}(e^{\lambda_\ell t}H_\ell), \mathrm{Im}(e^{\lambda_\ell t}H_\ell)$$

sont réelles et forment une base des solutions à valeurs dans $\mathbb{C}^n$. Elles forment donc également une base des solutions à valeurs dans $\mathbb{R}^n$ (il suffit de prendre des combinaisons avec coefficients réels).

2.3. Les champs linéaires de vecteurs de $\mathbb{R}^2$

Nous allons examiner, à titre d'illustration, les champs de vecteurs de $\mathbb{R}^2$. Pour simplifier, nous allons nous placer dans la situation générique où la matrice $A \in \mathrm{Mat}(2,2)$ du champ est inversible, et discuter en fonction des valeurs propres des différentes situations possibles. Dans tous les cas, l'origine est l'unique zéro du champ. La trajectoire par l'origine est constante et les autres trajectoires sont des courbes régulières.

1. *Les valeurs propres sont réelles et distinctes.*

   Soit λ_1, λ_2 les valeurs propres avec $\lambda_1 < \lambda_2$. La matrice A est diagonalisable dans la base des vecteurs propres v_1 (associé à λ_1) et v_2 (associé à λ_2). La trajectoire avec condition initiale $(x, y) \in \mathbb{R}^2$ s'écrit $\varphi(t, (x, y)) = xe^{\lambda_1 t}v_1 + ye^{\lambda_2 t}v_2$. D'un point de vue topologique, nous avons trois cas possibles représentés dans la figure 2.1 : le champ s'appelle une *source* si $0 < \lambda_1 < \lambda_2$, un *puits* si $\lambda_1 < \lambda_2 < 0$ et un *point de selle* si $\lambda_1 < 0 < \lambda_2$. Les sources ou puits sont appelés *nœuds*. La terminologie « point de selle » vient de l'exemple du gradient d'une fonction de $\mathbb{R}^2$ au voisinage d'un point de Morse de type selle.

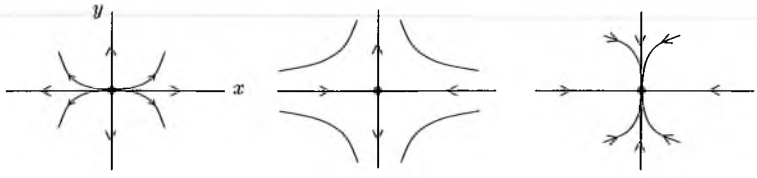


FIGURE 2.1. De gauche à droite, on montre une source, un point de selle et un puits.

2. *Les valeurs propres sont non réelles et conjuguées.*

   Soit $\lambda = \alpha + i\beta$ et $\bar{\lambda} = \alpha - i\beta$ les valeurs propres (on suppose que $\beta \neq 0$). On considère une base $\{H, \bar{H}\}$ de vecteurs propres conjugués. Posons $H = V + iW, \bar{H} = V - iW$ leur décomposition en partie réelle et imaginaire. De $AH = \lambda H = (\alpha + i\beta)(V + iW)$, on tire que $AV = \alpha V - \beta W$ et $AW = \beta V + \alpha W$. En écrivant l'équation différentielle dans la base $\{V, W\}$ avec des coordonnées (X, Y), c'est-à-dire si l'on pose $M = XV + YW$, on a $\dot{M} = \dot{X}V + \dot{Y}W = AM = A(XV + YW) = X(\alpha V - \beta W) + Y(\beta V + \alpha W) = (\alpha X + \beta Y)V + (-\beta X + \alpha Y)W$ et on obtient par identification l'expression

$$\begin{cases} \frac{dX}{dt} = \alpha X + \beta Y \\ \frac{dY}{dt} = -\beta X + \alpha Y \end{cases}.$$

   La matrice $\tilde{A}$ de cette équation différentielle est une *matrice de similitude* : elle est la matrice réelle correspondant à la multiplication complexe $(\alpha - i\beta)(X + iY)$. Il est donc judicieux d'identifier $\mathbb{R}^2$ à $\mathbb{C}$ en posant $z = X + iY$. L'équation différentielle prend alors la forme : $\dot{z} = \bar{\lambda}z$, qui est tout simplement une équation en dimension 1 complexe. Cette équation s'intègre en

$$z(t, z) = ze^{\bar{\lambda}t} = ze^{\alpha t}e^{-i\beta t}. \tag{2.12}$$

La dernière expression peut s'écrire, en vertu de la formule de Moivre, $ze^{\alpha t}(\cos\beta t - i\sin\beta t)$ (en remplaçant z par $X+iY$ et en développant, on retrouve dans ce cas particulier, les expressions données dans le paragraphe précédent). La formule (2.12) permet de tracer très facilement les trajectoires. On peut distinguer trois cas topologiquement distincts. Si $\alpha < 0$, les trajectoires hors l'origine sont des courbes qui spiralent vers l'origine : on dit que le champ est un *foyer de type puits.* Si $\alpha > 0$, les trajectoires hors l'origine sont des courbes qui spiralent en s'éloignant de l'origine : on dit que le champ est un *foyer de type source.* Enfin si $\alpha = 0$, les trajectoires sont des cercles entourant l'origine : on dit que le champ est un *centre.* Voir la figure 2.2 dans le cas où $\beta > 0$.

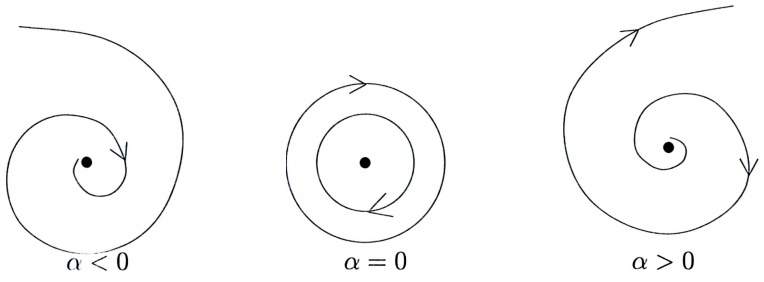


FIGURE 2.2

3. *Cas d'une valeur propre double non nulle.*

Nécessairement, la valeur propre $\lambda \neq 0$ doit être réelle (car la matrice est supposée réelle). La matrice A peut se mettre sous forme triangulaire, et l'équation différentielle prend alors la forme :

$$\begin{cases} \dfrac{dx}{dt} = \lambda x \\ \dfrac{dy}{dt} = \alpha x + \lambda y \end{cases},$$

avec un coefficient α que l'on peut choisir égal à 0 ou 1. Dans le cas $\alpha = 0$, on a une matrice diagonale, qui se traite comme dans le point (1) ci-dessus. Le champ est une source si $\lambda > 0$ et un puits si $\lambda < 0$. Supposons maintenant que $\alpha = 1$. On peut intégrer la première ligne. On obtient $x(t,x) = xe^{\lambda t}$. On reporte ensuite cette fonction dans la seconde ligne. On obtient une équation non autonome en y :

$$\frac{dy}{dt} = \lambda y + xe^{\lambda t},$$

qui s'intègre par la méthode de variation de la constante. On obtient après intégration $y(t,(x,y)) = e^{\lambda t}(xt+y)$. Le champ est un *nœud dégénéré,* de type puits si $\lambda < 0$ et de type source si $\lambda > 0$. Voir la figure 2.3 dans le cas où $\lambda < 0$ et $\alpha < 0$.

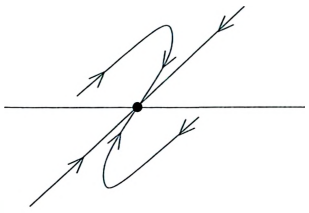


FIGURE 2.3

Remarque 2.4. Nous avons vu qu'un champ de vecteurs linéaire à matrice diagonalisable se réduit à une somme directe de champs de dimension 1 réels ou complexes. On peut donc obtenir le dessin des trajectoires pour toutes les équations génériques (à matrice inversible et diagonalisable sur $\mathbb{C}$), en faisant des produits des cas de dimension 1 et 2 réelle (la dimension 1 est complètement triviale ; remarquer que dans le point (1) ci-dessus, on avait des produits de champs de dimension 1). À titre d'exemple, un champ linéaire de dimension 3 sera toujours le produit d'un champ de dimension réelle 1 par l'un des cas de dimension réelle 2 décrits ci-dessus (figure 2.4).

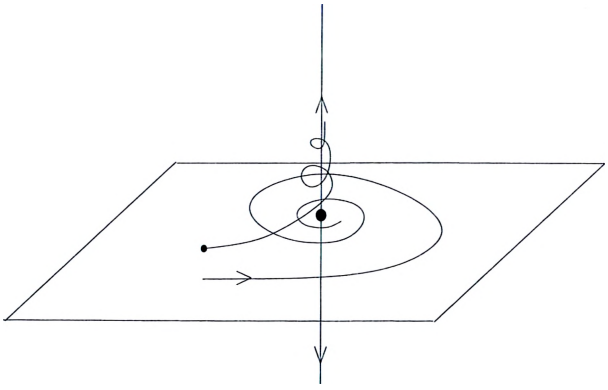


FIGURE 2.4. Cas où une valeur propre λ_1 est réelle positive et où les deux autres valeurs propres sont imaginaires conjuguées.

3

PROPRIÉTÉS GÉNÉRALES DES TRAJECTOIRES

3.1. Le principe du point fixe

On rappelle qu'un espace de Banach est un espace vectoriel normé dans lequel toute suite de Cauchy converge. L'espace $\mathbb{R}^n$, de dimension finie, est un espace de Banach. Il en est de même pour l'espace $\mathcal{C}(F)$ des fonctions continues sur un compact $F \subset \mathbb{R}^n$, avec la topologie de la convergence uniforme définie par la norme $||f||_0 = \text{Sup}\{|f(x)|; x \in F\}$, ou pour les espaces $\mathcal{C}^r(B)$ des fonctions de classe $\mathcal{C}^r$, $r \in \mathbb{N}$, sur une boule fermée $B \subset \mathbb{R}^n$, avec la topologie de la convergence uniforme en classe $\mathcal{C}^r$, définie par la norme $||f||_r = \text{Sup}\{|\partial_j f(x)|; x \in F, 0 \leq |j| \leq r\}$.

Définition 3.1. Soit M un espace métrique muni d'une distance notée $d(x, y)$. On dit que $f : M \mapsto M$ est une application lipschitzienne, de constante de Lipschitz $K \in \mathbb{R}^+$, si

$$d(f(x), f(y)) \leq K d(x, y), \quad \text{pour tout } x, y \in M.$$

On dit que f est une contraction lipschitzienne si $K < 1$.

Remarque 3.1. Une fonction lipschitzienne reste lipschitzienne si on remplace la métrique par une métrique équivalente. Par contre le fait d'être une contraction lipschitzienne pour une fonction dépend du choix de la métrique.

L'intérêt pour un espace E d'être de Banach est de permettre de trouver un point $x \in E$, solution d'un problème d'existence, en construisant une suite de Cauchy dont la limite sera solution du problème posé. Beaucoup de problèmes seront résolus par l'application du résultat suivant :

***Théorème 3.1* (Principe du point fixe).** *Soit $F \subset E$ un fermé dans un espace de Banach E et $f : F \mapsto F$ une contraction lipschitzienne. Alors l'application f a un et un seul point fixe x_0, c'est-à-dire tel que $f(x_0) = x_0$. Pour tout $x \in F$, on a $x_0 = \lim_{n\to+\infty} f^{\circ n}(x)$, où $f^{\circ n}$ désigne l'itérée n fois de l'application f.*

Démonstration. Montrons tout d'abord que, pour tout $x \in F$, la suite $(f^{\circ n}(x))$ est une suite de Cauchy de E. Pour tout $p \geq 1$ dans $\mathbb{N}$, on a :

$$||f^{\circ(p+1)}(x) - f^{\circ p}(x)|| \leq K||f^{\circ p}(x) - f^{\circ(p-1)}(x)||.$$

En itérant cette formule entre 1 et n pour tout entier $n \geq 1$, on obtient (avec la convention $f^{\circ 0} = Id$)

$$||f^{\circ(n+1)}(x) - f^{\circ n}(x)|| \leq K^n||f(x) - x||.$$

Soit maintenant $p \in \mathbb{N}$. Il suit de l'inégalité triangulaire que

$$||f^{\circ(n+p+1)}(x) - f^{\circ n}(x)|| \leq (K^n + \ldots + K^{n+p})||f(x) - x|| \leq \frac{K^n}{1-K}||f(x) - x||.$$

Comme $||f^{\circ(n+p+1)}(x) - f^{\circ n}(x)|| \to 0$ lorsque $n \to +\infty$, la suite $(f^{\circ n}(x))$ est une suite de Cauchy de E.

Soit $x_0 = \lim_{n\to+\infty} f^{\circ n}(x)$. Comme F est fermé , $x_0 \in F$. D'autre part, comme f est continue :

$$f(x_0) = f(\lim_{n\to+\infty} f^{\circ n}(x)) = \lim_{n\to+\infty} f^{\circ(n+1)}(x) = x_0,$$

ce qui prouve que x_0 est un point fixe de f. Finalement, montrons que ce point fixe est unique. Supposons que x_1 soit également un point fixe de f, on a :

$$||x_1 - x_0|| = ||f(x_1) - f(x_0)|| \leq K||x_1 - x_0||.$$

Comme $K < 1$, l'inégalité ci-dessus implique que $||x_1 - x_0|| = 0$, et donc que $x_0 = x_1$. ∎

3.2. Existence et unicité locales des trajectoires

On rappelle que si X est un champ de vecteurs sur un ouvert U de $\mathbb{R}^n$, une trajectoire de X par $x \in U$ est une courbe dérivable $\varphi : t \in I \to \varphi(t) \in U$, définie sur un intervalle ouvert I contenant $0 \in \mathbb{R}$, telle que $\varphi(0) = x$ et $\frac{d\varphi}{dt}(t) = X(\varphi(t))$ pour tout $t \in I$. Voici maintenant le *théorème de Cauchy d'existence et d'unicité des trajectoires dans le cas « Lipschitz »*, dont la preuve est inspirée de [6].

Théorème 3.2. *Soit X un champ de vecteurs sur un ouvert $U \subset \mathbb{R}^n$. On suppose qu'il existe une constante $K > 0$ telle que*

$$||X(x) - X(y)|| \leq K||x - y|| \quad \text{pour tout } x, y \in U,$$

c'est-à-dire que l'application $X : x \mapsto X(x)$ est Lipschitz de constante K. Alors on a :

1. Existence locale des trajectoires

   Soit $A \subset U$, un compact non vide. Alors il existe $\tau > 0$, dépendant de A, une fonction $\varphi : (t, x) \in [-\tau, \tau] \times A \mapsto U$, continue et dérivable par rapport à $t \in [-\tau, \tau]$, telle que, pour tout $x \in A$, la courbe

   $$\varphi(\cdot, x) : t \in]-\tau, \tau[\to \varphi(t, x) \in A$$

   soit une trajectoire par x. Cela signifie que $\varphi(0, x) = x$ et $\frac{\partial \varphi}{\partial t}(t, x) = X(\varphi(t, x))$ pour tout $(t, x) \in [-\tau, \tau] \times A$.

2. Unicité locale des trajectoires

   Soit $\varphi_1 : t \in I_1 \mapsto U$ et $\varphi_2 : t \in I_2 \mapsto U$ deux trajectoires par $x \in U$. Alors il existe un intervalle J de 0, $J \subset I_1 \cap I_2$, tel que $\varphi_1|_J \equiv \varphi_2|_J$.

Démonstration.
(1) Existence locale
L'équation des trajectoires

$$\frac{\partial \varphi}{\partial t}(t, x) = X(\varphi(t, x)), \tag{3.1}$$

avec $\varphi(0, x) = x$ et φ continue et dérivable en t, est équivalente à l'équation

$$\varphi(t, x) = x + \int_0^t X(\varphi(s, x))ds, \tag{3.2}$$

avec φ continue. En effet, remarquez que si une fonction continue vérifie l'équation (3.2), elle est dérivable en t et vérifie (3.1). Inversement, on obtient (3.2) par intégration de (3.1). On va résoudre l'équation intégrale (3.2). Le second membre de cette équation définit un opérateur fonctionnel

$$\varphi(t,x) \mapsto \mathcal{A}(\varphi)(t,x) = x + \int_0^t X(\varphi(s,x))ds, \tag{3.3}$$

et, d'après l'équation (3.2), on a $\mathcal{A}(\varphi) = \varphi$, c'est-à-dire que la trajectoire φ est un point fixe de $\mathcal{A}$. Nous allons trouver φ en appliquant le principe du point fixe 3.1. Nous devons pour cela définir convenablement un espace de Banach E et un fermé F sur lequel $\mathcal{A}$ va opérer comme une contraction lipschitzienne.

Pour tout $\tau > 0$ et compact $A \subset U$, on pose $A_\tau = [-\tau, \tau] \times A$. Soit $E = \mathcal{C}(A_\tau, \mathbb{R}^n)$, l'espace de Banach des applications continues, muni de la norme $||\varphi|| = \mathrm{Sup}\{||\varphi(t,x)||_{\mathbb{R}^n} \mid (t,x) \in A_\tau\}$, où $||\cdot||_{\mathbb{R}^n}$ est une norme quelconque de $\mathbb{R}^n$. Pour tout $x \in A$, *et avec un abus d'écriture classique*, on désigne également par x la projection $(t,x) \mapsto x$. Cette projection x appartient à l'espace E, avec, évidemment, $||x|| = ||x||_{\mathbb{R}^n}$. Soit un $\varepsilon > 0$.

Considérons

$$F = \{\varphi \in E \mid ||\varphi - x|| \leq \varepsilon \text{ et } \varphi(0,x) = x, \forall x \in A\}, \tag{3.4}$$

F est un fermé de E. Nous allons montrer que la paire $(F, \mathcal{A})$ vérifie les conditions du théorème 3.1 pour ε, τ assez petits.

1. *$\mathcal{A}$ est défini sur F, si $\varepsilon > 0$ est assez petit.*

   Comme A est compact, sa distance au bord de U, $\mathrm{dist}(A, \partial U)$, est positive et on peut choisir un ε tel que

   $$0 < \varepsilon < \mathrm{dist}(A, \partial U). \tag{3.5}$$

   On prend un tel ε pour définir l'ensemble F par (3.4). Si

   $$A(\varepsilon) = \{m \in \mathbb{R}^n \mid \mathrm{dist}(m, A) \leq \varepsilon\}$$

   est le ε-voisinage de A, on a $A(\varepsilon) \subset U$ pour un $\varepsilon > 0$ vérifiant (3.5). Maintenant, si $\varphi \in F$ on a $||\varphi(t,x) - x||_{\mathbb{R}^n} \leq \varepsilon$ pour tout $(t,x) \in [-\tau, \tau] \times A$, et donc $\varphi(t,x) \in A(\varepsilon)$; on peut par conséquent calculer $X(\varphi(t,x))$ quel que soit $t \in [-\tau, \tau]$, et définir $\mathcal{A}(\varphi)$.

2. $\mathcal{A}(F) \subset F$, *si* $\tau > 0$ *est assez petit.*

   Comme $A(\varepsilon)$ est un compact dans U et que X est continu, on a $K_1 = \mathrm{Sup}\{||X(y)||_{\mathbb{R}^n} \mid y \in A(\varepsilon)\} < +\infty$. Si $\varphi \in F$, on en déduit, d'après (3.3), que

$$||\mathcal{A}(\varphi)(t,x) - x||_{\mathbb{R}^n} = ||\int_0^t X(\varphi)ds||_{\mathbb{R}^n} \leq \int_0^t ||X(\varphi)||_{\mathbb{R}^n} ds \leq \tau K_1.$$

   Pour que $\mathcal{A}(\varphi) \in F$, il suffit donc de choisir τ tel que

$$\tau K_1 \leq \varepsilon, \tag{3.6}$$

   où le ε est la valeur choisie en (3.5).

3. $\mathcal{A}$ *est une contraction lipschitzienne de* F, *si* $\tau > 0$ *est assez petit.*

   Soit $\varphi, \psi \in F$, on a :

$$[\mathcal{A}(\varphi) - \mathcal{A}(\psi)](t,x) = \int_0^t \big[X(\varphi(s,x)) - X(\psi(s,x))\big] ds,$$

   d'où il suit que

$$||[\mathcal{A}(\varphi) - \mathcal{A}(\psi)](t,x)||_{\mathbb{R}^n} \leq \int_0^t ||X(\varphi(s,x)) - X(\psi(s,x))||_{\mathbb{R}^n} ds.$$

   La propriété de Lipschitz implique que

$$||X(\varphi(s,x)) - X(\psi(s,x))||_{\mathbb{R}^n} \leq K||\varphi(s,x) - \psi(s,x)||_{\mathbb{R}^n},$$

   et donc

$$||[\mathcal{A}(\varphi) - \mathcal{A}(\psi)](t,x)||_{\mathbb{R}^n} \leq K \int_0^t ||\varphi(s,x) - \psi(s,x)||_{\mathbb{R}^n} ds.$$

   En prenant les extremums, d'abord à droite puis à gauche dans cette inégalité, on obtient, dans la norme de l'espace E,

$$|||[\mathcal{A}(\varphi) - \mathcal{A}(\psi)]||| \leq \tau K||\varphi - \psi||.$$

   Il suffit donc de prendre $\tau > 0$ assez petit pour que :

$$L = \tau K < 1, \tag{3.7}$$

   afin que $\mathcal{A}$ soit une contraction lipschitzienne de constante L.

   Si ε, τ vérifient les trois conditions (3.5), (3.6) et (3.7), les conditions du théorème 3.1 sont vérifiées pour la paire $(\mathcal{A}, F)$: il existe une et une seule application $\varphi : (t,x) \mapsto \varphi(t,x)$ vérifiant $\mathcal{A}(\varphi) = \varphi$.

(2) Unicité locale

On peut appliquer le point (1), avec $A = \{x\}$, pour trouver une trajectoire φ par x, définie sur des intervalles $I = [-\tau, \tau]$ avec $\tau > 0$ assez petit. Si $\psi : t \in J \mapsto U$ est une autre trajectoire par x, on peut choisir τ assez petit pour que $I \subset J$. Alors $\psi|I$ est un autre point fixe de l'opérateur $\mathcal{A}$ et, comme ce point fixe est unique, on doit avoir $\varphi \equiv \psi|I$. ∎

Nous allons proposer une autre démonstration de l'unicité locale des trajectoires, en établissant une estimation de l'écart de deux trajectoires quelconques en fonction du temps t.

***Proposition 3.1* (Inégalité de Gronwall).** *Soit U, X et K comme dans l'énoncé du théorème 3.2. Soit φ_1 et φ_2 des trajectoires par les points x_1 et x_2 respectivement. Supposons que ces deux trajectoires soient définies sur l'intervalle de temps $[-\tau, \tau]$. Alors*

$$||\varphi_2(t) - \varphi_1(t)||_{\mathbb{R}^n} \leq ||x_2 - x_1||_{\mathbb{R}^n} e^{K|t|} \quad \text{pour} \quad \forall t \in [-\tau, \tau]. \tag{3.8}$$

En particulier, si $x_1 = x_2$, on obtient $\varphi_2(t) = \varphi_1(t)$ pour $\forall t \in [-\tau, \tau]$, ce qui donne le résultat d'unicité locale des trajectoires.

Démonstration. Si φ est une trajectoire de X, la courbe $t \mapsto \varphi(-t)$ est une trajectoire du champ $-X$. Il suffit donc de prouver (3.8) pour $t \in [0, \tau]$. On a :

$$\varphi_2(t) - \varphi_1(t) = (x_2 - x_1) + \int_0^t [X(\varphi_2(s)) - X(\varphi_1(s))]ds,$$

et, en appliquant la propriété de Lipschitz, on obtient

$$||\varphi_2(t) - \varphi_1(t)||_{\mathbb{R}^n} \leq ||x_2 - x_1||_{\mathbb{R}^n} + K \int_0^t ||\varphi_2(s) - \varphi_1(s)||_{\mathbb{R}^n} ds. \quad ∎$$

L'inégalité (3.8) sur $[0, \tau]$ suit alors du lemme suivant :

Lemme 3.1. *Soit $v : t \in [0, \tau] \mapsto \mathbb{R}$ une fonction positive, continue et telle que*

$$v(t) \leq a + K \int_0^t v(s)ds, \tag{3.9}$$

pour des constantes $a, K \geq 0$. Alors $v(t) \leq ae^{Kt}$ pour tout $t \in [0, \tau]$.

Démonstration. Posons $V(t) = \int_0^t v(s)ds$. L'inégalité (3.9) est équivalente à l'inégalité

$$\frac{dV}{dt}(t) \leq a + KV(t). \tag{3.10}$$

On peut écrire $V(t) = \Phi(t)e^{Kt}$ avec $\Phi(t) = V(t)e^{-Kt}$. De (3.10) on déduit que : $\frac{d\Phi}{dt}(t) \leq ae^{-Kt}$. Par intégration, on obtient

$$\int_0^t \frac{d\Phi}{ds}(s)ds = \Phi(t) - \Phi(0) \leq \frac{a}{K}(1 - e^{-Kt}).$$

Comme $\Phi(0) = V(0) = 0$, on a :

$$V(t) = \Phi(t)e^{Kt} \leq \frac{a}{K}(e^{Kt} - 1),$$

et donc, par définition de $V(t)$ et en appliquant (3.10) :

$$v(t) = \frac{dV}{dt}(t) \leq a + KV(t) \leq a + K \cdot \frac{a}{K}(e^{Kt} - 1) = ae^{Kt}. \quad \blacksquare$$

Remarquons que si le champ X est de classe $\mathcal{C}^k$, avec $k \geq 1$, il suit de la formule $\frac{d\varphi}{dt} = X(\varphi)$ que si φ est $\mathcal{C}^s$ avec $s < k$, alors φ est $\mathcal{C}^{s+1}$, donc φ est $\mathcal{C}^{k+1}$. En fait on peut démontrer la dépendance $\mathcal{C}^k$ par rapport à la variable (t, x) (voir [6] par exemple).

Théorème 3.3 **(Théorème de Cauchy en classe $\mathcal{C}^k$).** *Soit X un champ de vecteurs sur un ouvert $U \subset \mathbb{R}^n$. On suppose que X est de classe $\mathcal{C}^k$ avec $k \geq 1$, fini. Soit W un ouvert tel que $\overline{W}$ soit compact et contenu dans U. Alors il existe $\tau = \tau(W, k) > 0$ et une application $\varphi : (t, x) \in]-\tau, \tau[\times W \mapsto U$, de classe $\mathcal{C}^k$, telle que, pour tout $x \in W$, la courbe $t \mapsto \varphi(t, x)$ soit une trajectoire unique par x.*

Remarque 3.2.

1. Un champ X de classe $\mathcal{C}^1$ est localement lipschitzien. En fait, on peut établir le théorème 3.2 pour des champs de vecteurs localement lipschitziens (c'est-à-dire en restriction à tout ouvert relativement compact).

2. Si X est un champ $\mathcal{C}^\infty$, on peut lui appliquer le théorème 3.3 pour tout $k \in \mathbb{N}$. Comme en général le nombre $\tau(W, k) \to 0$ pour $k \to +\infty$, le théorème 3.3 ne s'applique pas directement à $k = +\infty$. Le résultat est tout de même vrai pour $k = +\infty$ comme on le montrera ci-après (paragraphe 3.3.2).

3. Cauchy a lui-même établi une version du théorème d'existence et d'unicité locales pour les champs de vecteurs analytiques.

4. Le champ $\frac{\partial}{\partial x} + 3|y|^{\frac{2}{3}}\frac{\partial}{\partial y}$ n'est pas localement Lipschitz sur $\mathbb{R}^2$. D'autre part, par chaque point $(x_0, 0)$, il passe deux trajectoires de classe $\mathcal{C}^\infty$: l'axe des x et la cubique, graphe de la fonction $y = (x - x_0)^3$. Ceci montre que le résultat d'unicité n'est pas vrai en général pour les champs X qui sont seulement continus et non localement Lipschitz.

Remarque 3.3. La notion de lipschitzité mérite quelques commentaires. Plaçons-nous pour simplifier dans le cas d'une fonction à une variable. Une fonction continue, linéaire par morceaux, avec un nombre fini de morceaux est lipschitzienne. Par exemple la fonction $f(x) = |x|$ est lipschitzienne. Cette fonction n'est pas dérivable à l'origine. La propriété d'être lipschitzienne ne nécessite donc pas que la fonction f soit dérivable partout.

Il existe même des fonctions qui sont dérivables partout et qui ne sont pas lipschitziennes. Cela est le cas si leur dérivée n'est pas bornée. Par exemple, soit la fonction f définie par $f(x) = x^2 sin(\frac{1}{x^2})$ pour $x \neq 0$ et par $f(x) = 0$ pour $x = 0$. Cette fonction est dérivable partout, mais sa dérivée n'est pas bornée même localement à l'origine et elle n'est donc pas de classe $\mathcal{C}^1$. Par contre, il suit du théorème des accroissements finis qu'une fonction de classe $\mathcal{C}^1$ en dimension quelconque est localement lipschitzienne, comme il a déjà été mentionné à propos des champs de vecteurs dans le point 1 de la remarque 3.2.

3.3. Flot d'un champ de vecteurs

3.3.1. Trajectoire maximale

Soit X un champ de vecteurs sur un ouvert U de $\mathbb{R}^n$. On suppose que X est au moins Lipschitz de façon que l'on puisse appliquer le théorème 3.2. Nous allons maintenant globaliser la notion de trajectoire.

Définition 3.2. On dit que $\varphi_x(t)$ est une trajectoire maximale de X par le point $x \in U$ si et seulement si $\varphi_x : t \in I_x =]-\tau_1(x), \tau_2(x)[\mapsto U$ est une trajectoire, telle que si $\varphi : t \in I \mapsto U$ est une autre trajectoire par x, alors $I \subset I_x$ et $\varphi \equiv \varphi_x|_I$ (figure 3.1). (La notation $\varphi_x(t)$ est provisoire.)

Proposition 3.2. *Pour tout $x \in U$, il existe un intervalle ouvert I_x, voisinage de 0, et une trajectoire $\varphi_x : I_x \mapsto U$ avec $\varphi(0) = x$ qui est maximale. De plus (φ_x, I_x) est unique.*

Démonstration. L'unicité de (φ_x, I_x) vient de la définition de la maximalité. En effet, si (φ, I) est une deuxième solution maximale, alors $I_x \supset I$ et $\varphi \equiv \varphi_x|_I$ et inversement $I \supset I_x$ et $\varphi_x \equiv \varphi|_{I_x}$ d'où $I_x = I$ et $\varphi \equiv \varphi_x$.

L'existence repose sur le lemme 3.2.

Lemme 3.2. *Si φ_1 et φ_2 sont deux trajectoires par x, définies respectivement sur I_1 et I_2, alors $\varphi_1 \equiv \varphi_2$ sur $I_1 \cap I_2$.*

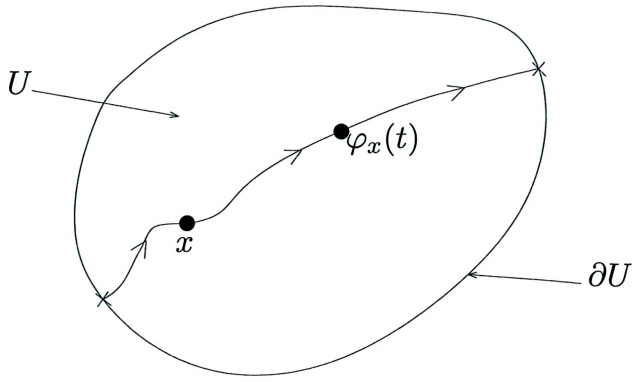


Figure 3.1. Exemple de trajectoire maximale.

Démonstration. Considérons l'ensemble $O = \{t \in I_1 \cap I_2 \mid \varphi_1(t) = \varphi_2(t)\}$. Alors O est un *fermé* parce que défini par l'ensemble des t vérifiant l'équation $\varphi_1(t) - \varphi_2(t) = 0$ avec φ_1 et φ_2 continues.

Mais O est aussi un *ouvert*. En effet, si $\{t_0\} \in O$, il suit du point (2) de la remarque 1.5 que $\tau \mapsto \varphi_1(t_0 + \tau)$ et $\tau \mapsto \varphi_2(t_0 + \tau)$ sont deux trajectoires par x_0, définies pour τ voisin de 0. Ces deux trajectoires doivent coïncider sur un voisinage J de 0, d'après le point (2) du théorème 3.2 (unicité locale). En conséquence $\{t_0\} + J \subset O$. Comme cela est vrai pour tout $t_0 \in O$, il en résulte que O est un *ouvert* comme voisinage de chacun de ses points.

L'ensemble O est donc à la fois *ouvert et fermé* dans $I_1 \cap I_2$ qui est connexe (car c'est un intervalle). D'autre part $O \neq \emptyset$ car $\{0\} \in O$. Il en résulte que $O = I_1 \cap I_2$ par la définition même de la connexité (sinon on aurait partagé $I_1 \cap I_2$ en deux ouverts disjoints et non vides, ce qui n'est pas possible). Ceci achève la preuve du lemme. ∎

Maintenant on peut achever la preuve de la proposition 3.2. On va construire la trajectoire maximale (φ_x, I_x) par une technique « marteau-pilon » consistant à considérer la réunion de toutes les trajectoires par x.

Fin de la preuve de la proposition 3.2. Soit $I_x = \bigcup\{J \mid J \text{ intervalle de définition d'une trajectoire par } x\}$. Posons $\varphi_x(t) = \varphi_J(t)$ si $t \in J$, J domaine de définition de $\varphi_J : J \mapsto U$. La fonction φ_x est *bien définie* (car si $t \in J \cap J'$, $\varphi_J(t) = \varphi_{J'}(t)$ d'après le lemme 3.2). Maintenant φ_x est une trajectoire (puisqu'en chaque t elle coïncide avec une trajectoire) et elle est *maximale* : en effet si (φ_J, J) est une trajectoire quelconque par x, on a par construction que $J \subset I_x$ et $\varphi_J = \varphi_x|_J$, et donc que φ_x est trajectoire maximale par x d'après la définition 3.2. Ce qui termine la preuve de la proposition 3.2. ∎

Dorénavant, par *trajectoire par* x, nous entendrons la trajectoire maximale par x et nous la noterons $\varphi(t,x)$ et non plus $\varphi_x(t)$. Nous pouvons maintenant traduire la propriété d'unicité sur les trajectoires maximales elles-mêmes :

Proposition 3.3. *Soit $x \in U$ et $y = \varphi(\tau,x)$ pour un certain $\tau \in]-\tau_1(x),\tau_2(x)[$. Alors $]-\tau_1(y),\tau_2(y)[=]-\tau_1(x),\tau_2(x)[-\{\tau\}$ et pour tout $t \in]-\tau_1(y),\tau_2(y)[$ on a la* formule d'addition *suivante :*

$$\varphi(t+\tau,x) = \varphi(t,\varphi(\tau,x)) = \varphi(t,y). \tag{3.11}$$

Démonstration. Pour t voisin de 0, les deux applications $t \mapsto \varphi(t,\varphi(\tau,x))$ et $t \mapsto \varphi(t+\tau,x)$ sont deux trajectoires par y, donc coïncident. Elles sont égales sur leur domaine de définition. Par maximalité, il vient $]-\tau_1(y),\tau_2(y)[\supset]-\tau_1(x),\tau_2(x)[-\{\tau\}$. On a l'inclusion opposée en inversant x et y. ∎

Définition 3.3. On appelle *orbite* de X par x l'*arc géométrique orienté* (par l'orientation du temps croissant), défini comme classe d'équivalence de la trajectoire maximale par x, modulo les changements de paramétrisation h à dérivée positive (on dit que h préserve l'orientation donnée par le temps).

Notation. On notera par γ_x^X l'orbite du champ X par le point x, ou simplement par γ_x s'il n'y a pas d'ambiguïté. On écrit souvent un peu abusivement $\gamma_x^X = \bigcup_t \varphi(t,x)$ (en fait, l'orbite n'est pas seulement un ensemble de points, mais plus précisément un arc géométrique orienté).

Il suit du théorème de l'unicité de chaque trajectoire maximale que les orbites sont les éléments d'une partition de U. En effet, si $y \in \gamma_x$, on a $\gamma_x = \gamma_y$, donc l'orbite γ_x est la classe d'équivalence de x par la relation d'équivalence : *appartenir à une même orbite.* L'espace quotient de U pour cette relation d'équivalence est appelé *espace des orbites* du champ de vecteurs et on désigne par *portrait de phase* la description de cet espace.

Soit maintenant $W = \bigcup_{x\in U} I_x \times \{x\} \subset \mathbb{R}\times U$ où I_x est le domaine de définition de la trajectoire maximale par x. L'ensemble W est la réunion des intervalles de définition de la solution maximale dans le produit $\mathbb{R} \times U$.

Définition 3.4. Le *flot* de X est l'application

$$\varphi : W \mapsto U,$$

définie par $I_x \ni t \mapsto \varphi(t,x)$, où $\varphi(t,x)$ est la solution maximale de X par $x \in U$.

Autrement dit, on considère la fonction définie en considérant l'ensemble des fonctions $t \mapsto \varphi(t,x)$ pour les différents points $x \in U$. On introduit cette notion car elle possède les bonnes propriétés qui sont présentées dans le prochain paragraphe.

Notation. Le flot sera noté dans la suite par $\varphi(t,x)$ ou simplement φ. La mention du flot lèvera toute ambiguïté, il suffira de revenir à la définition 3.4.

3.3.2. Propriétés différentiables du flot

Théorème 3.4. *Soit $W = \cup\{]-\tau_1(x), \tau_2(x)[\times\{x\}|\ x \in U\} \subset \mathbb{R}\times U$, le domaine du flot de X. Alors*

1. *W est un ouvert de $\mathbb{R}\times U$.*
2. *L'application flot $\varphi : W \mapsto U$ est continue si X est Lipschitz. Si X est de classe C^k, avec $1 \leq k \leq +\infty$, l'application φ est également de classe C^k.*

Remarque 3.4. Le théorème précédent est vrai pour $k = \infty$, contrairement au théorème local 3.3 qui n'était valable que pour les valeurs finies de k.

Preuve du théorème 3.4. Nous allons tout d'abord montrer simultanément les points (1) et (2) pour X Lipschitz ou de classe C^k, k fini. Nous on déduirons ensuite le cas $k = +\infty$ par un argument d'unicité.

(i) X Lipschitz ou bien de classe C^k, k fini.

Soit (t_0, x_0) un point quelconque dans W. Considérons $\sigma = \varphi([0,t_0], x_0) \subset U$. L'ensemble σ est compact (image du compact $[0,t_0]\times\{x_0\}$ par une application continue) ; on peut donc choisir un voisinage ouvert A de σ dans U tel que $\overline{A}$ soit compact et $\overline{A} \subset U$ (figure 3.2).

L'idée de la démonstration est la suivante : *on peut décomposer un parcours de longueur t arbitraire en N petits pas de longueur t/N pour lesquels on peut appliquer le théorème local 3.2, ou bien 3.3, selon le cas.*

Pour ce voisinage A de σ, il existe $\tau > 0$ donné par le théorème local. On choisit N tel que $N\tau > t_0$. En utilisant le flot local $\psi(t,x)$ sur $[-\tau,\tau]\times\bar{A}$, on

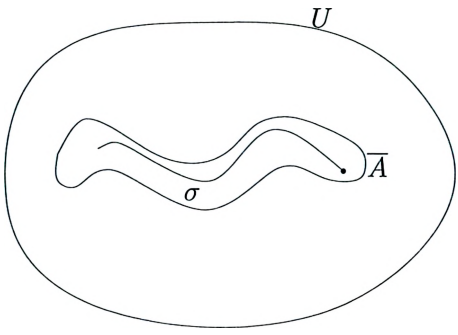


FIGURE 3.2

peut définir l'application $\psi : (t/N, x) \mapsto \psi(t/N, x)$ pour $(t, x) \in [0, N\tau] \times \overline{A}$. Rappelons que le τ du théorème local dépend du voisinage A et de k, mais qu'il est indépendant du point x choisi dans ce voisinage. Par construction, l'application $(t, x) \mapsto \psi(t/N, x)$ est continue, ou de classe C^k, selon le cas.

Maintenant les points $\psi(t_0/N, x_0)$, $\psi(t_0/N, \psi(t_0/N, x_0))\ldots$ et ainsi de suite jusqu'à N compositions, appartiennent, par récurrence, à σ. En effet le point $\psi(t_0/N, x_0) \in \sigma$ car localement les trajectoires des flots φ et ψ passant par x_0 sont confondues grâce au résultat d'unicité locale ; de même $\psi(t_0/N, \psi(t_0/N, x_0))$ appartient à σ car localement les trajectoires des flots φ et ψ passant par $\psi(t_0/N, x_0)$ sont confondues, et ainsi de suite.

Par continuité, les applications

$$\psi : (t/N, x) \mapsto \psi(t/N, x), \quad \psi : (t/N, \psi(t/N, x)) \mapsto \psi(t/N, \psi(t/N, x))\ldots$$

et ainsi de suite pour N compositions, sont définies pour $(t, x) \in I \times B$ où I est un voisinage de t_0 et $B \subset A$ est un voisinage de x_0. D'autre part, en raison de l'unicité des trajectoires locales et de la proposition 3.3, nous avons que $\psi(t/N, x) = \varphi(t/N, x)$, $\psi(t/N, \psi(t/N, x)) = \varphi(2t/N, x)\ldots$ et ainsi de suite pour N compositions, pour $(t, x) \in I \times B$. Finalement, en considérant exactement N compositions, nous obtenons que :

$$\varphi(t, x) \equiv \psi\Big(t/N, \psi\big(t/N, \psi(t/N, \ldots, x)\big)\ldots\Big) \quad (N \text{ compositions}), \tag{3.12}$$

pour tout $(t, x) \in I \times B$. Il résulte de ceci que le flot φ est défini sur le voisinage $I \times B$ du point (t_0, x_0) et donc que $I \times B \subset W$. Comme (t_0, x_0) est un point quelconque de W, ce domaine W est un ouvert de $\mathbb{R} \times U$. Ce qui démontre le point 1. Le terme de droite de la formule (3.12) est une composition d'applications continues si X est Lipschitz, d'après le théorème 3.2, et d'applications de classe C^k

si X est de classe $\mathcal{C}^k$, $k < \infty$, d'après le théorème 3.3. Il s'ensuit que le flot φ est continu ou de classe $\mathcal{C}^k$ selon le cas. Ce qui démontre le point 2.

(ii) X de classe $\mathcal{C}^\infty$.

Supposons que le champ X soit maintenant de classe $\mathcal{C}^\infty$. Le champ X est de classe $\mathcal{C}^k$, pour tout $k \geq 1$. En raison de l'unicité du flot en toute classe, le flot trouvé au point (i) en classe $\mathcal{C}^k$ est identique à celui trouvé au point (i) en classe $\mathcal{C}^{k+1}$. On a donc un et un seul flot, indépendant de la classe considérée. Ce flot unique étant de classe $\mathcal{C}^k$ quel que soit k, est de classe $\mathcal{C}^\infty$.

Remarque 3.5. Si X est un champ analytique, on montre que son flot est analytique.

3.3.3. Groupe à 1-paramètre

Définition 3.5. Soit X, de classe C^r sur U. On dit que le flot de X est *complet* (on dira aussi que le champ est complet) si $W = \mathbb{R} \times U$.

Autrement dit, on a alors $I_x = \mathbb{R}$ pour tout $x \in U$ et on pourra considérer $\varphi(t, x)$ pour tout $t \in \mathbb{R}$.

Les champs de vecteurs linéaires sont des exemples typiques de champs à flot complet. En effet, le flot du champ d'équation différentielle $\dot{x} = Ax$ est donné par la formule exponentielle $\varphi(t, x) = e^{tA}x$, formule qui est définie pour tout t et tout x. Une forme équivalente de la définition d'un flot complet est l'existence d'un temps d'intégration fini mais uniforme :

Lemme 3.3. *Supposons qu'il existe un $\delta > 0$ tel que pour tout $x \in U$ on ait $[-\delta, \delta] \subset I_x$, autrement dit que la trajectoire maximale en chaque point soit définie sur au moins un intervalle $[-\delta, \delta]$. Alors le champ X est à flot complet. La réciproque est évidemment triviale.*

Démonstration. Supposons qu'il existe un $\delta > 0$ comme dans l'énoncé. On peut écrire la formule d'addition (3.11) pour $t, \tau \in [-\delta, \delta]$ avec x quelconque. Chaque trajectoire maximale est donc définie sur l'intervalle $[-2\delta, 2\delta]$. De la même manière on montre, par récurrence, que chaque trajectoire maximale est définie sur $[-k\delta, k\delta]$, pour tout $k \in \mathbb{N}$, donc sur tout $\mathbb{R}$. ∎

Pour les champs de vecteurs à flot complet, la proposition 3.3 prend la forme très simple suivante, que l'on appellera *formule de translation* :

Proposition 3.4. *Considérons X à flot φ complet. Alors, pour tout $t, \tau \in \mathbb{R}$ et tout $x \in U$, on a :*

$$\varphi(t, \varphi(\tau, x)) = \varphi(t + \tau, x). \tag{3.13}$$

On rappelle qu'un difféomorphisme $\mathcal{C}^r$, sur un ouvert U, est une application inversible de U sur U, de classe $\mathcal{C}^r$ dans les deux sens (f et f^{-1} de classe $\mathcal{C}^r$). L'ensemble des difféomorphismes de classe $\mathcal{C}^r$ de U est donc un groupe pour la composition, groupe que l'on notera $\mathrm{Diff}^r(U)$. Nous avons alors le Théorèm 3.5

Théorème 3.5. *Soit X à flot complet, de classe C^r, alors :*

1. *Pour tout $t \in \mathbb{R}$, $\varphi_t : x \mapsto \varphi(t, x)$ est un difféomorphisme C^r de U sur U.*
2. *Quels que soient $t_1, t_2 \in \mathbb{R}$, $\varphi_{(t_1+t_2)} = \varphi_{t_1} \circ \varphi_{t_2}$.*

Remarque 3.6. Le point 2 du théorème 3.5 signifie que l'application $t \mapsto \varphi_t$ est un homomorphisme du groupe $\mathbb{R}$ muni de l'addition dans le groupe des difféomorphismes $\mathcal{C}^r$ sur U (d'après le point 1) muni de la loi de composition.

Preuve du théorème 3.5. On a tout d'abord la formule du point 2 grâce à la propriété de translation. En effet, pour tout $x \in U$:

$$\varphi_{(t_1+t_2)}(x) = \varphi(t_1 + t_2, x) = \varphi(t_1, \varphi(t_2, x)) = \varphi_{t_1}(\varphi_{t_2}(x)) = \varphi_{t_1} \circ \varphi_{t_2}(x).$$

Démontrons maintenant le point 1. Pour tout $x \in U$, nous avons

$$\varphi_0(x) = x = \varphi_{t-t}(x) = \varphi_t \circ \varphi_{-t}(x) = \varphi_{-t} \circ \varphi_t(x),$$

donc φ_t admet φ_{-t} comme inverse.

Or, pour tout t, l'application $x \mapsto \varphi_t(x)$ est $\mathcal{C}^r$, d'après la propriété 2 du flot, que t soit positif ou négatif. Donc φ_t et son inverse φ_{-t} sont $\mathcal{C}^r$. Ce qui termine la preuve. ∎

Définition 3.6. Soit φ un flot complet. L'application $t \mapsto \varphi_t$ sera aussi appelée *groupe à un paramètre* (de *difféomorphismes de U*).

On a donc un changement de point de vue : au lieu de considérer les trajectoires (maximales) une par une, on regarde pour chaque t, le *mouvement global* sur U défini par $x \mapsto \varphi_t(x)$. Pour $t = 0$, on a l'identité et, pour t croissant, la *déformation* peut être de plus en plus accentuée ; mais pour chaque t, φ_t reste un difféomorphisme (imaginer la déformation régulière d'une membrane élastique) (figure 3.3).

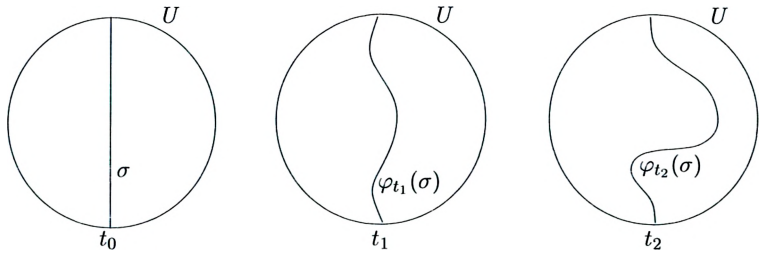


FIGURE 3.3. Déformation d'un segment σ par un flot φ_t pour des temps croissants $t_0 < t_1 < t_2$.

Prenons l'exemple du champ de vecteurs $X = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$ du plan $\mathbb{R}^2$ étudié dans le paragraphe 2.3. Son groupe à 1-paramètre est le *groupe des rotations du plan.*

Remarque 3.7. Si φ n'est pas complet, nous n'avons pas un groupe de difféomorphismes, mais un *pseudo-groupe de difféomorphismes :* pour chaque ouvert W, tel que $\overline{W}$ soit compact et contenu dans U, il existe des φ_t pour des $|t| < t_W$ avec φ_t difféomorphisme de W sur son image. Les φ_t se composent sur des domaines restreints : si W, W' sont deux domaines avec des temps correspondants $t_1 = t_W$ et $t_2 = t_{W'}$, $\varphi_{t_2} \circ \varphi_{t_1}$ sera défini et coïncidera avec $\varphi_{t_1+t_2}$ sur $\varphi_{t_1}^{-1}(\varphi_{t_1}(W) \cap W')$ (voir la figure 3.4) ; on rappelle que $\varphi_t^{-1} \equiv \varphi_{-t}$. Remarquez que l'on ne peut pas prendre $W = U$, car alors le seul temps d'intégration possible pour l'ouvert U est égal à $t_U = 0$ (en fait on a montré dans le lemme 3.3 que φ est complet si et seulement $t_U \neq 0$). La collection des $\{W, \varphi\}$, où $\overline{W} \subset U$ avec $\overline{W}$ compact, U ouvert et $\varphi : W \mapsto U$ est un difféomorphisme sur son image, est appelée le pseudo-groupe des difféomorphismes locaux sur U. Les $\{W, \varphi_t\}$ pour $|t| < t_W$ appartiennent à ce pseudo-groupe.

Les champs non complets sont nombreux. Donnons-en trois exemples.

1. Soit X un champ quelconque, non identiquement nul, sur un ouvert U. Considérons un point $x_0 \in U$ tel que $X(x_0) \neq 0$. Alors le champ Y obtenu en restreignant X à $V = U - \{x_0\}$ n'est pas à flot complet. On le voit de la façon suivante. Soit φ le flot de X sur U et un temps $t_0 > 0$ tel que $y_0 = \varphi(-t_0, x_0) \neq x_0$, et donc que $y_0 \in V$ (un tel temps t_0 existe car $X(x_0) \neq 0$; sinon on aurait $\varphi(t, x_0) \equiv x_0$ et l'orbite serait constituée du seul point x_0, ce qui n'est possible que si $X(x_0) = 0$). On a évidemment : $\varphi(t_0, y_0) = x_0 \notin V$. Remarquons que le flot de Y sur V s'obtient par restriction du flot de X à une partie du domaine du flot de X (car $V \subset U$). En conséquence, le fait que $\varphi(t_0, y_0) \notin V$ implique que l'intervalle de définition I_{y_0} de la trajectoire

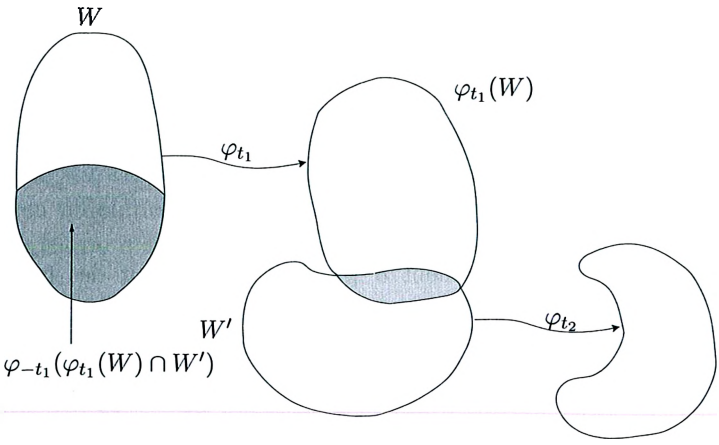


FIGURE 3.4

dans V de Y par le point y_0 ne peut pas contenir la valeur t_0 (en effet tous les points de la trajectoire de Y par y_0 doivent appartenir à V par définition). L'intervalle I_{y_0} est donc de la forme $]-\tau_2(y_0), \tau_1(y_0)[$ avec $\tau_1(y_0) \leq t_0$ (en fait il est facile de voir que $\tau_1(y_0) = t_0$). En conséquence, $I_{y_0} \neq \mathbb{R}$ et le flot de Y n'est pas complet.

2. *Équation de Newton* Cette équation différentielle est définie dans $U \subset \mathbb{R}^{6n}$, espace de phase des positions-vitesses $(x, \dot{x})$. Précisément $U = \mathbb{R}^{6n} - \Delta$ où Δ est l'ensemble des chocs : $\{x_i = x_j | i \neq j\}$.

   Le potentiel tend vers l'infini lorsqu'on se rapproche de Δ et l'on peut avoir des chocs en temps fini à partir de certaines conditions initiales $(x, \dot{x})$ $(I_{(x,\dot{x})} \neq]-\infty, \infty[)$, par exemple si on lâche deux particules $(n = 2)$ à une distance non nulle mais avec des vitesses initiales nulles. Plus généralement, il est possible pour un nombre quelconque n de particules, de trouver des conditions initiales telles que le mouvement aboutisse à un choc d'au moins deux des particules en un temps fini, c'est-à-dire que l'on peut arriver à la frontière du domaine en un *temps fini* avec une loi en $1/x^2$ (un choc intervient lorsque la trajectoire atteint la frontière ∂U et l'instant du choc est la borne supérieure de I_x).

   Le flot des équations de Newton n'est donc pas complet.

3. Soit le champ $X = x^2 \frac{\partial}{\partial x}$ sur $\mathbb{R}$. L'équation différentielle de ce champ $\frac{dx}{dt} = x^2$ a pour flot explicite

$$\varphi(t, x_0) = \frac{x_0}{1 - x_0 t}. \tag{3.14}$$

Alors si $x_0 > 0$, on a $\varphi(t, x_0) \to \infty$ lorsque $t \to \frac{1}{x_0}$. Le flot n'est donc pas défini en $\frac{1}{x_0}$ et $I_{x_0} =]-\infty, \frac{1}{x_0}[$. De manière analogue, $I_{x_0} =]\frac{1}{x_0}, \infty[$ si $x_0 < 0$. Par contre $I_0 =]-\infty, +\infty[= \mathbb{R}$ car la formule (3.14) donne $\varphi(t, 0) \equiv 0$. Le flot n'est donc pas complet et son domaine de définition, $W = \cup_{x \in \mathbb{R}} I_x \times \{x\} \subset \mathbb{R}^2$, est la région connexe ouverte située entre les deux branches de l'hyperbole $\{xt = 1\}$ (région hachurée dans la figure 3.5). La partie droite de la figure montre le portrait de phase.

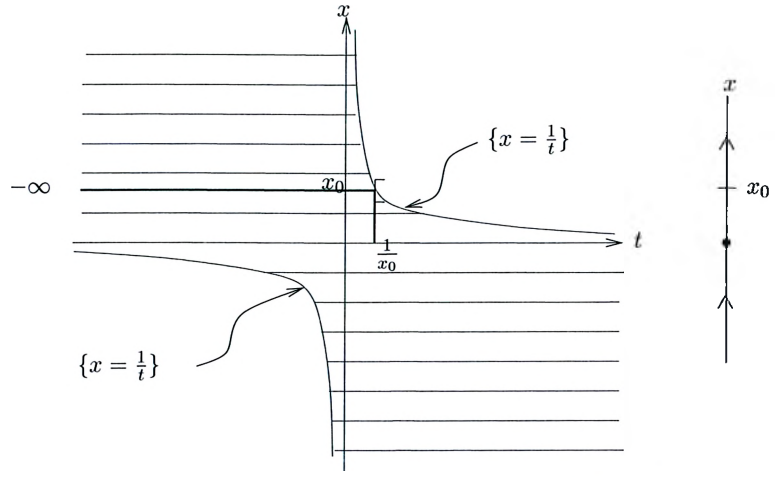


FIGURE 3.5. Le domaine W du flot de $x^2 \frac{\partial}{\partial x}$.

3.3.4. Équivalence à des champs de vecteurs à flot complet

Pour simplifier, nous allons supposer que les champs de vecteurs considérés sont de classe C^∞. Nous allons montrer que si X n'est pas complet sur $U = \mathbb{R}^n$, il est facile de trouver une fonction f telle que $f(x) > 0$ pour tout $x \in \mathbb{R}^n$, de classe C^∞, telle que $Y = fX$ soit complet.

Tout d'abord, remarquons que si deux champs sont proportionnels par un facteur positif, ils ont les mêmes orbites :

Proposition 3.5. *Soit X un champ de vecteurs sur un ouvert U et f une fonction de classe C^∞ telle que $f(x) > 0$ pour tout $x \in U$ (on écrira $f > 0$). Alors le champ de vecteurs $Y = fX$ a les mêmes orbites que le champ X.*

Remarque 3.8. Soit φ^X et φ^Y les flots de X et Y respectivement. Considérons un point $x \in U$ quelconque, et soit I_x et J_x les intervalles de définition des

trajectoires (maximales) de X et Y respectivement. Soit γ_x^X et γ_x^Y les orbites par x, respectivement des champs X et Y. Il suit de la définition 3.3, que $\gamma_x^X = \gamma_x^Y$ (avec X et Y définis comme dans la proposition 3.5) si et seulement s'il existe un difféomorphisme $h : t \mapsto h(t)$ de I_x sur J_x, avec $h'(t) > 0$ pour tout t, tel que $\varphi^X(t,x) = \varphi^Y(h(t),x)$.

Preuve de la proposition 3.5. Il suffit de montrer que si $x \in U$, γ_x^X orbite de X par x et γ_x^Y orbite par x de Y, alors $\gamma_x^X \subset \gamma_x^Y$ (en effet comme $X = \frac{Y}{f}$, le même raisonnement montrerait que $\gamma_x^Y \subset \gamma_x^X$, d'où on conclurait que $\gamma_x^X = \gamma_x^Y$).

Pour un même point x, soit $\varphi^X : t \in I \mapsto U$ (resp. $\varphi^Y : s \in J \mapsto U$) la trajectoire maximale de X (resp. Y). On cherche un difféomorphisme $s = h(t)$ de I dans J, à dérivée positive, tel que

$$\varphi^X(t) = \varphi^Y \circ h(t), \quad \forall t \in I.$$

Supposons le problème résolu (c'est-à-dire que h existe !). Par dérivation on a :

$$\frac{d\varphi^X}{dt}(t) = \frac{dh}{dt}(t)\ \frac{d\varphi^Y}{ds}\big(h(t)\big). \tag{3.15}$$

(L'ordre des facteurs dans le terme de droite de (3.15) est justifié par le fait que $\frac{dh}{dt}(t)$ est un scalaire et que $\frac{d\varphi^Y}{ds}\big(h(t)\big)$ est un vecteur.)

Puisque $\varphi^Y(s)$ est la trajectoire de Y, l'équation (3.15) implique que

$$X\big(\varphi^X(t)\big) = \frac{d\varphi^X}{dt}(t) = \frac{dh}{dt}(t)\ Y\big(\varphi^Y \circ h(t)\big) = \frac{dh}{dt}(t)\ Y\big(\varphi^X(t)\big), \tag{3.16}$$

pour tout $t \in I$. Puisque $Y = fX$, on a aussi que

$$Y\big(\varphi^X(t)\big) = f\big(\varphi^X(t)\big)\ X\big(\varphi^X(t)\big), \tag{3.17}$$

pour tout $t \in I$. En substituant (3.17) dans (3.16), on obtient que

$$X\big(\varphi^X(t)\big) = \frac{dh}{dt}(t)\ f\big(\varphi^X(t)\big)\ X\big(\varphi^X(t)\big). \tag{3.18}$$

On suppose maintenant que $X(x) \neq 0$ (sinon $\varphi^X(t,x) \equiv \varphi^Y(t,x) \equiv x$ et dans ce cas on peut prendre $h(t) \equiv t$). Pour tout $t \in I$, on a alors que $X(\varphi^X(t)) \neq 0$.

En effet si $X(\varphi^X(t_0)) = 0$ pour un certain t_0, l'orbite de $\varphi^X(t_0)$ se réduirait à ce point. Mais comme cette orbite est aussi celle de x, on aurait que $\varphi^X(t) \equiv x$ et donc $X(x) = 0$, contrairement à l'hypothèse (nous reviendrons sur cette question de la classification des orbites dans le chapitre 4).

Comme $X(\varphi^X(t)) \neq 0$ pour tout $t \in I$ et que $f > 0$, on déduit de (3.18) que $\frac{dh}{dt} = \frac{1}{f \circ \varphi^X}$. La fonction h est donc égale à

$$h(t) = \int_0^t \frac{du}{f \circ \varphi^X(u)}. \tag{3.19}$$

Cette formule définit un difféomorphisme h de classe $\mathcal{C}^\infty$ préservant l'orientation, car f et φ^X sont de classe $\mathcal{C}^\infty$ et que $\frac{dh}{dt} = \frac{1}{f \circ \varphi^X} > 0$.

Soit h la fonction définie par (3.19) et $\varphi^Y(s)$ la trajectoire par x de Y. Alors la courbe $\varphi^X(t) = \varphi^Y\big(h(t)\big)$ vérifie (3.16) et donc également (3.15). On en conclut que $\varphi^X(t)$ est trajectoire par x de X et donc que $\gamma_x^X \subset \gamma_x^Y$. ∎

Définition 3.7. Deux champs de vecteurs X, Y sur U, comme dans la proposition précédente, sont dits $\mathcal{C}^\infty$-*équivalents* (la relation : *il existe une fonction* $f > 0$ *de classe* $\mathcal{C}^\infty$ *telle que* $Y = fX$ est manifestement une relation d'équivalence dans l'espace des champs de vecteurs $\mathcal{C}^\infty$ sur U).

Il existe un critère suffisant basé sur la compacité, pour que la trajectoire (maximale) de X par x soit complète (c'est-à-dire $I_x = \mathbb{R}$). L'exemple des champs linéaires montre que ce critère n'est pas nécessaire.

***Lemme 3.4* (critère de complétude).** *Soit* $x \in U$ *et* $I_x =]-\tau_1(x), \tau_2(x)[$ *le domaine de définition de la trajectoire maximale par* x. *Soit* $\gamma_x^+ = \bigcup_{t \geq 0} \varphi(t, x) = \varphi([0, \tau_2(x)[, x)$, *la* demi-orbite positive *par* x. *Si* γ_x^+ *est bornée, alors* $\tau_2(x) = +\infty$. *Si l'orbite* γ_x *elle-même est bornée, alors* $I_x = \mathbb{R}$.

Démonstration. Nous allons montrer que $\tau_2(x) = +\infty$ si γ_x^+ est bornée. Supposons au contraire que $\tau_2(x) < +\infty$ et considérons une suite $(t_i)_i$ convergeant vers $\tau_2(x)$ dans $[0, \tau_2(x)]$. Comme $\overline{\gamma_x^+}$ est compact et que $t \mapsto \varphi(t, x)$ est continue, il existe $y \in \overline{\gamma_x^+}$ et une sous-suite de $(t_i)_i$, que nous noterons encore $(t_i)_i$, telle que

$$\varphi(t_i, x) \to y \in \overline{\gamma_x^+}.$$

Par le théorème de Cauchy 3.2, il existe un voisinage compact B de y et un $\tau > 0$ tels que, pour tout $z \in B$, la trajectoire par z est définie sur $[-\tau, \tau]$ (figure 3.6).

Revenons à la suite $(t_i)_i$. Si i est assez grand, on a : $|t_i - \tau_2(x)| < \tau/2$ et $\varphi(t_i, x) \in B$. On peut alors prolonger la trajectoire de $\varphi(t_i, x)$ jusqu'à $\tau + t_i > \tau_2(x)$. D'où la contradiction avec l'hypothèse que $\tau_2(x) < +\infty$, il suit donc que $\tau_2(x) = +\infty$.

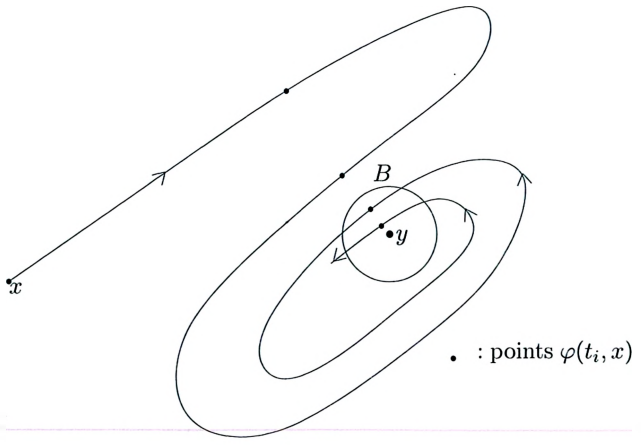


FIGURE 3.6

Supposons maintenant que l'orbite γ_x elle-même soit bornée. On peut répéter le raisonnement précédent à la demi-orbite négative $\gamma_x^- = \bigcup_{t\leq 0} \varphi(t,x)$, qui est la demi-orbite positive par x du champ $-X$. On obtient ainsi que $-\tau_1(x) = -\infty$ et donc que $I_x = \mathbb{R}$. ∎

Exemple d'application. Supposons que l'on a une boule B de rayon fini, telle que le champ soit rentrant sur ∂B. De cette condition suit que les trajectoires issues des points de B ne peuvent pas sortir de B, ce qui implique que la demi-orbite γ_x^+ est contenue dans le borné B pour tout x dans B. On peut donc lui appliquer le critère de complétude. Il en résulte que X est complet sur B *pour les* $t \geq 0$ (ce qui veut dire que $[0, +\infty[\subseteq I_x$, pour tout $x \in B$).

Grâce au critère de complétude que l'on vient d'établir, on déduit le corollaire 3.1

Corollaire 3.1. *Si X est un champ borné sur $\mathbb{R}^n$, alors le flot de X est complet.*

Démonstration. L'hypothèse signifie qu'il existe $M > 0$ tel que $\forall y \in \mathbb{R}^n$, on ait $\|X(y)\| < M$.

Soit alors x un point quelconque dans $\mathbb{R}^n$ et $\varphi(t,x)$ la trajectoire maximale par x définie sur $I_x =]-\tau_1(x), \tau_2(x)[$. Montrons par exemple que $\tau_2(x) = +\infty$ (la démonstration serait la même pour prouver que $\tau_1(x) = -\infty$ en remplaçant X par $-X$).

Faisons un raisonnement par l'absurde en supposant que $\tau_2(x) < +\infty$. Utilisons sur $[0,t] \subset [0,\tau_2(x)[$ le théorème des accroissements finis

$$\|\varphi(t,x) - x\| \leq \sup_{\theta\in[0,t]} \left\|\frac{\partial\varphi}{\partial\theta}(\theta,x)\right\| \cdot |t|\,.$$

Or nous avons $\frac{d\varphi}{d\theta}(\theta,x) = X(\varphi(\theta,x))$ et $\|X(y)\| \leq M$, $\forall y \in \mathbb{R}^n$. Il vient donc

$$\|\varphi(t,x) - x\| \leq M \cdot |t| \leq M\tau_2(x), \qquad \forall t \in [0,\tau_2(x)[.$$

Le flot φ est donc borné et, d'après le critère précédent, on a $\tau_2(x) = +\infty$. Il en résulte une contradiction et la preuve du corollaire. ∎

Maintenant il est facile de construire une fonction $f > 0$, de classe C^∞, telle que $Y = fX$ soit complet si X ne l'est pas.

Corollaire 3.2. *Tout champ de vecteurs C^∞ sur $\mathbb{R}^n$ est C^∞-équivalent à un champ à flot complet.*

Démonstration. On définit $f : x \mapsto f(x) = \frac{1}{\sqrt{1+\|X(x)\|^2}}$. Alors évidemment $f > 0$ et f est C^∞ si X est C^∞. Le champ $Y(x) = \frac{1}{\sqrt{1+\|X(x)\|^2}}X(x)$ est naturellement borné sur $\mathbb{R}^n$ car $\|Y(x)\| < 1$ et, d'après le corollaire 3.1, le champ Y est complet. ∎

Pour un champ X défini sur un ouvert U quelconque de $\mathbb{R}^n$, il est plus difficile de construire une fonction f telle que fX soit à flot borné. Par contre, il suit trivialement du corollaire 3.1 que tout champ de vecteurs *à support compact* est complet (le support d'un champ X est le complémentaire de l'ouvert des points où le champ est nul). Cela étant, on peut toujours multiplier un champ par une fonction positive C^∞, égale à 1 sur un compact quelconque de U et à support compact dans U. Cette remarque implique le résultat suivant :

Proposition 3.6. *Soit X un champ de vecteurs C^∞ sur un ouvert U de $\mathbb{R}^n$ et W un sous-ouvert de U, à fermeture compacte dans U. Alors il existe une fonction f de classe C^∞, positive sur U, à support compact dans U et identique à 1 sur W telle que fX soit un champ complet sur U.*

Le champ $Y = fX$ est C^∞-équivalent à X sur l'ouvert W (mais pas sur l'ouvert U en général). Ce résultat permettra de faire des *études locales* sur X dans des ouverts W, en remplaçant X par des champs complets.

3.3.5. Exemples de flots

1. En dimension 1

L'exemple de base est celui du champ linéaire $\lambda x \frac{\partial}{\partial x}$, de flot $\varphi(t,x) = xe^{\lambda t}$. Pour $\lambda \neq 0$, on a deux portraits de phase différents, selon le signe de λ (figure 3.7).

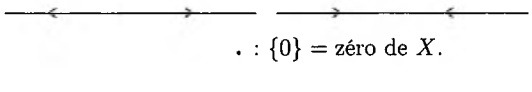


FIGURE 3.7. Portraits de phase du champ $\lambda x \frac{\partial}{\partial x}$: à gauche pour $\lambda > 0$, à droite pour $\lambda < 0$.

Le champ de vecteurs général de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ est de la forme $f(x)\frac{\partial}{\partial x}$, où f est une fonction de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$. Comme un ouvert quelconque de $\mathbb{R}$ est une réunion dénombrable d'intervalles ouverts disjoints et qu'un intervalle ouvert de $\mathbb{R}$ est difféomorphe à $\mathbb{R}$, l'étude d'un champ sur un ouvert quelconque se ramène à l'étude des champs sur $\mathbb{R}$.

Soit $\Sigma_f = \{x \in \mathbb{R} \mid f(x) = 0\}$, l'ensemble des zéros de f. Chacun de ces points est un zéro (ou point singulier) du champ X et il est une orbite (singulière) du champ. Ceci est trivial, cependant ce sera vu en détail dans le prochain chapitre. Le complémentaire de Σ est un ouvert de $\mathbb{R}$ et donc une réunion dénombrable d'intervalles ouverts (éventuellement en nombre fini). Certains de ces intervalles peuvent avoir $+\infty$ et/ou $-\infty$ comme extrémité. On a donc :

$$\mathbb{R} - \Sigma_f = \bigcup_n A_n.$$

où n décrit un ensemble fini, ou bien $\mathbb{N}$, et où chaque A_n est un intervalle ouvert. Montrons que chaque intervalle A_n est une orbite de X. Soit $x \in A_n =]\alpha, \beta[$ et soit γ_x l'orbite de x. Tout d'abord, comme les différentes orbites sont deux à deux disjointes, on a $\gamma_x \subset]\alpha, \beta[$. Ensuite il suit, de l'unicité des trajectoires, que γ_x est un intervalle ouvert : $\gamma_x =]a, b[\subset]\alpha, \beta[$. Nous voulons montrer que $]a, b[=]\alpha, \beta[$. Montrons par exemple que $b = \beta$. Si cela n'était pas, on aurait $a < b < \beta$ et b serait fini. De plus, $X(b) \neq 0$ (puique $b \notin \Sigma_f$) et l'orbite par b contiendrait tout un intervalle $]b-\tau, b+\tau[\subset]\alpha, \beta[$. Cette orbite aurait tout l'intervalle $]b-\tau, b[$ en commun avec l'orbite de x. À nouveau, en raison de l'unicité des trajectoires, cette orbite par x devrait contenir l'intervalle $[b, b+\tau[$, ce qui est en contradiction avec la définition du point b.

Les orbites de X sont donc, d'une part, les zéros de la fonction f et, d'autre part, les intervalles ouverts A_n, composantes connexes du complémentaire

de l'ensemble Σ_f des zéros de f. Il est donc immédiat de déduire le portrait de phase de X, du graphe de f : figure 3.8.

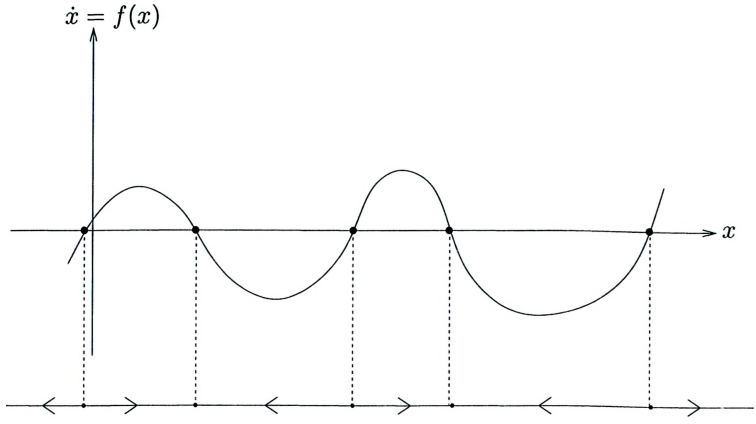


FIGURE 3.8. Portrait de phase du champ $f(x)\frac{\partial}{\partial x}$.

Remarquez que l'on peut obtenir *par quadrature* le flot de X sur chaque intervalle A_n. Soit A_n un de ces intervalles et $x \in A_n$. L'équation différentielle des trajectoires, $\frac{\partial \varphi}{\partial t}(t,x) = f(\varphi(t,x))$, s'intègre en donnant $t = \int_x^{\varphi(t,x)} \frac{d\varphi}{f(\varphi)}$ où $\varphi(t,x) \in A_n$. On obtient le flot $\varphi(t,x)$ en inversant la fonction $t(\xi) = \int_x^{\xi} \frac{d\varphi}{f(\varphi)}$ définie à x fixé sur l'intervalle A_n. Cette inversion est possible car $\frac{dt}{d\xi}(\xi) = \frac{1}{f(\xi)} \neq 0$ si $\xi \in A_n$. Le domaine de définition I_x de la trajectoire maximale par x est l'intervalle ouvert $I_x = t(A_n)$ sur lequel la trajectoire est la fonction $\varphi(t)$, inverse de la fonction $t(\varphi)$. On retrouve que l'orbite de x est l'intervalle $A_n = \varphi(I_x)$ puisque $t(\xi)$ est un difféomorphisme de A_n sur I_x.

Remarque 3.9. Les propriétés qualitatives (par exemple la stabilité des points singuliers) qui seront étudiées dans les chapitres suivants ne dépendent que du portrait de phase, qui est déduit de l'étude du signe de f. L'étude de ces propriétés qualitatives ne nécessite pas l'intégration du champ. Cette remarque illustre parfaitement les idées qualitatives de Poincaré (voir [22]).

2. En dimension supérieure ou égale à 2

En général il n'y a pas de possibilité d'obtenir le flot de X par simple quadrature. C'est justement la raison pour laquelle on développe la théorie qualitative dont il sera question à partir du prochain chapitre. Rappelons cependant que l'on sait intégrer simplement les champs de vecteurs linéaires

et en déduire leur portrait de phase. On a décrit les portraits de phase des champs linéaires génériques dans le chapitre précédent. On verra, dans le chapitre III-3, que l'intérêt des champs de vecteurs génériques vient de ce qu'ils servent de modèle local, aux voisinages des points singuliers génériques de champs de vecteurs quelconques.

Certains champs de dimension supérieure à 2 se ramènent à des champs de vecteurs en dimension 1, et sont donc étudiables de façon élémentaire.

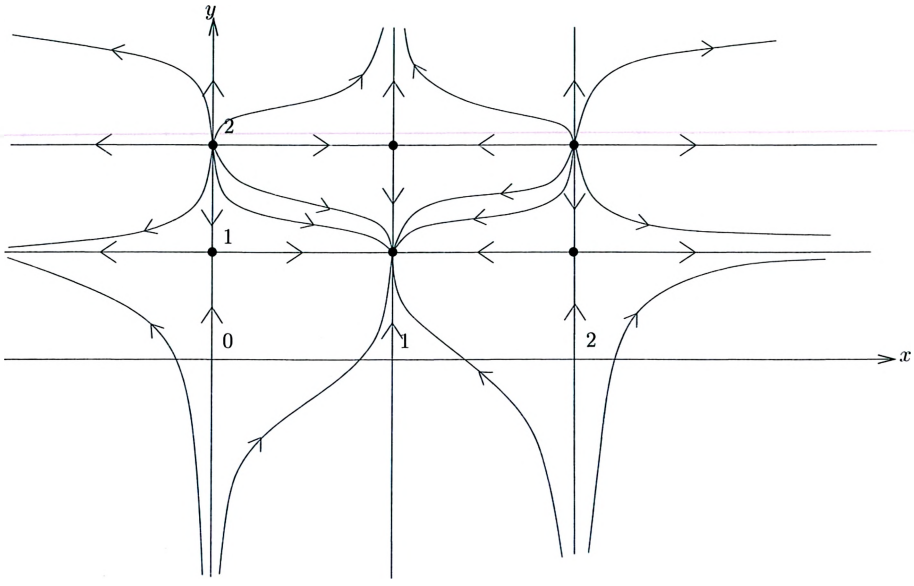


FIGURE 3.9. Portrait de phase du champ X. On n'a dessiné que quelques orbites représentatives.

Considérons par exemple le champ de $\mathbb{R}^2$ *à variables séparées* : $X = f(x)\frac{\partial}{\partial x} + g(y)\frac{\partial}{\partial y}$ avec f, g des fonctions de classe C^∞. Si $\varphi_1(t,x)$ et $\varphi_2(t,y)$ sont les flots des deux champs unidimensionnels $f(x)\frac{\partial}{\partial x}$ et $g(y)\frac{\partial}{\partial y}$, le flot de X est égal à $\varphi(t,(x,y)) = (\varphi_1(t,x), \varphi_2(t,y))$. On en déduit facilement le portrait de phase de X à partir des portraits de phase des champs unidimensionnels (dans la figure 3.9 on a dessiné le portrait de phase du champ $X = 2x(x-1)(x-2)\frac{\partial}{\partial x} + (y-1)(y-2)\frac{\partial}{\partial y}$).

4

ANALYSE QUALITATIVE DES TRAJECTOIRES

4.1. Champ sur une variété, intégrale première

Dans cette section, nous allons étendre la notion de champ de vecteurs aux variétés différentiables. Pour cela, nous allons utiliser les notions d'espace tangent en un point d'une variété, et de différentielle pour une application différentiable entre deux variétés, notions introduites dans le chapitre I-2. Nous examinerons le cas particulier d'un champ de vecteurs tangent à une sous-variété d'un ouvert $U \subset \mathbb{R}^n$ et la notion d'intégrale première. Sauf mention expresse du contraire, et cela ne sera pas explicitement précisé, toute variété, toute application, tout champ de vecteurs, etc. sont supposés de classe $\mathcal{C}^\infty$.

Définition 4.1. Soit M une variété de dimension p, munie d'un atlas $\{(U_i, \varphi_i)\}_{i\in I}$. Un champ de vecteurs X sur M sera défini par la donnée d'une famille de champs de vecteurs $(X_i)_{i\in I}$ telle que X_i est défini sur chaque ouvert $\Omega_i = \varphi_i(U_i) \subset \mathbb{R}^p$, et telle que les différents champs X_i soient compatibles dans les intersections de cartes au sens suivant : si $U_i \cap U_j \neq \emptyset$, alors $X_j = (\varphi_j \circ \varphi_i^{-1})_*(X_i)$ sur $\varphi_i(U_i \cap U_j) \subset \Omega_i$ (figure 4.1). (L'image $G_*(Z)$ par difféomorphisme d'un champ de vecteurs Z sur un ouvert euclidien a été définie dans le paragraphe 1.2.)

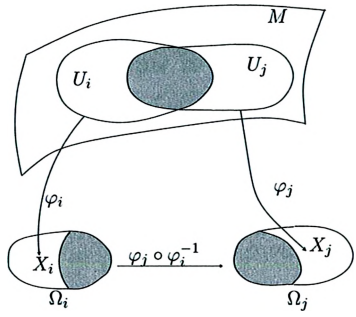


FIGURE 4.1

L'inconvénient de cette définition est de dépendre du choix de l'atlas. Pour nous débarrasser de cette limitation, nous allons utiliser la notion de vecteur tangent en un point quelconque de M, ainsi que la notation de l'espace (ou fibré) tangent $TM = \coprod_x T_x M$, où $T_x M$ est l'espace tangent au point $x \in M$ (figure 4.2).

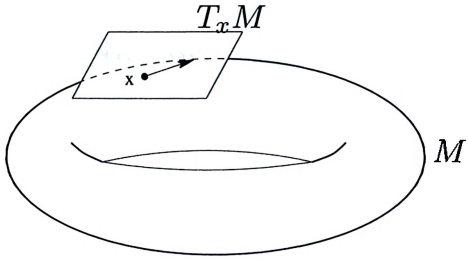


FIGURE 4.2

Considérons un champ de vecteurs X, comme dans la définition 4.1. Soit $m \in U_i$, U_i étant le domaine d'une des cartes de cette définition. Dans cette carte, le vecteur $X_i(\varphi_i(m))$ est identifié à un vecteur $X(m) \in T_m M$. D'autre part, si (U_j, φ_j) est une autre carte telle que $m \in U_i \cap U_j$, la condition de compatibilité mentionnée dans la définition 4.1 implique que ce vecteur $X(m)$ est le même pour les deux cartes. On associe ainsi au champ de vecteurs X une application $m \in M \mapsto X(m) \in TM$, telle que $X(m) \in T_m M$ pour tout $m \in M$. Pour traduire cette condition, on dit que cette application est une section du fibré tangent. On identifie le champ X avec la section $m \mapsto X(m)$ ainsi définie. Cette façon de considérer un champ de vecteurs comme une section de TM est indépendante du choix des cartes. Elle est malheureusement un peu abstraite et, pour travailler avec le champ de vecteurs X, il sera indispensable de se rappeler que, dans chaque carte, ce champ s'identifie avec un champ de vecteurs sur un ouvert euclidien.

Toutes les notions introduites pour les champs de vecteurs sur les ouverts euclidiens s'étendent sans difficulté aux champs de vecteurs sur les variétés. Tout d'abord, on a de façon directe que les théorèmes d'existence et d'unicité locales 3.2 et 3.3 sont valables sur les ouverts domaines de cartes. Il en résulte l'existence et l'unicité des trajectoires maximales et la notion de flot. La même démonstration que celle donnée dans le théorème 3.4 est valable pour le flot d'un champ sur une variété M : le flot est une application définie sur un ouvert W de $\mathbb{R} \times M$, et il est de même classe de différentiabilité que le champ. Le critère de complétude (lemme 3.4) se démontre comme dans le cas euclidien. On en déduit le résultat suivant :

Proposition 4.1. *Tout champ de vecteurs sur une variété compacte est à flot complet.*

Soit G un difféomorphisme entre deux variétés M et N. On peut maintenant utiliser la notion de différentielle introduite dans le paragraphe I-1.1.1 pour définir l'image d'un champ par G, exactement de la même façon qu'entre deux ouverts euclidiens (voir définition 1.2) :

Définition 4.2. Soit G un difféomorphisme entre deux variétés M et N. Soit X un champ de vecteurs sur M. Alors, on appelle champ de vecteurs image de X par G, le champ Y défini par

$$Y(y) = dG(G^{-1}(y))[X(G^{-1}(y))], \quad \forall y \in N. \tag{4.1}$$

On note le champ Y par $G_*(X)$.

Remarque 4.1.

1. En utilisant des cartes, on démontre que la formule (4.1) définit bien un champ sur N.

2. La formule (4.1) est la même que celle utilisée pour les difféomorphismes entre ouverts euclidiens (paragraphe 1.2). Par la même démonstration que dans ce cas, on montre que G envoie les orbites de X sur celles de $Y = G_*(X)$. Plus précisément, si $\varphi^X(t,x)$ est le flot de X et $\varphi^Y(t,y)$ est le flot de Y, on a :
$$G \circ \varphi^X(t,x) = \varphi^Y(t, G(x)),$$
pour tout (t,x) dans le domaine de définition du flot de X.

3. Si M est une variété, on désigne par $\mathcal{X}(M)$ l'ensemble de tous les champs de vecteurs sur M. Cet ensemble est naturellement un espace vectoriel, car, en utilisant la structure vectorielle de chaque espace tangent T_xM, on peut définir la somme de deux champs de vecteurs et le produit par un réel d'un champ de vecteurs en posant : $(X+Y)(x) = X(x) + Y(x)$ et $(\lambda X)(x) = \lambda X(x)$. Clairement, l'application $X \mapsto G_*(X)$ est linéaire. Cette application a pour inverse $Y \to (G^{-1})_*(Y)$: G_* est donc un isomorphisme linéaire de $\mathcal{X}(M)$ sur $\mathcal{X}(N)$.

Exemple des champs sur les tores. Considérons le tore $T^n = \mathbb{R}^n/\mathbb{Z}^n$ comme il a été défini dans le sous-paragraphe I-2.1.3 et soit $\rho : \mathbb{R}^n \mapsto T^n$ la projection canonique. On remarque que, pour tout $m \in \mathbb{R}^n$, la différentielle $d\rho(m)$ est un isomorphisme de $T_m\mathbb{R}^n \equiv \mathbb{R}^n$ sur $T_{\rho(m)}T^n$. Si X est un champ de vecteurs sur T^n, alors on peut le relever en un champ $\widetilde{X}$ sur $\mathbb{R}^n$ en posant : $\widetilde{X}(m) = d\rho(m)^{-1}(X(\rho(m)))$. Ce champ de vecteurs est invariant par translation par les éléments de $\mathbb{Z}^n$. Inversement, si $\widetilde{X}$ est un champ de vecteurs sur $\mathbb{R}^n$ invariant par les translations entières, il se projette en un champ de vecteur X de T^n vérifiant : $X(\rho(m)) = d\rho(m)[\widetilde{X}(m)]$. *On peut donc identifier les champs de vecteurs sur T^n, avec les champs de $\mathbb{R}^n$ invariants par les translations par les éléments de $\mathbb{Z}^n$.* Par exemple, tout champ de vecteurs constant sur $\mathbb{R}^n$ s'identifie avec un champ de vecteurs de T^n. On appelle ces champs de vecteurs : *champs de Kronecker du tore.* Ils jouent un rôle essentiel en mécanique. On y reviendra dans la prochaine section. On peut généraliser cette identification à toute variété construite par quotient d'une autre par l'action d'un groupe discret (les translations de $\mathbb{Z}^n$ dans le cas du tore T^n), agissant librement et proprement (voir le sous-paragraphe I-2.1.3).

Nous allons maintenant revenir à une situation plus concrète en considérant des sous-variétés d'ouverts euclidiens. Dans ce cas, nous avons un moyen plus direct de nous donner un champ sur M :

Définition 4.3. Soit M une sous-variété de l'ouvert $U \subset \mathbb{R}^n$ et X un champ de vecteurs sur U. On dit que X est tangent à M si, pour tout $m \in M$, on a $X(m) \in T_mM$.

Considérons un champ tangent à une sous-variété M, comme dans la définition 4.3. En considérant un recouvrement de M par des ouverts de trivialisation, puis l'atlas associé de M, il est clair que la restriction $X|M$ de X aux points de M est un champ de vecteurs sur la variété plongée M. En effet, si (W_i, Φ_i) est l'un des ouverts de trivialisation, il lui correspond la carte $(W_i \cap M, \varphi_i)$ avec $\varphi_i : W \cap M \mapsto \mathbb{R}^p \times \{0\} \subset \mathbb{R}^p \times \mathbb{R}^{n-p}$, et le champ $X|M$ est donné par la collection

des champs $X_i = (\Phi_i)_*(X|\mathbb{R}^p \times \{0\})$ (remarquez que, en utilisant la définition de champ image introduite plus haut, on peut écrire : $X_i = (\varphi_i)_*(X|W_i \cap M)$. On peut montrer qu'inversement, tout champ X sur une variété M à topologie à base dénombrable s'obtient, par restriction, par un plongement i de M dans un espace euclidien $\mathbb{R}^n$, d'un champ de vecteurs de $\mathbb{R}^n$ tangent à la sous-variété $i(M)$. Ce qui fait qu'il ne serait pas restrictif de ne considérer que de tels champs tangents à des sous-variétés.

Définition 4.4. Soit X un champ de vecteurs sur un ouvert U de $\mathbb{R}^n$. On dit que la fonction différentiable f est une *intégrale première* de X si, en tout point $x \in U$, on a $X \cdot f(x) = 0$.

Soit φ le flot de X. Il suit du lemme 1.1 que $X \cdot f(\varphi(t,x)) = \frac{\partial(f\circ\varphi)}{\partial t}(t,x)$. Donc, si f est intégrale première, $f \circ \varphi(t,x)$ est constante en temps, et on peut écrire que $f \circ \varphi(t,x) = f(x)$ pour tout t ; autrement dit, la fonction f est laissée invariante par le flot du champ. Si α est une valeur régulière de f, le champ X est tangent à l'hypersurface $V = f^{-1}(\alpha)$.

Remarque 4.2. Ce phénomène arrive fréquemment. Il n'est pas rare que des surfaces soient tangentes au flot d'un champ, par exemple si le champ admet des intégrales premières : pensez aux équations différentielles de la mécanique.

4.2. Type topologique des trajectoires

Pour simplifier l'étude, on va supposer que X est de classe $\mathcal{C}^\infty$ et complet sur une variété M. Nous avons vu, par exemple au corollaire 3.2, que tout champ de $\mathbb{R}^n$ de classe $\mathcal{C}^\infty$ est $\mathcal{C}^\infty$-équivalent à un champ complet, et, comme toutes les notions qualitatives que nous allons introduire ne dépendent que de l'espace des orbites, et non pas du temps explicitement, on ne perd rien en faisant cette hypothèse.

Soit donc $\varphi : W = R \times M \mapsto M$ le flot de X, et γ_x la trajectoire par x.

Définition 4.5. *Le groupe des périodes en* x, $P_x = \{t \in \mathbb{R} \mid \varphi(t,x) = x\}$, est l'ensemble des temps de retour en x, de la trajectoire par x.

Propriétés de P_x.

(a) $P_x = P_y$, *si* $y \in \gamma_x$: P_x *ne dépend que de l'orbite* γ_x *et non du point sur l'orbite* γ_x.

En effet, si $t \in P_x$, on a $\varphi(t,x) = x$; par ailleurs, si $y \in \gamma_x$, il existe un τ tel que $y = \varphi(\tau, x)$. Alors (formule de translation) il vient :

$$y = \varphi(\tau, x) = \varphi(\tau, \varphi(t,x)) = \varphi(t+\tau, x) = \varphi(t, \varphi(\tau, x)) = \varphi(t, y),$$

autrement dit $t \in P_y$ si $y \in \gamma_x$. Il est clair que l'on montrerait de même que $t \in P_y$ entraîne $t \in P_x$ si $x \in \gamma_y$: par l'unicité des orbites, $y \in \gamma_x$ est équivalent à $x \in \gamma_y$ (c'est la même orbite !). On en conclut que $P_x = P_y$, si $y \in \gamma_x$.

(b) *P_x est un sous-groupe additif de $\mathbb{R}$* : si $t_1, t_2 \in P_x$ alors $\varphi_{t_1 - t_2}(x) = \varphi_{t_1} \circ \varphi_{-t_2}(x) = \varphi_{t_1}(x) = x$, ce qui entraîne que $t_1 - t_2 \in P_x$.

(c) *P_x est fermé dans $\mathbb{R}$*. En effet P_x est défini par l'équation $\{\varphi(t,x) - x = 0\}$ où $t \mapsto \varphi(t,x) - x$ est une application continue, donc P_x est fermé comme image réciproque du fermé $\{0\}$.

P_x est donc un *sous-groupe additif fermé* de $\mathbb{R}$. La proposition suivante caractérise ces sous-groupes :

Proposition 4.2. *Soit G un sous-groupe additif fermé de $\mathbb{R}$. On a alors les trois possibilités suivantes : $G = \{0\}$, $G = \mathbb{R}$, ou bien G est le sous-groupe discret*

$$\mathbb{Z}T = \{nT; n \in \mathbb{Z}\} \text{ pour un } T > 0.$$

Démonstration. Supposons $G \neq \{0\}$. Soit $G_+ = \{t \mid t > 0 \text{ et } t \in G\}$. Cet ensemble n'est pas vide car $G \neq \{0\}$ par hypothèse et possède des éléments positifs.

Alors $T = \text{Inf } G_+$ est fini : on a $T > 0$ ou bien $T = 0$.

Premier cas : $T > 0$. Soit $t \in G$, on peut écrire $t = pT + r$ pour un $p \in \mathbb{N}$ et un reste r, $0 \leq r < T$. Puisque $pT \in G$ et $t \in G$, alors $r \in G$, car G est un sous-groupe additif de $\mathbb{R}$. Comme $T \in G$ est l'infimum des $t \in G$, $t > 0$, on doit avoir $r = 0$. Il en résulte que $t = pT$ et donc que $G = \mathbb{Z}T$ (le sous-groupe additif de $\mathbb{R}$ engendré par T, qui est égal à l'ensemble $\{\ldots -2T, -T, 0, T, 2T, \ldots\}$).

Deuxième cas : $T = 0$. On peut trouver des $t_i > 0$, $t_i \in G_+$ avec $(t_i)_i \to 0$. Soit $t \in \mathbb{R}$ quelconque. Pour tout i, il existe un nombre $p_i \in \mathbb{Z}$ tel que $p_i t_i \leq t < (p_i + 1)t_i$. Posons $g_i = p_i t_i$. On a $|g_i - t| < t_i$, ce qui signifie que $(g_i)_i \to t$ puisque $t_i \to 0$; comme G est fermé et que $g_i \in G$, il vient $t \in \overline{G} = G$. Puisque $t \in \mathbb{R}$ est arbitraire, $G = \overline{G} = \mathbb{R}$, ce qu'il fallait démontrer. ∎

On peut caractériser les deux cas $P_x = \mathbb{R}$ et $\mathbb{Z}T$, $T > 0$ de la façon suivante :

Proposition 4.3. *Nous avons les équivalences suivantes :*

1. $P_x = \mathbb{R} \iff X(x) = 0 \iff \gamma_x = \{x\}$.

2. $P_x = \mathbb{Z}T$, $T > 0 \iff \gamma_x$ *est homéomorphe au cercle* S^1 *(en fait* γ_x *est* difféomorphe *au cercle* S^1 *; on dit aussi que* γ_x *est une courbe compacte simple de classe* C^∞ *dans* M*).*

Définition 4.6. Un point x tel que $X(x) = 0$ est dit *point critique* du champ X, ou *zéro* de X, ou *point d'équilibre*, ou *point singulier* ou *état stationnaire* du champ X, et $\{x\}$ est appelé *orbite singulière* de X. Si $P_x = \mathbb{Z}T$, $T > 0$, l'orbite γ_x est appelée *orbite périodique* du champ. Dans ce dernier cas, T est appelé *la période (minimale)* de γ_x.

Remarque 4.3. Il suit du point 1 de la proposition 4.3, que si $X(x) \neq 0$ alors, pour tout $y \in \gamma_x$, on a aussi $X(y) \neq 0$. En effet, si $X(y) = 0$, alors $\{y\}$ serait une orbite singulière disjointe de l'orbite γ_x car cette dernière orbite contient au moins un point x non singulier (rappelons qu'en vertu du théorème d'unicité de Cauchy, les orbites de X sont deux à deux disjointes).

Preuve de la proposition 4.3. Le point 1 est presque évident. Nous avons :

$$\gamma_x = \{x\} \iff P_x = \mathbb{R} \iff \varphi(t,x) = x,\ \forall t \in \mathbb{R} \iff \dots$$

$$\dots \iff \frac{\partial \varphi}{\partial t}(t,x) = 0,\ \forall t \in \mathbb{R} \iff X(\varphi(t,x)) = 0,$$

et en particulier $X(x) = 0$. Inversement, si $X(x) = 0$, la courbe $\varphi(t) \equiv x$ vérifie $\varphi(0) = x$ et $\frac{\partial \varphi}{\partial t}(t) = 0$, $\forall t \in \mathbb{R}$. Cette courbe est donc la trajectoire par x, ce qui implique que $\gamma_x = \{x\}$ et conclut le point 1.

Supposons maintenant que $P_x = \mathbb{Z}T$ avec $T > 0$. Posons $\varphi(t) = \varphi(t,x)$. Cette application est périodique de période T. Elle *se factorise* donc par la projection canonique Π de $\mathbb{R}$ sur $\mathbb{R}/\mathbb{Z}T$. Ce quotient $\mathbb{R}/\mathbb{Z}T$ est une variété de dimension 1, difféomorphe à $S^1 = \mathbb{R}/\mathbb{Z}$ par le difféomorphisme $u \in \mathbb{R}/\mathbb{Z} \mapsto T.u \in \mathbb{R}/\mathbb{Z}T$. (on identifiera maintenant le quotient $\mathbb{R}/\mathbb{Z}T$ avec S^1). On rappelle que S^1 est difféomorphe au cercle trigonométrique par le difféomorphisme $t \in S^1 \mapsto e^{2i\pi t} \in \Gamma = \{z \in \mathbb{C}, \mid z \mid = 1\}$. La factorisation permet de définir une application $\overline{\varphi} : S^1 \mapsto U$, telle que $\varphi = \overline{\varphi} \circ \Pi$ (figure 4.3).

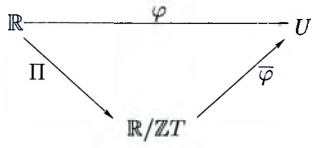


FIGURE 4.3

Comme φ est continue, il en est de même pour $\overline{\varphi}$. De plus, cette application est injective car T est la période minimale et que l'on a identifié S^1 avec $\mathbb{R}/\mathbb{Z}T$. Comme $\overline{\varphi}$ est une application injective et continue du compact S^1 sur son image γ_x, c'est un homéomorphisme de S^1 sur l'orbite γ_x.

L'application φ étant de classe $\mathcal{C}^\infty$, il suit, de la définition même de la structure de variété sur S^1, que $\overline{\varphi}$ est aussi de classe $\mathcal{C}^\infty$. Maintenant, observons que φ (et donc $\overline{\varphi}$) est de rang 1, c'est-à-dire que pour tout t, $\frac{d\varphi}{dt}(t) \neq 0$. En effet, on a par hypothèse que $\frac{d\varphi}{dt}(0) = X(x) \neq 0$ et donc $\frac{d\varphi}{dt}(t) \neq 0$ pour tout t, comme on l'a vu dans la remarque 4.3. L'application $\overline{\varphi}$ est donc une immersion injective $\mathcal{C}^\infty$ de la variété compacte S^1 dans M : c'est donc un plongement de classe $\mathcal{C}^\infty$, et γ_x est une courbe simple compacte de classe $\mathcal{C}^\infty$ de U (c'est-à-dire une courbe $\mathcal{C}^\infty$-difféomorphe au cercle S^1).

La réciproque : γ_x homéomorphe à $S^1 \Rightarrow P_x = \mathbb{Z}T$, est plus délicate à démontrer. On l'admettra. Ce qui termine la preuve de la proposition 4.3.

Le cas $P_x = \{0\}$ correspond à l'obtention d'une orbite γ_x *apériodique*. C'est le cas qui peut présenter une topologie la plus compliquée. La trajectoire définie par φ est de rang un, c'est-à-dire que, pour tout t, $X(\varphi(t)) \neq 0$; l'orbite est donc une *immersion injective de classe* $\mathcal{C}^\infty$ de $\mathbb{R}$ dans U. Cette orbite est une courbe en bijection avec $\mathbb{R}$, mais sa topologie peut être très compliquée car φ n'est pas nécessairement un plongement (en fait la topologie induite par U peut être strictement moins fine que la topologie de $\mathbb{R}$). Un exemple de ce phénomène sera donné dans le prochain chapitre où l'on expliquera en détail la construction *du flot irrationel sur le tore* T^2, qui est flot d'un champ de Kronecker particulier. Ce champ peut être obtenu par restriction à une sous-variété de $\mathbb{R}^3$ ou de $\mathbb{R}^4$. Ce dernier cas apparaît dans les équations linéaires de la mécanique, à deux degrés de liberté. Les notions qualitatives d'ensemble limite, récurrence... que nous présenterons dans le prochain chapitre, ont été essentiellement introduites pour étudier de telles orbites.

4.3. Théorème du voisinage tubulaire

Soit X un champ de vecteurs, défini sur une variété M de dimension n, que l'on supposera, par commodité, de classe $\mathcal{C}^\infty$. Soit $p \in M$ un point où $X(p) \neq 0$. Un tel point est appelé *point régulier du champ* X, par opposition aux points singuliers où le champ s'annule. Nous allons montrer qu'au voisinage d'un point régulier, le champ a pour modèle le champ constant $\frac{\partial}{\partial x_1}$.

Définition 4.7. Une *section locale* Σ (non nécessairement plane) de X en p est une sous-variété difféomorphe au disque unité D^{n-1} de $\mathbb{R}^{n-1}$ (disque ou boule de rayon 1 dans $\mathbb{R}^{n-1}$) telle que $p \in \text{int}\Sigma$ ($\text{int}\Sigma$ désigne l'intérieur de Σ, égal à $\Sigma - \partial\Sigma$ où $\partial\Sigma$ est le bord de Σ). On suppose que $X(x)$ est *transverse* à $T_x\Sigma$ (espace tangent de dimension $(n-1)$) pour tout $x \in \Sigma$: on écrira $X \pitchfork_x \Sigma$ pour tout $x \in \Sigma$.

Nous pouvons illustrer cette définition par la figure 4.4 suivante.

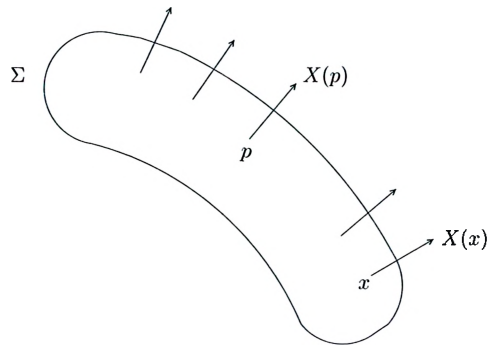


FIGURE 4.4

Remarque 4.4. Dire que $X(x)$ est transverse à $T_x\Sigma$ signifie que la droite $\mathbb{R}\{X(x)\} \subset T_xM$ est transverse au $(n-1)$-plan $T_x\Sigma$ au sens défini dans la partie I. Ici, cela revient à dire que $\mathbb{R}\{X(x)\} \oplus T_x\Sigma = T_xM$ (pour x fixé, on désigne par $\mathbb{R}\{X(x)\}$ la droite vectorielle $\{uX(x) \mid u \in \mathbb{R}\}$).

Proposition 4.4. *Pour tout point p régulier de X, on peut trouver une section locale.*

Démonstration. Tout d'abord, en se restreignant à une carte contenant p, on peut supposer que X est défini sur un ouvert U de $\mathbb{R}^n$. Considérons un hyperplan H de $\mathbb{R}^n$ de dimension $(n-1)$ passant par p, transverse à $X(p)$; il est toujours possible de trouver un tel hyperplan en complétant le vecteur $X(p) \neq 0$ en une base vectorielle de $\mathbb{R}^n$ (voir figure 4.5). Alors, puisque l'application $x \in H \mapsto X(x)$ est continue, il existe $\varepsilon > 0$ tel que $X(x) \pitchfork_x H$ pour $\|x - p\| \leq \varepsilon$ ($\|\cdot\|$ norme euclidienne de $\mathbb{R}^n$).

Pour obtenir une section locale par p, il suffira alors de prendre le disque plan $\Sigma = \{x \in H \mid \|x - p\| \leq \varepsilon\}$ qui est bien difféomorphe au disque unité D^{n-1} de $\mathbb{R}^{n-1}$. Ce qui termine la preuve. ∎

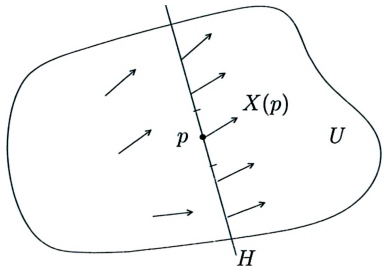


FIGURE 4.5

Remarque 4.5. Dans la pratique, on peut avoir besoin de sections locales qui soient difféomorphes au disque sans être planes (dans la carte choisie), la construction précédente peut être aisément généralisée. Considérons par exemple une hypersurface V dans M (c'est-à-dire une sous-variété de dimension $(n-1)$), passant par p et transverse à $X(p)$ en ce point. Soit $\mathrm{dist}(x,y)$ une distance sur M et définissant la topologie (par exemple associée à une métrique riemannienne, voir la définition I-2.15). Alors, si $\varepsilon > 0$ est assez petit, l'ensemble $\Sigma = \{x \in V \mid \mathrm{dist}(p,x) \leq \varepsilon\}$ est une section à X par p.

Nous voulons montrer que, au voisinage de chaque point p régulier, on peut trouver une *carte* (voisinage paramétré) dans laquelle le champ est trivial (autrement dit est un *champ constant*). Introduisons pour ce faire la définition importante suivante :

Définition 4.8. Soit $p \in M$ un point régulier de X. Un *voisinage tubulaire* (flow-box en anglais) B de X, au voisinage de p, est un voisinage avec une paramétrisation (c'est-à-dire avec un difféomorphisme C^∞)

$$\Phi : (Y,u) \in D^{n-1} \times [-\tau,+\tau] \longmapsto B \subset M,$$

tel que $\Phi_* \left(\frac{\partial}{\partial u}\right) = X$ et $\Phi(0,0) = p$ (figure 4.6).

Il est utile de faire une série de remarques à propos des notations et du sens de cette définition.

Remarque 4.6.

1. On entend, par difféomorphisme défini sur le fermé $D^{n-1} \times [-\tau,+\tau]$, une application qui s'étend en un difféomorphisme d'un voisinage ouvert de $D^{n-1} \times [-\tau,+\tau]$ sur son image, comme il a été défini dans la partie I.

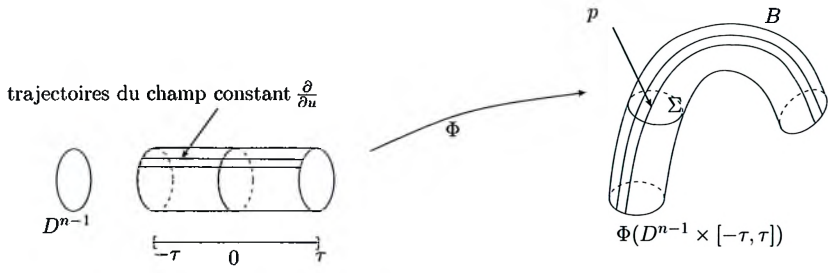


FIGURE 4.6. Voisinage tubulaire B.

2. La notation $\Phi_* X$ désigne le transport du champ X par le difféomorphisme Φ (voir paragraphe 1.2). La relation $\Phi_* \left(\frac{\partial}{\partial u}\right) = X$ signifie que le champ de vecteurs X se lit $\frac{\partial}{\partial u}$ dans des coordonnées locales $(Y, u) \in D^{n-1} \times [-\tau, \tau]$ de $\mathbb{R}^n$.

   On rappelle qu'un difféomorphisme, d'un ouvert de $\mathbb{R}^n$ sur son image dans une variété M, est la terminologie moderne pour le *choix d'un système de coordonnées locales*, que l'on considère maintenant être l'*inverse d'une application de carte* : ici, par exemple, $(Y, u) \in D^{n-1} \times [-\tau, \tau]$ sont de nouvelles coordonnées où le champ va s'écrire *comme le champ constant* $\frac{\partial}{\partial u}$.

3. La classe de Φ est la même que celle de X, c'est-à-dire de classe $\mathcal{C}^\infty$. Dans ce qui suit, il suffirait que X soit de classe $\mathcal{C}^1$, ainsi que Φ, pour pouvoir établir l'existence de voisinages tubulaires et des applications de Poincaré.

4. Le champ $\frac{\partial}{\partial u}$ est un champ constant de composantes $(0, \ldots, 0, 1)$ (horizontal sur la figure 4.6), c'est-à-dire de composantes nulles en y dans D^{n-1} et 1 en u. Par intégration en temps, ce champ $\frac{\partial}{\partial u}$ fournit la solution $u(t) = u_0 + t, Y(t) = Y_0$, à partir de la condition initiale (Y_0, u_0). Si on prend $u_0 = 0$, le temps t joue le rôle de la coordonnée u.

5. Il suit, de cette définition, que $\Sigma = \Phi(D^{n-1} \times \{0\})$ est une *section locale* par le point p, car $D^{n-1} \times \{0\}$ est transverse à $\frac{\partial}{\partial u}$ et Φ est un difféomorphisme (conservant donc les propriétés différentielles).

6. On dit que le voisinage tubulaire B exprimé en les coordonnées $(Y, u) \in D^{n-1} \times [-\tau, \tau]$ est une « carte de trivialisation » du champ de vecteurs X, puisqu'en ces coordonnées, X prend l'expression la plus simple possible, à savoir un champ constant en la coordonnée u.

Nous allons maintenant établir l'existence de voisinages tubulaires au voisinage de tout point régulier p d'un champ de vecteurs X, défini sur une variété M quelconque de dimension n.

Proposition 4.5. *En tout point régulier $p \in M$ d'un champ de vecteurs X, on peut trouver des voisinages tubulaires.*

Démonstration. Par le choix d'une carte contenant p, on peut supposer que X est défini sur un ouvert U de $\mathbb{R}^n$. Choisissons une section locale passant par le point p image de l'immersion injective $\psi : v \in D^{n-1} \mapsto \psi(v) \in M$, avec $v = (v_1, \ldots, v_{n-1})$ et $\psi(0) = p$. Soit φ le flot de X. Considérons l'application définie par $\Psi : (v, u) \mapsto \Psi(v, u) = \varphi(u, \psi(v))$. Cette application de classe C^∞ est définie, pour $u_0 > 0$ assez petit, sur l'ouvert $D^{n-1} \times]-u_0, u_0[$. Soit

$$Y_i = \Psi_*\Big(\frac{\partial}{\partial v_i}\Big)(0,0) = \frac{\partial \Psi}{\partial v_i}(0,0) \in \mathbb{R}^n,$$

pour $i = 1, \ldots, n-1$. Comme $\psi(D^{n-1})$ est une section locale par p, le système de vecteurs $\{Y_1, \ldots, Y_{n-1}, X(p)\}$ est, par la définition 4.7, une base de $\mathbb{R}^n$. Nous allons évaluer la différentielle $d\Psi(0,0)$ sur la base $\{\frac{\partial}{\partial v_1}, \ldots, \frac{\partial}{\partial v_{n-1}}, \frac{\partial}{\partial u}\}$. Comme $\varphi(0, x) \equiv x$, on a $d\Psi(0,0)[\frac{\partial}{\partial v_i}] = Y_i$, pour $i = 1, \ldots, n-1$ et $d\Psi(0,0)[\frac{\partial}{\partial u}] = X(p)$. L'application linéaire $d\Psi(0,0)$ est donc un isomorphisme linéaire de $\mathbb{R}^n$.

On peut donc appliquer le théorème de l'inverse I-1.3 à Ψ au point $0 \in \mathbb{R}^n$: on peut choisir le disque D^{n-1} et u_0 tel que Ψ soit un difféomorphisme de classe C^∞ sur son image $B \subset U$.

Il reste à vérifier la relation $\Psi_*\left(\frac{\partial}{\partial u}\right) = X$. On a :

$$\Psi_*\left(\frac{\partial}{\partial u}\right)(v,u) = d\Psi(v,u)[\frac{\partial}{\partial u}](v,u) = \ldots$$

$$\ldots = \frac{\partial \Psi}{\partial u}(v,u) = \frac{\partial \varphi}{\partial t}(u, \psi(v)) = X(\varphi(u, \psi(v))).$$

Comme, par définition, $\Psi(v,u) = \varphi(u, \psi(v))$, cela signifie que $\Psi_*\left(\frac{\partial}{\partial u}\right) = X|_B$. Ce qui termine la démonstration. ∎

Remarque 4.7. À cause de la formule $\Psi_*\left(\frac{\partial}{\partial u}\right) = X$, on voit que les lignes $\{v =$ constante$\}$ sont envoyées sur les trajectoires de X. C'est le flot lui-même qui sert à définir la carte de trivialisation locale. On a choisi la carte de façon que, partant d'une section transverse au champ, on se déplace le long des trajectoires de telle manière que les lignes ainsi obtenues soient, par définition de la nouvelle carte, des droites parallèles à l'axe Ou, dans les coordonnées de la nouvelle variable $q = (v, u) \in D^{n-1} \times [-\tau, \tau]$.

Remarque 4.8. Le voisinage tubulaire que l'on peut construire, en un point p régulier, est loin d'être unique et peut être adapté aux besoins. Au voisinage tubulaire $B = \Psi(D^{n-1} \times [-u_0, u_0])$ est associée une section locale $\Sigma = \Psi(D^{n-1} \times \{0\})$ et le temps $\tau = u_0$. Les deux choix suivants sont intéressants dans la pratique :

(a) On peut choisir Σ arbitraire et donc, par exemple, aussi large que possible (cette idée sera utilisée pour construire une section globale du champ, lorsque cela est possible). En effet, si Σ est une section quelconque (c'est-à-dire une sous-variété compacte connexe de dimension $(n-1)$ transverse à X en chacun de ses points), on peut trouver $\tau > 0$ et une boîte B difféomorphe à $\Sigma \times [-\tau, \tau]$ avec $\Phi|_{\Sigma \times \{0\}} \equiv Id$ (ce qui signifie que $\Phi(m, 0) = m$ pour tout $m \in \Sigma$). Le voisinage tubulaire $\Sigma \times [-\tau, \tau]$ est un épaississement de la section locale Σ. Évidemment, plus Σ est pris grand, plus τ doit être choisi petit. Le voisinage tubulaire prend alors la forme d'une *large plaque mince parallèle à la section* Σ (figure 4.7).

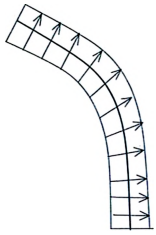


FIGURE 4.7. Voisinage en forme de plaque mince. En trait gras, la section Σ. Le champ X est signalé par des flèches.

(b) À l'inverse, on peut choisir un segment de trajectoire $\Gamma = \varphi([-\tau, \tau], p)$ avec $\tau > 0$ quelconque plongé dans U (si l'orbite de p est apériodique, on peut prendre τ arbitrairement grand). On peut alors trouver un voisinage tubulaire B tel que $\Gamma = \Psi(\{0\} \times [-\tau, \tau])$, où $\{0\} \in \mathbb{R}^{n-1}$. Pour construire un tel voisinage, on doit prendre une section locale Σ de diamètre suffisamment petit, d'autant plus petit que τ est choisi grand. Le voisinage tubulaire prend alors la forme d'un *long tube autour du segment de trajectoire* Γ (figure 4.8).

Dans ces deux cas, l'existence du voisinage tubulaire ne peut pas se déduire uniquement du théorème de l'inverse et nécessite un argument global. Par exemple, dans le cas (a), on peut appliquer le théorème de l'inverse en chacun des points de Σ, pour montrer que l'application Φ est localement inversible. Elle sera un difféomorphisme global à condition d'être *injective*. Cette injectivité sera

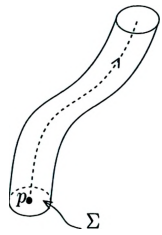


FIGURE 4.8. Voisinage en forme de tube. En trait pointillés, le segment de trajectoire Γ.

vérifiée pour $\tau > 0$ choisi assez petit. Elle se déduit du fait que Σ est une sous-variété plongée dans M.

4.4. Indice des points singuliers isolés

Définition 4.9. Soit X un champ de vecteurs sur une variété M de dimension n, et $p \in M$ un point singulier de X ($X(p) = 0$). Dans des coordonnées locales, $x \in \Omega$ ouvert de $\mathbb{R}^n$ contenant p, telles que p ait pour coordonnées 0, le champ s'écrit

$$X(x) = dX(0)(x) + o(||x||).$$

On appelle *spectre des valeurs propres de X en p* l'ensemble des valeurs propres de la matrice jacobienne $A = dX(0)$ comptées avec leur multiplicité.

Cette définition est justifiée par le fait que le spectre des valeurs propres de X en p ne dépend pas de la carte choisie, en effet :

Lemme 4.1. *Par un changement de coordonnées locales en p, la partie linéaire A est transformée par une conjugaison linéaire $A \mapsto PAP^{-1}$ où $P \in GL(n, \mathbb{R})$.*

Démonstration. Supposons que $y = G(x)$ soit un changement local de coordonnées avec, pour simplifier, $G(0) = 0$. Le champ image $Y = G_*(X)$ est donné par

$$Y(y) = dG(G^{-1})[X(G^{-1})].$$

En posant $dX(0) = A$ et $P = dG(0)$, on a :

$$Y(y) = (P + O(||y||))[A(P^{-1}(y) + o(||y||))] = PAP^{-1}(y) + o(||y||),$$

d'où il résulte que $dY(0) = PAP^{-1}$. ∎

Définition 4.10. On dit que le point singulier p du champ X est *non dégénéré* si aucune des valeurs propres de X en p n'est nulle. Cela est équivalent à dire que la partie linéaire de X en p, dans tout système de coordonnées locales, est une matrice inversible.

Dans la section 2.3, on a classifié les champs linéaires non dégénérés de $\mathbb{R}^2$ en fonction de la distribution des valeurs propres, en point de selle (si les valeurs propres sont réelles de signe contraire), nœud (si elles sont réelles de même signe) et centre ou foyer (si elles sont complexes conjuguées).

Lemme 4.2. *Si p est un point singulier non dégénéré de X, alors ce point est isolé parmi les points singuliers de X (autrement dit, il existe un voisinage W de p dans M, tel que p soit l'unique point singulier de X contenu dans W).*

Démonstration. Restreignons X sur une carte de coordonnées $x \in \Omega$ voisinage de p de coordonnées 0. Dire que p est non dégénéré signifie que $dX(0)$ est une matrice inversible. On peut donc appliquer le théorème de l'inverse à l'application $x \mapsto X(x)$ au voisinage de $p = 0$: il existe un voisinage W de p dans Ω tel que X soit une bijection sur son image. Comme $X(0) = 0$, on a $X(x) \neq 0$ pour tout $x \in W - \{0\}$. ∎

Remarque 4.9. L'exemple du champ $x\frac{\partial}{\partial x} + y^2\frac{\partial}{\partial y}$ montre qu'un point singulier, ici le point $(0,0)$, peut être isolé sans pour cela être non dégénéré.

Nous allons nous restreindre dorénavant aux champs X de vecteurs sur les surfaces, c'est-à-dire sur les variétés de dimension deux. Pour un point singulier *isolé* p de X, nous allons définir un invariant important appelé l'*indice de X en p*. Rappelons tout d'abord que si γ est un *lacet* ou *chemin fermé* de $\mathbb{R}^2 - \{0\}$ (donné par une application continue de $S^1 = \mathbb{R}/\mathbb{Z}$ dans $\mathbb{R}^2 - \{0\}$), on appelle *indice de γ par rapport à 0* : $\mathrm{Ind}_0(\gamma)$, le nombre de tours que γ fait autour de $\{0\}$, tours comptés dans le *sens direct* (le sens direct en mathématique, dit aussi sens trigonométrique, est le sens inverse à celui des aiguilles d'une montre).

De manière plus précise, on peut définir cet indice pour un lacet différentiable par morceaux par la formule

$$\mathrm{Ind}_0(\gamma) = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z} \in \mathbb{Z}, \tag{4.2}$$

où $\mathbb{R}^2$ est identifié avec $\mathbb{C}$ (le 0 dans la notation indique que l'indice est calculé par rapport à l'origine). On vérifie que l'indice du lacet $t \mapsto e^{2\pi i n t}$, qui tourne n fois autour de l'origine est bien égal à n, ce qui justifie cette définition analytique. La figure 4.9 montre un lacet d'indice 2 par rapport à l'origine.

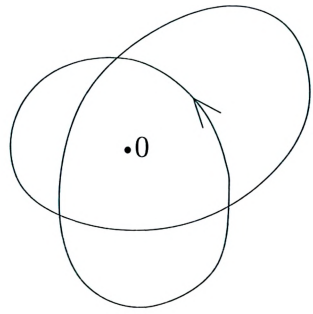


FIGURE 4.9. Lacet d'indice 2 en $0 \in \mathbb{R}^2$.

Remarque 4.10. Le choix de la paramétrisation du lacet γ par $t \mapsto e^{2\pi i t}$ vient du fait que l'on considère ce lacet comme application de $S^1 = \mathbb{R}/\mathbb{Z}$, et donc comme une application périodique *de période* 1. Évidemment, la formule intégrale (4.2), et donc la définition de l'indice, sont indépendantes du choix de la paramétrisation et en particulier de la période. On pourrait par exemple paramétrer γ par $t \mapsto e^{it}$, qui est une application de période 2π.

On peut étendre la définition de l'indice aux lacets continus, en remarquant que toutes les approximations différentiables d'un lacet continu γ, suffisamment $\mathcal{C}^0$-proches de γ ont le même indice, et en définissant l'indice de γ comme égal à l'indice de ces approximations différentiables.

Une propriété capitale de l'indice est d'être invariant par *homotopie*. Si M et N sont deux espaces topologiques et f, g des applications continues de M dans N, une homotopie entre f et g est une application continue $F : (x, u) \in M \times [0,1] \mapsto N$ telles que $f(x) \equiv F(x,0)$ et $g(x) \equiv F(x,1)$. Si, par exemple, M est compact et $N = \mathbb{R}^n$, on peut interpréter une homotopie comme un chemin continu entre f et g, dans l'espace de Banach $\mathcal{C}(M, \mathbb{R}^n)$ des applications continues, espace muni de la topologie de la convergence uniforme. On dira que deux lacets continus de $\mathbb{R}^2 - \{0\}$ sont homotopes, s'il sont définis par deux applications homotopes (à travers les applications continues de S^1 dans $\mathbb{R}^2 - \{0\}$).

Définition 4.11. Un lacet dans $\mathbb{R}^2 - \{0\}$, homotope à un point, est un *lacet constant*.

Il est clair, d'après (4.2), qu'un lacet constant est d'indice nul.

Proposition 4.6. *Si les deux lacets γ_0 et γ_1 sont homotopes dans $\mathbb{R}^2 - \{0\}$, alors* $\mathrm{Ind}_0(\gamma_0) = \mathrm{Ind}_0(\gamma_1)$.

Démonstration. Soit $F : S^1 \times [0,1] \mapsto \mathbb{R}^2 - \{0\}$ une homotopie continue entre γ_0 et γ_1. On peut approximer F par une application différentiable H, qui sera une homotopie entre deux lacets différentiables Γ_0 et Γ_1, C^0-proches de γ_0 et γ_1 respectivement. Par définition, on a $\mathrm{Ind}_0(\gamma_0) = \mathrm{Ind}_0(\Gamma_0)$ et $\mathrm{Ind}_0(\gamma_1) = \mathrm{Ind}_0(\Gamma_1)$. Notons Γ_u le lacet donné par $t \in S^1 \mapsto H(t,u)$. Par la formule (4.2), il est clair que l'application $u \mapsto \mathrm{Ind}_0(\Gamma_u)$ est continue. Comme cette fonction est à valeurs dans $\mathbb{Z}$, elle doit être constante. On a donc $\mathrm{Ind}_0(\Gamma_0) = \mathrm{Ind}_0(\Gamma_1)$, ce qui achève la démonstration. ∎

Remarque 4.11. Il est facile de montrer qu'inversement, si deux lacets de $\mathbb{R}^2 - \{0\}$ ont même indice par rapport à l'origine, alors ils sont homotopes dans $\mathbb{R}^2 - \{0\}$.

Nous allons maintenant définir l'indice d'un champ de vecteurs en un point singulier isolé d'une surface :

Définition 4.12. Soit un champ de vecteurs X sur une surface et p un point singulier. Soit Ω un ouvert de coordonnées d'une carte, dans laquelle le point $p = 0$ est isolé parmi les points singuliers. On supposera que Ω est difféomorphe à $\mathbb{R}^2$, ce qui n'est pas une restriction. Soit un $r > 0$ assez petit, pour que le cercle de rayon r et centré en $p = 0$ soit contenu dans Ω. Alors l'indice du champ X en p est égal, par définition, à l'indice du lacet $\Gamma_r : t \mapsto X(re^{2\pi it})$ de $\mathbb{R}^2 - \{0\}$ (autrement dit, du lacet obtenu en restreignant les valeurs du champ au cercle de rayon r, centré à l'origine). Cet indice appartient à $\mathbb{Z}$ et sera noté $\mathrm{Ind}_p X$.

L'indice ainsi défini est indépendant du choix de r. En effet si $r_1 < r_2$ sont deux rayons assez petits, l'application $F : (t,u) \mapsto F(t,u) = X([ur_1 + (1-u)r_2]e^{2\pi it})$ est une homotopie entre Γ_{r_1} et Γ_{r_2} dans $\mathbb{R}^2 - \{0\}$. L'indice est aussi indépendant de la carte choisie, à condition que l'application de carte préserve l'orientation. En effet, si $G : \Omega \mapsto \Omega' = \mathbb{R}^2$ est un changement de cartes, avec $G(0) = 0$, l'image du cercle de rayon r sera une courbe simple entourant l'origine et qui tend vers l'origine lorsque $r \to 0$. Or, dans le calcul de l'indice, on peut évidemment remplacer le cercle de rayon r par une telle courbe. Ainsi l'indice $\mathrm{Ind}_p X$ est intrinsèquement défini, ce qui justifie la notation : $\mathrm{Ind}_p X$. C'est de plus, par la définition 4.12, un *invariant local* : il ne dépend que des valeurs de X dans un voisinage arbitraire du point singulier p et, pour le calculer, on peut toujours supposer que le champ est défini sur $\mathbb{R}^2$. Remarquez que l'indice ne change pas si on remplace le champ par un champ équivalent, et même si on remplace X par $-X$. En effet, si Γ_r est un lacet de $\mathbb{R}^2 - \{0\}$ associé à X, le lacet $-\Gamma_r$ associé à $-X$ est obtenu à partir du lacet Γ_r par une rotation d'angle π. Or deux lacets de $\mathbb{R}^2 - \{0\}$, différents par une

rotation, sont homotopes comme il suit facilement de la proposition 4.6. Enfin, si le point singulier est non dégénéré, l'indice ne dépend que de sa partie linéaire :

Proposition 4.7. *Supposons que X soit défini sur $\mathbb{R}^2$ avec un point singulier non dégénéré à l'origine : $X(x) = X_0(x) + o(||x||)$, où X_0 est la partie linéaire de X. Alors* $\mathrm{Ind}_0 X_0 = \mathrm{Ind}_0 X$.

Démonstration. On peut écrire : $X(x) = X_0(x) + ||x||\varepsilon(x)$ où $\varepsilon : x \mapsto \varepsilon(x)$ est une application qui tend vers 0 pour $x \to 0$. D'autre part, puisque X_0 est donnée par une matrice inversible, il existe $a > 0$ tel que $||X_0(x)|| \geq a||x||$. Choisissons $r > 0$ assez petit pour que $||\varepsilon(x)|| < \frac{a}{2}$ si $||x|| \leq r$. Pour tout $u \in \mathbb{R}$, posons $X_u(x) = X_0 + u||x||\varepsilon(x)$. Alors $F(t,u) = X_u(re^{2\pi it})$ est une homotopie à valeurs dans $\mathbb{R}^2 - \{0\}$, et le résultat suit de la proposition 4.6 et de la définition 4.12 de l'indice d'un champ. ∎

Il est très facile de calculer les indices des champs linéaires non dégénérés. En fait, par des arguments élémentaires d'homotopie, il suffit de considérer les deux cas : $X_\pm = x\frac{\partial}{\partial x} \pm y\frac{\partial}{\partial y}$. Par exemple, si $X = \lambda x\frac{\partial}{\partial x} - \mu y\frac{\partial}{\partial y}$ avec $\lambda, \nu \in \mathbb{R}^+$, et un point de selle linéaire quelconque de X, la famille $X_t = (1-t)X + tX_-$, pour $0 \leq t \leq 1$, va donner l'homotopie souhaitée. Appliquons la définition 4.12, alors l'indice $\mathrm{Ind}_0 X_+$ est celui du lacet $t \mapsto e^{2\pi it}$, et l'indice de X_- est celui du lacet $t \mapsto e^{-2\pi it}$. On a donc : $\mathrm{Ind}_0 X_\pm = \pm 1$.

On nommera, dorénavant, un point singulier non dégénéré d'un champ (non linéaire) selon la nature de sa partie linéaire. Par exemple, un point de selle d'un champ est un point dont la partie linéaire est de type selle, c'est-à-dire avec des valeurs propres réelles de signe opposé. Le calcul de l'indice des champs linéaires $X_\pm$, combiné avec la proposition 4.7, permet de calculer l'indice pour un point singulier isolé général :

Proposition 4.8. *Soit p un point singulier non dégénéré d'un champ de vecteurs X sur une surface. Alors* $\mathrm{Ind}_p X = -1$ *si p est un point de selle, et il est égal à $+1$ dans tous les autres cas (la figure 4.10).*

Par exemple, considérons sur $\mathbb{R}^2$ le champ de vecteurs $X(x,y) = y\frac{\partial}{\partial x} - \sin x\frac{\partial}{\partial y}$, qui est le champ représentatif de l'équation du pendule dans l'espace de phase. Ce champ admet un point singulier en $(0,0)$ avec comme spectre de valeurs propres : $\{i, -i\}$. On a donc que $\mathrm{Ind}_{(0,0)} X = 1$. Le point $(0,\pi)$ est au contraire un point de selle de spectre $\{1,-1\}$. On a donc $\mathrm{Ind}_{(0,\pi)} X = -1$.

Remarque 4.12. L'indice en un point singulier isolé mais dégénéré peut prendre des valeurs autres que ± 1. Reprenons l'exemple de la remarque 4.9. Soit le champ $X(x,y) = x\frac{\partial}{\partial x} + y^2\frac{\partial}{\partial y}$ pour lequel l'origine de $\mathbb{R}^2 \approx \mathbb{C}$ est un point à la fois

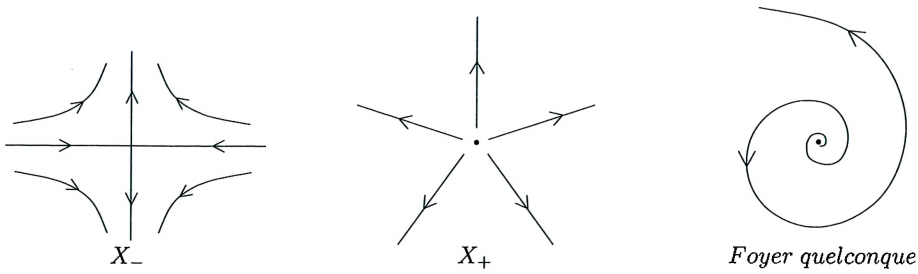


FIGURE 4.10. Lacets d'indice 2 par rapport à l'origine $\{0\}$.

singulier et dégénéré. *Alors l'indice de ce champ par rapport à l'origine est égal à* 0.

En effet, pour calculer l'indice de ce champ, il suffit de calculer l'indice par rapport à l'origine du lacet $\gamma : t \mapsto X(e^{2\pi it}) = (\cos 2\pi t, \sin^2 2\pi t)$ (figure 4.11).

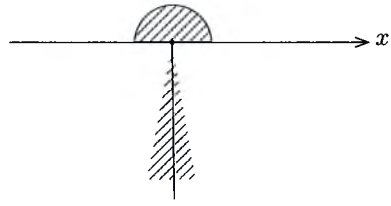


FIGURE 4.11. Le domaine d'existence du lacet γ est à l'extérieur du domaine hachuré. Évidemment, le cône de sommet $\{0\}$ (resp. la demi-circonférence) peut être choisi avec un angle au sommet (resp. un rayon) arbitrairement petit.

Or le lacet γ, ne coupant pas le demi-axe des y négatifs : $Oy_- = \{(0, y) \in \mathbb{R}^2 \mid y \leq 0\}$, est homotope dans $\mathbb{R}^2 - \{0\}$ à un lacet constant, et il est donc d'indice nul (c'est une conséquence directe, comme on l'a vu, de la définition 4.11). Évidemment, le lacet γ peut être enfermé dans une région plus petite, mais ce qui va importer, comme on va le voir, est qu'il soit contenu dans un ouvert homéomorphe à $\mathbb{R}^2$ de $\mathbb{R}^2 - \{0\}$.

L'affirmation qu'un lacet ne coupant pas Oy_- est d'indice nul peut s'établir de la façon suivante. Choisissons un homéomorphisme $\Psi : \mathbb{R}^2 - Oy_- \mapsto \mathbb{R}^2$. On peut prendre $\Psi = \Psi_1 \circ \Psi_2$ avec $\Psi_1(z) = \log(z)$ (on choisit la branche du log complexe telle que $\log(1) = 0$ et alors $\Psi_1(\mathbb{R}^2 - \{0\}) = B = \{z = x + iy \in \mathbb{C} | -\frac{\pi}{2} < y < \frac{3\pi}{2}\}$), et Ψ_2 comme étant un difféomorphisme de B sur $\mathbb{R}^2$ (Ψ_2 est très facile à exhiber) (la figure 4.12 montre la construction de Ψ).

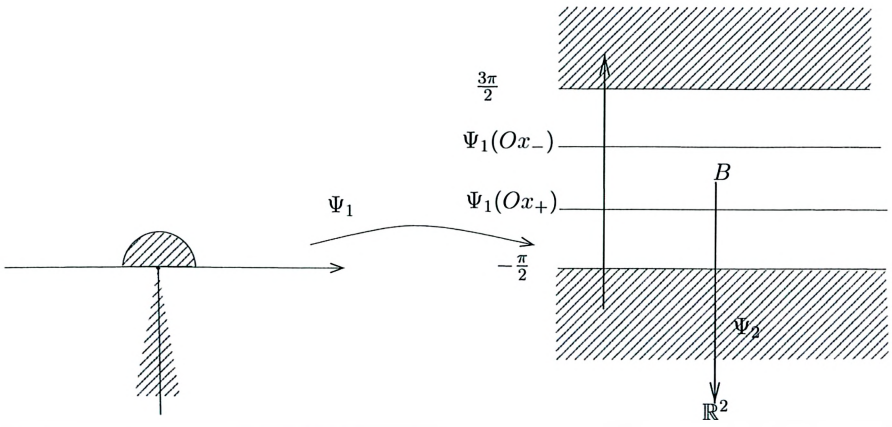


FIGURE 4.12. L'image de la partie hachurée de la figure gauche par $\Psi_1 = \log(z)$ est la bande B non hachurée de la figure droite.

Pour $r > 0$ quelconque, on définit la fonction φ sur $[0, +\infty)$ par : $\varphi(t) = e^{\frac{1}{(t-r)(t-2r)}}$ pour $t \in]r, 2r[$, et par $\varphi(t) = 0$ pour $t \in [0, r] \cup [2r, +\infty)$. Soit maintenant : $\widetilde{\varphi}(t) = \frac{\int_0^t \varphi(s)ds}{\int_0^{+\infty} \varphi(s)ds}$. Cette dernière fonction est croissante, $\mathcal{C}^\infty$ et à valeurs dans l'intervalle $[0, 1]$. Elle est égale à 0 sur l'intervalle $[0, r]$ et à 1 sur l'intervalle $[2r, +\infty)$. On lui associe l'application h_1, $\mathcal{C}^\infty$ de $\mathbb{R}^2$ dans $\mathbb{R}^2$, définie par : $h_1(x) = \widetilde{\varphi}(||x||)x$. Si B_ρ désigne le disque de rayon $\rho > 0$ centré à l'origine de $\mathbb{R}^2$, on remarque que l'application h_1 envoie le disque B_r sur l'origine et est égale à l'identité en dehors du disque B_{2r}. Soit maintenant l'homotopie $H : (x, u) \in \mathbb{R}^2 \times [0, 1] \mapsto H(x, u) = (1 - u)x + uh_1(x) \in \mathbb{R}^2$ reliant l'identité pour $u = 0$ à l'application h_1 pour $u = 1$. Comme $H(x, u) \equiv x$ pour $x \notin B_{2r}$, l'application

$$G : (x, u) \mapsto G(x, u) = \Psi^{-1} \circ H(\Psi(x), u)$$

se prolonge continûment en une application, encore notée G, de $\mathbb{R}^2 \times [0, 1]$, à valeurs dans $\mathbb{R}^2$, telle que $G(0, u) \equiv 0$. Supposons maintenant que r soit choisi assez grand pour que $\Psi \circ \gamma(S^1) \subset B_r$. Alors l'application $G : (\gamma, u) \mapsto G(\gamma, u)$ est une homotopie dans $\mathbb{R}^2 - \{0\}$ entre le lacet γ et le lacet constant γ_1 défini par $t \mapsto \gamma_1(t) \equiv \Psi^{-1}(0) \in \mathbb{R}^2 - \{0\}$. Comme un lacet constant dans $\mathbb{R}^2 - \{0\}$ est trivialement d'indice nul, γ est aussi d'indice nul.

Si X est un champ de vecteurs à singularités isolées sur une surface *compacte* M, le nombre des singularités est fini, car tout sous-ensemble discret d'un espace compact est fini.

On peut alors définir un indice global pour X en faisant la somme des indices aux points singuliers :

$$\mathrm{Ind}_M X = \sum \{\mathrm{Ind}_p X \mid X(p) = 0\}.$$

Dans cette définition, on suppose que $\mathrm{Ind}_M X = 0$ si le champ X n'a aucune singularité. Le principal résultat concernant l'indice est le théorème suivant, dont on pourra trouver une démonstration élémentaire dans [18].

Théorème 4.1 **(Théorème de Poincaré-Hopf).** *Soit M une surface compacte. Alors l'indice $\mathrm{Ind}_M X$ ne dépend pas du champ de vecteurs X à singularités isolées.*

Le nombre qui apparaît dans l'énoncé ne dépend que de la surface compacte. Il est clair d'autre part que la compacité est une hypothèse indispensable dans l'énoncé, car sur une surface non compacte, $\mathbb{R}^2$ par exemple, on peut trouver des champs dont tous les zéros sont non dégénérés mais en nombre infini.

On appelle ce nombre la *caractéristique d'Euler de M*. Cette caractéristique est notée $\mathcal{X}(M)$, et le théorème 4.1 se traduit par la *formule de Poincaré-Hopf* pour un champ de vecteurs X, à points singuliers isolés, sur une surface M :

$$\mathcal{X}(M) = \mathrm{Ind}_M X. \tag{4.3}$$

Il y a bien d'autres façons de calculer et d'interpréter $\mathcal{X}(M)$ et, en particulier, on peut montrer que la caractéristique d'Euler est un *invariant topologique,* c'est-à-dire invariant par homéomorphisme, ce qui ne suit pas de façon évidente de notre définition (on trouvera une définition directe de $\mathcal{X}(M)$, par exemple dans [9]).

Le théorème de Poincaré-Hopf nous donne un moyen, très pratique, de calculer la caractéristique d'Euler $\mathcal{X}(M)$, en considérant des champs de vecteurs particuliers sur la surface compacte M. Par exemple, on peut considérer le champ nord-sud sur la sphère qui a un point source au pôle Nord, un point puits au pôle Sud, et dont les autres trajectoires vont du pôle Nord au pôle Sud en suivant les méridiens (figure 4.13).

On trouve donc que $\mathcal{X}(S^2) = 2$. Comme le tore T^2 possède des champs de vecteurs sans points singuliers (par exemple les champs constants non nuls), on a $\mathcal{X}(T^2) = 0$. On peut aussi trouver des champs particuliers sur toutes les surfaces compactes, mais c'est un peu plus délicat.

Inversement, la formule de Poincaré-Hopf (4.3) donne une contrainte *a priori*, sur tout champ que l'on peut mettre sur une surface. Par exemple, si $\mathrm{Ind}_M X \neq 0$, alors tout champ de vecteurs sur M doit avoir des points singuliers. C'est bien évidemment le cas de la sphère S^2. Remarquez qu'un tel champ sur la sphère a au moins deux points singuliers, si les points singuliers sont non dégénérés.

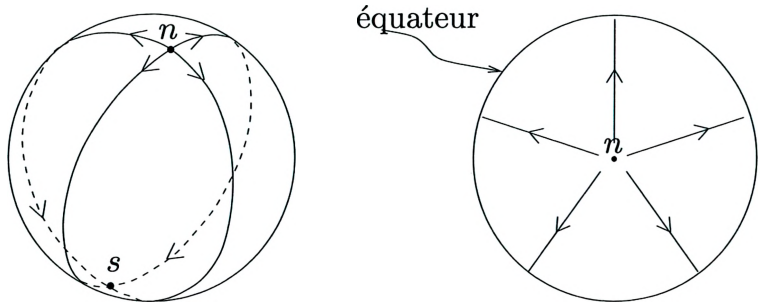
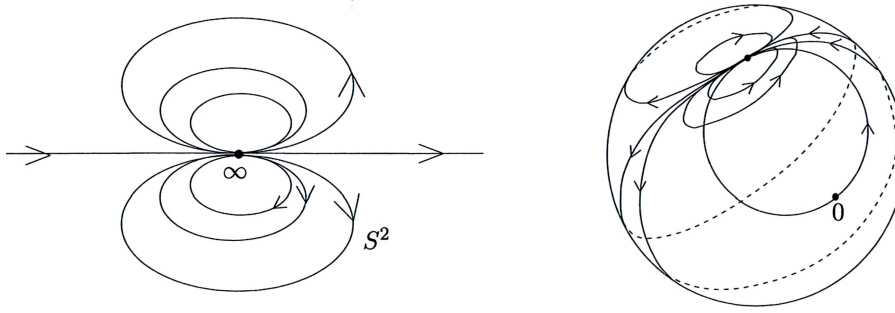


FIGURE 4.13. La sphère coupée selon l'équateur.

Par contre, il existe un champ sur S^2, avec un seul point singulier d'indice égal à 2, obtenu en compactifiant $\mathbb{R}^2$ par adjonction du point ∞ et en étendant à équivalence près le champ $\frac{\partial}{\partial x}$ de $\mathbb{R}^2$: l'unique point singulier sera le point ∞. Le champ $\frac{\partial}{\partial x}$ a pour équation différentielle dans la coordonnée complexe z

$$\dot{z} = 1. \tag{4.4}$$

On étudie le voisinage de l'infini en faisant $Z = \frac{1}{z}$ (l'infini correspond à $Z = 0$). Dans cette coordonnée Z, l'équation (4.4) s'écrit $\dot{Z} = \frac{-1}{z^2} = -Z^2$. Le point infini est l'unique point singulier du champ $\frac{\partial}{\partial x}$ étendu à S^2 (voir figure 4.14). L'indice du point singulier $Z = 0$ est l'indice de la courbe $t \mapsto -e^{4\pi i t} = e^{(4\pi i t + \pi i)}$ qui est d'indice 2 (voir formule (4.2)).



a. Dans la carte Z.

b. Sur la sphère.

FIGURE 4.14. Prolongement du champ $\frac{\partial}{\partial x}$.

On peut étendre le théorème de Poincaré-Hopf, ainsi que la définition de la caractérisque d'Euler, aux surfaces compactes à bord. Par exemple, on a $\mathcal{X}(D^2) = 1$, si D^2 est le disque (considérer le champ radial sur D^2). Ainsi, tout champ de vecteurs, transverse ou tangent au bord d'un disque, a au moins un

point singulier dans le disque. On se servira plus loin de ce résultat en conjonction avec le théorème de Poincaré-Bendixson 5.2. En général, sur une telle surface, la formule de Poincaré-Hopf s'étend aux champs X, à singularités isolées, qui sont soit transverses, soit tangents à chacune des composantes du bord. Pour calculer la caractéristique d'Euler d'une surface à bord, il suffit de prendre le *double de la surface,* en identifiant deux copies de la surface sur le bord et en prenant le champ X sur une copie et le champ $-X$ sur l'autre copie, en identifiant les deux bords par un difféomorphisme (voir la figure 4.15 qui illustre le double du disque avec pour champ X : le champ radial).

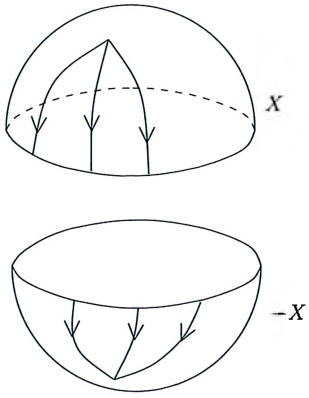


FIGURE 4.15

On peut effectuer le recollement de façon à obtenir un champ différentiable. Il en résulte que l'indice global de X sur une surface à bord ne dépend que de la surface. On appelle encore caractéristique d'Euler l'invariant ainsi obtenu. Il est évidemment égal *à la moitié de celui de la surface double.* Plus généralement, si deux surfaces compactes à bord M_1 et M_2 ont un bord difféomorphe (c'est-à-dire si leurs bords ont le même nombre de composantes connexes), on peut les recoller sur leur bord commun pour obtenir une surface sans bord M. La caractéristique d'Euler de cette surface M est alors la somme des caractéristiques de M_1 et de M_2. On considère les champs X_1 et X_2 sur M_1 et M_2, à singularités isolées, transverses au bord et se recollant comme un champ de M.

Ainsi, la caractéristique d'Euler d'un disque est égale à $+1$: pour le voir, il suffit de considérer le champ radial $x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$ qui est transverse au bord, avec un point singulier de type source à l'origine. Comme le double d'un disque est une sphère, on retrouve ainsi que $\mathcal{X}(S^2) = 2$. La caractéristique d'Euler d'un anneau est égale à 0, car le double de l'anneau est un tore qui est de caractéristique nulle. Celle du tore moins un disque ouvert est égale à -1 (on obtient

le tore, de caractéristique 0, en recollant un disque, et on a vu qu'un disque est de caractéristique égale à 1).

La caractéristique d'Euler d'une surface est en relation avec son *genre*, le chapitre III-4 fera souvent référence à cette notion. Nous allons en dire quelques mots ici, en nous restreignant aux surfaces compactes, orientables, connexes et sans bord.

Le genre d'une surface se définit comme suit :

Définition 4.13. Le *genre* g d'une surface connexe est le nombre *maximum* de courbes fermées simples, sans points communs, que l'on peut tracer sur cette surface sans la déconnecter (c'est-à-dire que le complémentaire de ces courbes reste connexe).

Concrètement, si l'on considère que la surface est en papier, le genre est le nombre maximal de découpages (de « coups de ciseaux ») que l'on peut effectuer, *sans que* la surface soit découpée en plusieurs morceaux. Il est donc clair que la sphère S^2 est de genre 0 et que le tore T^2 est de genre 1 (avec un coup de ciseau, on ne découpe pas le tore (en deux) ; avec deux, on le peut). Cette notion est topologique : deux surfaces n'ayant pas le même genre ne sont pas homéomorphes, ceci permet de distinguer la sphère du tore. Plus généralement pour le « tore à g trous » T^2_g, le genre est précisément égal au nombre des trous. Par exemple, le tore usuel T^2 est le tore à 1 trou T^2_1, et, pour $g = 2$, on obtient une surface en forme de « bretzel ». Il est d'autre part facile de trouver sur T^2_g un champ de vecteurs avec exactement $2g - 2$ points de selle.

Il en résulte la relation importante reliant la caractéristique d'Euler $\mathcal{X}(M)$ au genre g de la surface (voir [17]) :

$$\mathcal{X}(T^2_g) = 2 - 2g. \tag{4.5}$$

Toute surface compacte, orientable, connexe et sans bord, est homéomorphe à l'une des variétés T^2_g (on considère que S^2 est le tore à 0 trou T^2_0) et on peut montrer que ces surfaces sont deux à deux non homéomorphes [9]. Il résulte donc de la formule (4.5) que toute surface compacte, orientable, connexe et sans bord, est caractérisée, à homéomorphisme près, par son genre ou bien sa caractéristique d'Euler. Par exemple, pour la sphère S^2, vue comme T^2_0, le genre est égal à 0 et $\mathcal{X}(S^2) = 2$; pour le tore $T^2 = T^2_1$, le genre est égal à 1 et $\mathcal{X}(T^2) = 0$.

5

RÉCURRENCE

Soit X un champ de vecteurs sur une variété M. On suppose que X est un champ de vecteurs à flot complet, c'est-à-dire à flot défini sur $\mathbb{R} \times M$. On veut s'intéresser au comportement des trajectoires pour $t \to +\infty$ (ou $-\infty$).

Définition 5.1 **(Ensembles ω-limite et α-limite d'un point).** Soit $x \in M$. On note

$$\omega_X(x) = \{y \in M \mid \exists\ (t_i)_i \to +\infty,\ \ \varphi(t_i, x) \to y\}$$

et

$$\alpha_X(x) = \{y \in M \mid \exists\ (t_i)_i \to -\infty,\ \ \varphi(t_i, x) \to y\}.$$

Autrement dit, l'ensemble $\alpha(x)$ est l'ensemble $\omega(x)$ pour le champ $-X$. S'il n'y a pas d'ambiguïté, on écrira simplement $\omega(x), \alpha(x)$.

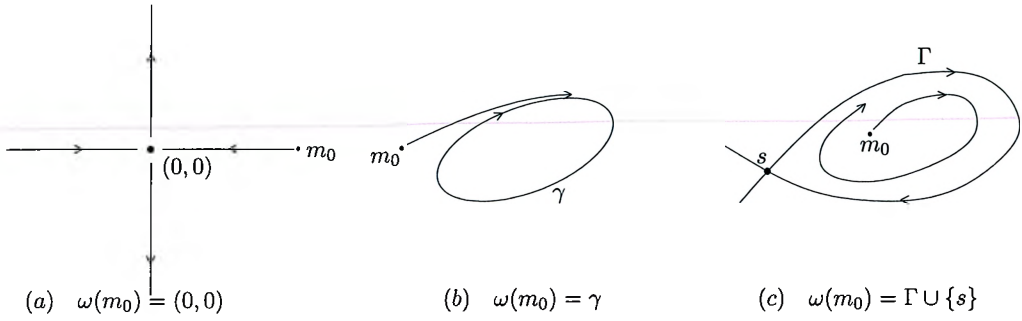
L'ensemble ω-limite du point x est l'ensemble des limites possibles de points de la trajectoire par x, pour des t tendant vers $+\infty$ (même définition pour l'ensemble α-limite du point x pour des t tendant vers $-\infty$).

Exemples d'ensembles limites.

(a) Considérons le champ de vecteurs linéaire $X = -x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$. Si $m_0 = (x_0, 0)$, on a $w(m_0) = \{(0,0)\}$; si $m_0 = (0, y_0)$, on a $\alpha(m_0) = \{(0,0)\}$. Tous les autres ensembles limites sont vides.

(b) L'exemple précédent montre qu'un ensemble limite $\omega(m)$ peut se réduire à un seul point singulier m_0 du champ, et alors on a nécessairement que $\varphi(t, m) \to m_0$ pour $t \to +\infty$. Inversement, si $\varphi(t, m) \to m_0$ pour $t \to +\infty$, alors $\omega(m) = \{m_0\}$ et $X(m_0) = 0$ (si $X(m_0) \neq 0$, l'orbite de m_0 ne serait pas réduite au point

m_0 et serait contenue dans $\omega(m_0)$ comme on le verra ci-après, ce qui contredirait l'hypothèse que $\varphi(t,m) \to m_0$). En général, $m_0 \in \omega(x) \not\Rightarrow \varphi(t,m) \to m_0$ car $\omega(m)$ n'est pas réduit au seul point m_0. Dans la figure 5.1, on montre quelques exemples d'ensembles ω-limites possibles : un point singulier comme dans l'exemple ci-dessus, une orbite périodique, une *connection de selle* formée de 2 orbites (un point singulier s de type selle et une orbite, dont les 2 ensembles limites sont ce point s).



(a) $\omega(m_0) = (0,0)$ (b) $\omega(m_0) = \gamma$ (c) $\omega(m_0) = \Gamma \cup \{s\}$

FIGURE 5.1. Exemples d'ensembles limites.

5.1. Propriétés des ensembles limites

Comme on l'a signalé plus haut : $\omega_{-X}(x) = \alpha_X(x)$; il suffira donc d'étudier les ensembles $\omega(x)$ associés à X. On choisit une distance $\mathrm{dist}(x,y)$ définissant la topologie de M, par exemple la distance associée à une métrique riemannienne (voir la définition I-2.15).

Proposition 5.1. *Si $y \in \gamma_x$, alors $\omega(x) = \omega(y)$. Autrement dit, l'ensemble $\omega(x)$ ne dépend que de l'orbite γ de x. Si $\gamma_x = \gamma$, on pourra donc noter par $\omega(\gamma)$ l'ensemble $\omega(x)$.*

Démonstration. Comme $y \in \gamma_x \Leftrightarrow x \in \gamma_y$, il suffit de montrer que $x \in \gamma_y \Rightarrow \omega(x) \subset \omega(y)$. Supposons donc que $x = \varphi(\tau, y)$ pour un certain $\tau \in \mathbb{R}$ (ce qui traduit que $x \in \gamma_y$), et considérons $z \in \omega(x)$, c'est-à-dire : $z = \lim_i \varphi(t_i, x)$ pour une suite $(t_i)_i \to +\infty$. On a $\varphi(t_i, x) = \varphi(t_i, \varphi(\tau, y)) = \varphi(t_i + \tau, y)$. Comme la suite $(t_i + \tau)_i \to +\infty$, on obtient que : $z = \lim_i \varphi(t_i + \tau, y) \in \omega(y)$. Ce qui conclut. ∎

Proposition 5.2. *L'ensemble $\omega(x)$ est fermé dans U.*

Démonstration. Soit $(y_j) \to y$ avec $y_j \in \omega(x)$. Considérons un $n \in \mathbb{N}$ quelconque. On peut trouver $j(n)$ tel que $\mathrm{dist}(y_{j(n)}, y) \leq \frac{1}{2n}$. Pour tout j, par définition de $\omega(x)$, il existe $(t_{ij})_i \to +\infty$ telle que $\varphi(t_{ij}, x) \to y_j$ pour $i \to +\infty$. En particulier, il existe $i(n) \in \mathbb{N}$ tel que $\mathrm{dist}(\varphi(t_{i(n)j(n)}, x), y_j(n)) \leq \frac{1}{2n}$. Par inégalité triangulaire, nous obtenons que, pour tout $n \in \mathbb{N}$, on a $\mathrm{dist}(\varphi(t_{i(n)j(n)}, x), y) \leq \frac{1}{n}$ et donc que $y \in \omega(x)$. ∎

Remarque 5.1. En fait on a la définition équivalente suivante, pour tout $T_0 \in \mathbb{R}$:

$$\omega(x) = \bigcap_{T \geq T_0} \overline{\{\varphi(t,x) \mid t \geq T\}}. \tag{5.1}$$

Cette définition montre directement que $\omega(x)$ est fermé (comme intersection de fermés).

Définition 5.2. On dit que $K \subset M$ est *invariant par le flot* $\varphi(t,x)$ si $\varphi_t(K) = \varphi(t,K) \subset K$ pour tout $t \in \mathbb{R}$. En appliquant le difféomorphisme φ_{-t} à cette inclusion, on obtient que $K \subset \varphi_{-t}(K)$ pour tout t. Ainsi, K est invariant par le flot si et seulement si $\varphi_t(K) = K$ pour tout $t \in \mathbb{R}$. Remarquez que K est invariant par le flot du champ X si et seulement s'il est une réunion d'orbites de X.

De la même façon, on peut définir l'invariance pour les temps positifs ou pour les temps négatifs.

Proposition 5.3. *L'ensemble $\omega_X(x)$ est invariant par le flot du champ X.*

Démonstration. Écrivons $\omega_X(x) = \omega(x)$. Si $y \in \omega(x)$, il existe une suite $(t_i)_i \to \infty$ telle que $\varphi(t_i, x) \to y$; mais alors, pour tout t fixé, par continuité de φ et la formule de translation, on a : $\varphi(t,y) = \lim_i \varphi(t, \varphi(t_i, x)) = \lim_i \varphi(t + t_i, x) \in \omega(x)$ lorsque $(t_i)_i \to +\infty$. On a donc : $\varphi(t,y) \in \omega(x)$ pour tout t, et ceci pour tout $y \in \omega(x)$. Ce qui démontre que $\varphi(t, \omega(x)) \subset \omega(x)$. ∎

Ainsi, pour tout $t \in \mathbb{R}$, on a $\omega(x) = \omega(\varphi(t,x)) = \varphi(t, \omega(x))$ (transporté de $\omega(x)$ par φ_t).

Proposition 5.4. *Si $y \in \omega(x)$, alors $\omega(y) \subset \omega(x)$.*

Démonstration. Cette propriété suit immédiatement du fait que $\omega(x)$ est fermé et invariant par le flot. ∎

La figure 5.1-(c) montre un exemple d'ensemble limite $\omega(m)$ où $s \in \omega(m)$ et $\omega(s) = \{s\} \neq \omega(m)$.

Remarque 5.2. Il peut se faire que $\omega(x)$ soit vide (cf. $X = \frac{\partial}{\partial x}$ sur $\mathbb{R}$). Par contre, si la demi-orbite positive $\gamma_x^+ = \{\varphi(t,x) \mid t \geq 0\}$ a une fermeture *dans* M qui est compacte, alors nécessairement $\omega(x) \neq \emptyset$ est compact. En effet, dans ce cas, si $(t_i)_i \to +\infty$, on peut extraire une sous-suite $(t_{i(n)})_n$ telle que $\varphi(t_{i(n)}, x)$ converge dans M. D'autre part $\omega(x)$ est compact comme partie fermée du compact $\overline{\gamma_x^+}$. Dans le cas $M = \mathbb{R}^n$, il suffit que γ_x^+ soit bornée.

Définition 5.3. Si $K \subset M$ est un fermé, on définit la distance de $x \in M$ à K par $\text{dist}(x, K) = \text{Inf}_{y \in K} \text{dist}(x, y)$.

Comme K est fermé, $\text{dist}(x, K) = 0$ si et seulement si $x \in K$.

Définition 5.4. Soit K un fermé de M. On dit alors que $\varphi(t,x) \to K$ pour $t \to +\infty$ si :

$$\text{dist}(\varphi(t,x), K) \to 0 \text{ pour } t \to +\infty.$$

On peut encore définir cette convergence de la façon suivante : pour tout $\varepsilon > 0$, soit $V_\epsilon(K) = \{y \in M \mid \text{dist}(y, K) < \epsilon\}$ le ε-voisinage de K, alors :

$$\varphi(t,x) \to K \iff \{\forall \epsilon > 0,\ \exists T_\varepsilon \text{ tel que : } t \geq T_\varepsilon \Rightarrow \varphi(t,x) \in V_\epsilon(K)\}.$$

Remarque 5.3. Si $K_1 \subset K_2$ sont deux fermés de M et si $\varphi(t,x) \to K_1$, alors on a aussi que $\varphi(t,x) \to K_2$. Le plus grand fermé de M avec cette propriété est évidemment la variété M elle-même ! L'intérêt de la proposition suivante est de montrer, au contraire, que $\omega(x)$ est le plus petit fermé avec cette propriété.

Proposition 5.5. *Soit $x \in M$ et supposons que $\overline{\gamma_x^+}$ soit compact dans M (rappelons que, sous cette condition, $\omega(x)$ est compact non vide). Alors $\varphi(t,x) \to \omega(x)$ pour $t \to +\infty$ et, de plus, si K est un fermé tel que $\varphi(t,x) \to K$ pour $t \to +\infty$, alors $\omega(x) \subset K$. Autrement dit, $\omega(x)$ est le plus petit fermé K qui soit limite de $\varphi(t,x)$ pour $t \to +\infty$.*

Démonstration. La seconde affirmation est triviale : si K est un fermé tel que $\varphi(t,x) \to K$ pour $t \to +\infty$, alors K doit contenir toutes les limites des suites $(\varphi(t_i, x))_i$, avec $(t_i)_i \to +\infty$, c'est-à-dire que $\omega(x) \subset K$.

Pour démontrer que $\varphi(t,x) \to \omega(x)$, faisons un raisonnement par l'absurde. Supposons que $\varphi(t,x) \not\to \omega(x)$. Alors il existe $\epsilon > 0$ et une suite $(t_i)_i \to +\infty$

tels que $\varphi(t_i, x) \notin V_\epsilon(\omega(x))$. Mais comme $\overline{\gamma_x^+} \bigcap (M - V_\epsilon(\omega(x)))$ est compact, on peut extraire une sous-suite convergente de cette suite telle que $\varphi(t_{ij}, x) \to y \notin V_\epsilon(\omega(x))$. Comme évidemment $\omega(x) \subset V_\epsilon(\omega(x))$, on a donc $y \notin \omega(x)$; or l'existence de la sous-suite implique que $y \in \omega(x)$, nous avons une contradiction. Ce qui démontre la proposition. ∎

On voit donc que l'ensemble $\omega(x)$, lorsqu'il est compact non vide, est le plus petit ensemble qui attire la trajectoire.

Proposition 5.6. *Soit $x \in M$ et supposons que $\overline{\gamma_x^+}$ soit compact dans M. Alors $\omega(x)$ est connexe.*

Démonstration. Faisons un raisonnement par l'absurde. Nous savons que $\omega(x)$ est compact non vide. Si $\omega(x)$ est non connexe, on peut l'écrire comme union de deux compacts K et L non vides disjoints : $\omega(x) = K \cup L$ avec $K \cap L = \emptyset$. Si $\varepsilon > 0$ est assez petit, $V_\varepsilon(K) \cap V_\varepsilon(L) = \emptyset$ et $\overline{V_\varepsilon(K) \cup V_\varepsilon(L)} \subset M$ est compact. Choisissons un tel ε. Soit $B = \overline{V_\varepsilon(K) \cup V_\varepsilon(L) \cup \gamma_x^+} \subset M$. Comme $\varphi(t, x) \to \omega(x) \subset B$, il existe un T tel que $\varphi(t, x) \in B$ si $t > T$. Comme K et L sont contenus dans $\omega(x)$, on peut trouver deux suites $(t_i)_i, (t_i')_i \to +\infty$ avec $t_i, t_i' > T$, telles que $\varphi(t_i, x) \in V_\varepsilon(K)$ et $\varphi(t_i', x) \in V_\varepsilon(L)$ pour tout i. On suppose que les deux suites sont choisies telles que $t_i < t_i' < t_{i+1}$ pour tout i. Par raison de connexité du segment d'orbite $\varphi([t_i, t_i'], x)$, pour tout i il existe un temps τ_i avec $t_i < \tau_i < t_i'$, tel que $\varphi(\tau_i, x)$ appartienne au compact $C = B - V_\varepsilon(K) \cup V_\varepsilon(L)$ (si un tel τ_i n'existait pas, on ne pourrait pas avoir $V_\varepsilon(K) \cap V_\varepsilon(L) = \emptyset$). En extrayant une sous-suite de ces temps τ_i, on obtient une suite convergente dans C. Il en résulte que $\omega(x) \cap C \neq \emptyset$, ce qui est une contradiction, puisque $\omega(x) \subset V_\varepsilon(K) \cup V_\varepsilon(L) \subset B$, avec inclusion stricte de $V_\varepsilon(K) \cup V_\varepsilon(L)$ dans B, entraîne que $\omega(x) \cap C = \emptyset$. ∎

Remarque 5.4. La nécessité de faire l'hypothèse que M soit compacte dans la proposition précédente n'est pas très intuitive. La figure 5.2 illustre cette nécessité. Dans cette figure, la trajectoire par x tourne dans le plan en formant une spirale, pour t croissant indéfiniment vers l'infini, en restant confinée entre les deux droites horizontales parallèles D et D'. Il est clair que $\overline{\gamma_x^+}$ n'est pas bornée donc non compacte. L'ensemble ω-limite est la réunion de ces deux droites D et D', réunion qui est un ensemble non connexe.

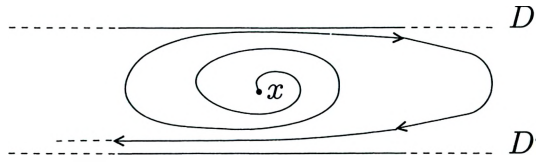


FIGURE 5.2. Si $\overline{\gamma_x^+}$ n'est pas compact, $\omega(x)$ peut être non connexe.

Définition 5.5. On pose :

$$L_+(X) = \overline{\bigcup_x \omega(x)}\,,\ L_-(X) = \overline{\bigcup_x \alpha(x)}\,,\ L(X) = L_+(X) \cup L_-(X).$$

L'ensemble $L(X)$ est appelé *ensemble limite* du champ X. Il est invariant par le flot de X, et c'est le plus petit fermé vers lequel tendent toutes les trajectoires, aussi bien pour les temps positifs que pour les temps négatifs.

5.2. Orbites récurrentes

Soient $x \in M$ et l'orbite $\varphi(t,x)$ du champ X.

Définition 5.6. On dit que x est *positivement récurrent* si $x \in \omega(x)$. Autrement dit, s'il existe une suite $(t_i)_i \to +\infty$ telle que $\varphi(t_i,x) \to x$. Cela signifie aussi que, quel que soit le voisinage V de x, quel que soit T, il existe $t \geq T$ tel que $\varphi(t,x) \in V$. Si un point est positivement récurrent, il en est de même pour chaque point de son orbite. On dira que l'orbite est positivement récurrente. La trajectoire d'un point positivement récurrent x revient indéfiniment dans V (pour des temps allant vers $+\infty$) : *c'est le phénomène de récurrence.*

On définit de la même façon les points et orbites *négativement récurrents.*

Remarque 5.5. Usuellement on dit *récurrent* pour positivement récurrent. De même, pour dire que le point x est négativement récurrent pour le champ X, on dit souvent que x est positivement récurrent pour le champ $-X$.

Remarque 5.6. Tous les exemples de points récurrents proposés dans la suite (récurrences triviales, flot irrationnel) sont récurrents à la fois positivement et négativement.

Exemples triviaux de récurrence

1. Si $X(x) = 0$ (i.e. x est un point critique), alors $\omega(x) = \alpha(x) = \{x\}$ (car on a évidemment : $\varphi(t, x) = x$, pour tout t).

2. Il est aussi immédiat de voir que si l'orbite γ_x est périodique, alors $\omega(x) = \alpha(x) = \gamma_x$. D'une part comme γ_x est compact, on a $\omega(x), \alpha(x) \subseteq \gamma_x$. Inversement, si $y \in \gamma_x$, on peut écrire que $y = \varphi(\tau, x)$ pour un certain $\tau \in \mathbb{R}$, et si T est la période de γ_x, il s'ensuit que $y = \varphi(t + iT, x)$ pour tout $i \in \mathbb{Z}$. On a donc que $y \in \omega(x)$ et $y \in \alpha(x)$ (prendre $t_i = \tau + iT$, pour $i \to \pm\infty$).

Définition 5.7. Les points singuliers et les orbites périodiques sont appelés *orbites récurrentes triviales.*

Nous allons maintenant détailler un exemple d'orbite récurrente non triviale.

Flot irrationnel sur le tore

Nous allons donner un exemple d'*orbite apériodique* dans $\mathbb{R}^3$ ou $\mathbb{R}^4$ dont la fermeture remplit une surface, c'est-à-dire une sous-variété de dimension 2. Cette sous-variété sera difféomorphe au *tore* T^2 obtenu comme quotient de $\mathbb{R}^2$ par le groupe $\mathbb{Z}^2$ des translations entières (voir paragraphe I-2.1.2).

(1) Construction du flot sur le tore

Rappelons que pour obtenir l'espace quotient $T^2 = \mathbb{R}^2/\mathbb{Z}^2$, on peut se limiter à un domaine fondamental de l'action du groupe $\mathbb{Z}^2$ sur $\mathbb{R}^2$, c'est-à-dire un compact de $\mathbb{R}^2$ qui est traversé par chaque orbite de l'action (les orbites sont les classes d'équivalence de la relation d'équivalence ρ de passage au quotient). On peut prendre comme domaine fondamental le carré $[0, 1] \times [0, 1]$. Le quotient C/ρ est obtenu par les identifications sur le bord du carré par les formules : $(x, 0) \sim (x, 1)$ et $(0, y) \sim (1, y)$. Comme C est compact, on obtient un homéomorphisme de C/ρ avec T^2. Cet espace quotient C/ρ est donc une représentation topologique de T^2 (figure 5.3).

Pour commencer, nous allons construire un champ de vecteurs sur T^2. Rappelons qu'un champ X sur T^2 est déterminé par un champ $\widehat{X}$ sur $\mathbb{R}^2$ invariant par les translations entières de $\mathbb{Z}^2$, c'est-à-dire les translations de type $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+m \\ y+p \end{pmatrix}$ avec $(m, p) \in \mathbb{Z}^2$. Les orbites de X sont alors les orbites de $\widehat{X}$ considérées dans l'espace T^2.

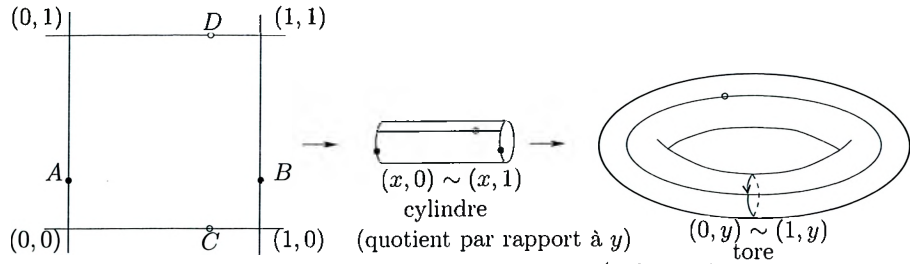


FIGURE 5.3

Nous allons considérer l'exemple particulier d'un champ constant (champ de Kronecker) : $\widehat{X}_\alpha = \frac{\partial}{\partial x} + \alpha \frac{\partial}{\partial y}$ dont l'équation différentielle est :

$$\begin{cases} \dot{x} = 1 \\ \dot{y} = \alpha \end{cases}.$$

Nous avons $\dot{y} - \alpha\dot{x} \equiv 0$, ce qui signifie que la fonction $y - \alpha x$ est une intégrale première de $\widehat{X}_\alpha$. Les orbites sont donc les droites de pente α, d'équation $\{y = \alpha x + \beta\}$, pour β quelconque dans $\mathbb{R}$ (figure 5.4). Selon la valeur de α, on obtient deux types d'orbites.

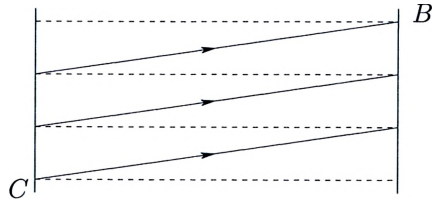


FIGURE 5.4

i) *Les orbites sont périodiques si* $\alpha \in \mathbb{Q}$ (on pose $\alpha = \frac{p}{q}$ avec $p \in \mathbb{Z}$, $q \in \mathbb{N}$ et p, q premiers entre eux). La période est égale à q : chaque orbite « se ferme » en faisant q tours dans le sens des x et p tours dans le sens des y, avec une période égale à q. En effet si $x \mapsto x + q$, on a $qy = p(x + q) + q\beta$ soit

$$y = \frac{p}{q}x + \beta + p \ ;$$

comme p est entier, le système « se ferme » par une translation de q en x (ce qui revient dans l'identification précédente par le tore à faire q tours dans le sens des

x) et une translation de p en y (à savoir p tours dans le sens des y). Par ailleurs, l'intégration du champ $\widehat{X}_{\frac{p}{q}}$ donne

$$\varphi(t,m) : \begin{cases} x(t,m) = t + x_0 \\ y(t,m) = \frac{p}{q}t + y_0 \end{cases}$$

si $m = (x_0, y_0) \in \mathbb{R}^2$. On vérifie facilement que si $t \in]0, q[$, on a :

$$\varphi(t,m) - m \notin \mathbb{Z}^2,$$

ce qui implique que la (plus petite) période est égale à q. La figure 5.5 illustre le cas $q = 2$ et $p = 1$ (on tourne deux fois dans le sens des x lorsqu'on tourne une fois dans le sens des y).

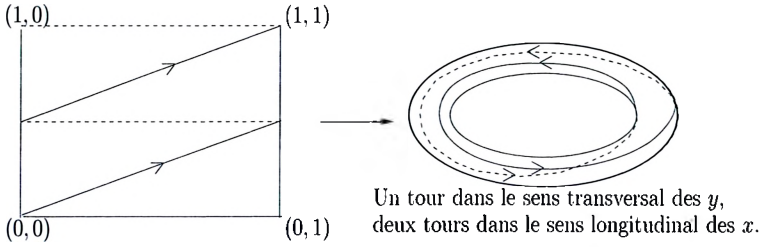


FIGURE 5.5. Cas $q = 2, p = 1$.

Pour p et q quelconques, on a intérêt à visualiser les orbites tout d'abord dans le rectangle $T_{p,q} = [0,p] \times [0,q] \subset \mathbb{R}^2$. L'orbite de $(0,0)$ par exemple est la droite joignant $(0,0)$ à (p,q) (c'est-à-dire $y = \frac{q}{p}x$: on tourne p fois dans le sens des x lorsqu'on tourne q fois dans le sens des y (figure 5.6)). Ramenée par translations entières dans le carré $C = T_{1,1} = [0,1] \times [0,1]$ (le domaine fondamental du quotient $T^2 = \mathbb{R}^2/Z^2$), cette orbite est représentée, si $p = 3$ et $q = 2$, par 3 segments de droites parallèles. Projetée dans le tore T^2, cette orbite coupe 3 fois le cercle $\Gamma_1 = \mathbb{R}/\mathbb{Z} \times \{0\}$ et 2 fois le cercle $\Gamma_2 = \{0\} \times \mathbb{R}/\mathbb{Z}$.

ii) *Les orbites sont apériodiques si* $\alpha \notin \mathbb{Q}$. Dans ce cas, les orbites ne se ferment pas. On dit que l'on a affaire à un *flot irrationnel* sur le tore. L'ensemble des points de $\gamma_0 \cap \Gamma_1$ (où Γ_1 est le cercle du tore associé à $x = 0$ et γ_0 est l'orbite par le point $0 = (0,0)$) est dense dans Γ_1. En effet cet ensemble est la projection, dans le cercle $\mathbb{R}/\mathbb{Z}$, de l'ensemble des points $G = \{m + n\alpha | (m,n) \in \mathbb{Z}^2\}$ qui est dense dans $\mathbb{R}$, si et seulement si α est un nombre irrationnel. En effet G est un sous-groupe additif de $\mathbb{R}$, non réduit à $\{0\}$. Si sa fermeture $\overline{G}$ était un groupe discret, il serait de la forme $\mathbb{Z}\{T\}$ pour un $T \neq 0$. On aurait alors $G = \overline{G}$ et il existerait $(m,n) \in \mathbb{Z}^2$ tel que $m + n\alpha = 0$. Ceci signifierait que α serait rationnel, ce qui n'est pas le

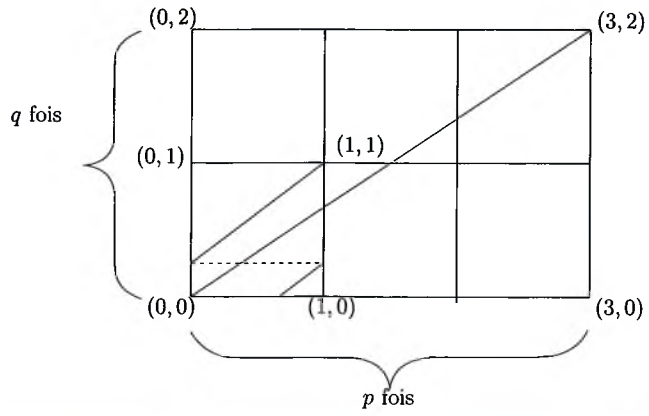


FIGURE 5.6. Cas $p = 3$, $q = 2$.

cas. D'après la proposition 4.2, si G n'est pas discret, il doit être dense dans $\mathbb{R}$. Il en résulte aisément que l'orbite γ_0 elle-même est dense dans le tore : $\overline{\gamma}_0 = T^2$. Le même résultat est vrai quelle que soit l'orbite. Chaque orbite de points du tore est donc dense dans T^2.

(2) On peut plonger T^2 dans $\mathbb{R}^3$

Pour cela, on fait tourner un cercle du plan (x, y) autour de l'axe $0z$ (figure 5.7). On peut ensuite prolonger X_α, α irrationnel, à $\mathbb{R}^3$ (nous ne le ferons pas ici). On obtient alors un exemple de champ X dans $\mathbb{R}^3$ ayant une orbite apériodique γ, telle que $\overline{\gamma}$ est une surface plongée dans $\mathbb{R}^3$, difféomorphe au tore T^2.

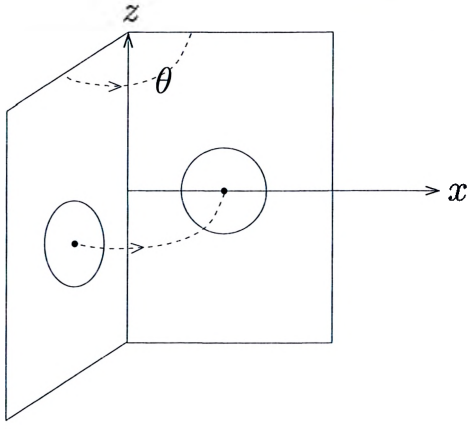


FIGURE 5.7

Dans $\mathbb{R}^3$, la topologie des orbites apériodiques peut donc être très compliquée. Remarquez que dans l'exemple décrit, chaque boule ouverte de $\mathbb{R}^3$ coupant le tore plongé, intersecte chaque orbite en une infinité d'intervalles ouverts : il en résulte que la topologie induite par $\mathbb{R}^3$ sur chaque orbite est strictement moins fine que la topologie de $\mathbb{R}$.

(3) Plongement du flot irrationel dans $\mathbb{R}^4$

Considérons le système mécanique formé de deux oscillateurs linéaires non couplés. On peut se figurer, par exemple, le système formé par deux ressorts élastiques sans aucune interaction. L'équation différentielle d'un tel système est formée par la juxtaposition de deux équations différentielles linéaires du deuxième ordre :

$$\begin{cases} \ddot{x}_1 + \omega_1^2 x_1 = 0 \\ \ddot{x}_2 + \omega_2^2 x_2 = 0 \end{cases}.$$

Les nombres $\frac{\omega_1}{2\pi}$ et $\frac{\omega_2}{2\pi}$ sont les fréquences des deux équations. On rappelle que l'on peut passer dans l'espace de phase $\mathbb{R}^4$ de coordonnées (x_1, y_1, x_2, y_2) pour obtenir l'équation différentielle d'un champ de vecteurs linéaire :

$$\begin{cases} \dot{x}_1 = y_1 \\ \dot{y}_1 = -\omega_1^2 x_1 \\ \dot{x}_2 = y_2 \\ \dot{y}_2 = -\omega_2^2 x_2 \end{cases}. \tag{5.2}$$

Ce champ admet deux intégrales premières $E_1 = \frac{1}{2}(\omega_1^2 x_1^2 + y_1^2)$ et $E_2 = \frac{1}{2}(\omega_2^2 x_2^2 + y_2^2)$. Si $m = (x_1^0, y_1^0, x_2^0, y_2^0) \in \mathbb{R}^4$ est une condition initiale telle que $a = E_1(x_1^0, y_1^0) > 0$ et $b = E_2(x_2^0, y_2^0) > 0$, la trajectoire par m est portée par un tore T_{ab} plongé dans $\mathbb{R}^4$, produit des deux ellipses $\{E_1(x_1, y_1) = a\} \subset \mathbb{R}^2 \times \{0\}$ et $\{E_2(x_2, y_2) = b\} \subset \{0\} \times \mathbb{R}^2$. On peut paramétrer T_{ab} par $(\theta_1, \theta_2) \in \mathbb{R}^2/\mathbb{Z}^2$ en posant

$$x_1 = \frac{\sqrt{a}}{\omega_1}\cos\theta_1, \quad y_1 = \sqrt{a}\sin\theta_1, \quad x_2 = \frac{\sqrt{b}}{\omega_2}\cos\theta_2, \quad y_2 = \sqrt{b}\sin\theta_2.$$

L'équation différentielle (5.2) se restreint sur T_{ab} en l'équation du champ constant : $\dot{\theta}_1 = \omega, \quad \dot{\theta}_2 = \omega_2$. Ce champ est équivalent (à la division par ω_1 près) au champ $\frac{\partial}{\partial\theta_1} + \alpha\frac{\partial}{\partial\theta_2}$ où $\alpha = \frac{\omega_2}{\omega_1}$ est le rapport des deux fréquences. Le flot du système restreint à chaque tore T_{ab} est donc irrationnel dès que le rapport des fréquences est irrationnel.

Fin de l'exemple du flot irrationnel sur le tore.

Voici maintenant un théorème d'existence pour les orbites récurrentes. Ce théorème est non trivial car il nécessite l'utilisation du lemme de Zorn (voir par exemple [13]).

Proposition 5.7. *Si la demi-orbite positive γ_x^+ est à fermeture compacte dans M, l'ensemble $\omega(x)$ contient au moins une orbite récurrente.*

Démonstration. On considère l'ensemble $\mathcal{A}$ de tous les compacts invariants non vides contenus dans $\omega(x)$. Comme $\omega(x)$ est lui-même un compact invariant et non vide (cf. la proposition 5.3 ainsi que la remarque 5.2), on a $\mathcal{A} \neq \emptyset$. Dans $\mathcal{A}$ on considère l'ordre partiel d'inclusion au sens large.

Montrons que $\mathcal{A}$ est un *ensemble inductif décroissant.* Cela signifie que si $K_1 \supseteq K_2 \supseteq \ldots$ est une chaîne totalement ordonnée d'éléments, il existe un élément $K \in \mathcal{A}$ plus petit que tous les éléments de la chaîne. Pour montrer cette propriété, il suffit de montrer que $K = \bigcap_i K_i \in \mathcal{A}$. En effet, K est compact et invariant comme intersection de compacts invariants. Pour montrer que $K \neq \emptyset$, il suffit d'utiliser la définition suivante des ensembles compacts : un ensemble A est compact, si et seulement si, pour toute famille de fermés $\{F_i\}_{i\in I}$ de A, dont toutes les sous-familles finies ont une intersection non vide, on a également que $\bigcap_{i\in I} F_i \neq \emptyset$. Ici, pour une chaîne totalement ordonnée $\{K_i\}_{i\in I}$ d'éléments de $\mathcal{A}$, une partie finie est totalement ordonnée, donc de la forme $K_{i_1} \supseteq \ldots \supseteq K_{i_\ell}$. Il en résulte que $\bigcap_{j=1}^{\ell} K_{i_j} = K_{i_\ell} \neq \emptyset$. De la définition de compact rappelée plus haut, il suit donc que $\bigcap_{i\in I} K_i \neq \emptyset$. Donc $\mathcal{A}$ est un ensemble inductif décroissant.

Alors par le *lemme de Zorn*, on peut trouver un élément *minimal* K_0 dans $\mathcal{A}$ (peut-être non unique). Rappelons qu'un ensemble minimal K_0 est un ensemble qui est tel que si $K' \subseteq K_0$, alors $K' = K_0$. Puisque K_0 est non vide, il existe $z \in K_0$. Montrons que z est nécessairement récurrent (c'est-à-dire $z \in \omega(z)$). Pour cela, considérons $\omega(z)$. C'est un compact invariant non vide : $\omega(z) \in \mathcal{A}$. De plus, $z \in K_0$ implique que $\overline{\gamma_z} \subset K_0$, car K_0 est fermé et invariant. Comme $\omega(z) \subset \overline{\gamma_z}$, on a $\omega(z) \subset K_0$. L'ensemble K_0 étant minimal, il suit que $\omega(z) = K_0$ et donc que $z \in \omega(z)$, ce qui signifie que z est un point récurrent. Ce qui conclut. ∎

Dans les exemples de la figure 5.1, le point $(0,0)$, l'orbite γ et le point de selle s sont des orbites récurrentes pour les figures (a), (b) et (c) respectivement. L'exemple (c) montre que tous les points d'un ensemble ω-limite ne sont pas nécessairement récurrents (à l'exception de s, aucun autre point de γ est récurrent).

La proposition 5.7 permet éventuellement de montrer que le champ a des orbites récurrentes (il suffit que X possède une partie compacte invariante non vide). Dans le prochain paragraphe, nous aurons l'occasion d'utiliser ce critère

d'existence pour les champs de vecteurs du plan. Comme la preuve de la proposition 5.7 utilise le lemme de Zorn, elle est non constructive : elle ne donne pas explicitement une orbite récurrente.

5.3. Récurrence pour les champs de vecteurs d'un ouvert de la sphère

Par souci de simplicité, on supposera dans cette section que X est de classe $\mathcal{C}^\infty$ et de flot complet. Rappelons que les orbites récurrentes trivialement sont les points singuliers et les orbites périodiques. On a vu avec le flot irrationnel sur T^2 (que l'on peut étendre en un champ de $\mathbb{R}^3$) un exemple de récurrence non triviale. Au contraire, nous allons voir que, sur $\mathbb{R}^2$, toutes les récurrences sont triviales.

Dans ce paragraphe, on supposera que X est un champ de vecteurs défini sur un ouvert de la sphère S^2. Un tel ouvert est, soit la sphère toute entière, soit un ouvert de $\mathbb{R}^2$. En fait, il se trouve que la situation est plus simple pour la sphère, qui a l'avantage d'être compacte, et, en particulier, tous les champs de vecteurs y sont à flot complet par la proposition 4.1. On essayera donc, lorsque cela sera possible, de se ramener à la considération de champs de vecteurs sur la sphère.

Par exemple, tout champ de vecteurs polynomial de $\mathbb{R}^2$ est équivalent à un champ qui s'étend à la sphère en un champ analytique. Précisons un peu cette construction. On identifie $\mathbb{R}^2$ à $\mathbb{C}$ par l'identification habituelle $(x, y) \in \mathbb{R}^2 \sim z = x + iy \in \mathbb{C}$. Nous avons montré, au paragraphe I-2.1.4, que l'on peut identifier la sphère S^2 à l'espace projectif $P^1(\mathbb{C})$ (voir l'item 5 du paragraphe I-2.1.3), défini par les deux cartes $\Pi_1 = (\mathbb{C}, z)$, $\Pi_2 = (\mathbb{C}, Z)$, et le changement de carte $Z = \frac{1}{z}$ défini pour $z \in \mathbb{C} \setminus \{0\}$, avec pour image $\mathbb{C} \setminus \{0\} = \{Z \neq 0\}$. La sphère $S^2 = \mathbb{C} \cup \{\infty\}$ où $\infty = \{Z = 0\}$ dans la carte Π_2, apparaît donc comme la compactification de $\mathbb{C}$ (la carte Π_1) obtenue par adjonction du point ∞.

Considérons maintenant un champ polynomial X sur $\mathbb{R}^2$. En utilisant la variable complexe z, un tel champ a pour équation différentielle : $\dot{z} = P(z, \bar{z})$, où P est un polynôme à coefficients complexes en $z, \bar{z}$. Supposons que le degré de P soit égal à $n \in \mathbb{N}$. La fonction $\left|P\left(z, \bar{z}\right)\right| \sim |z|^n$ pour $|z| \to +\infty$ et tend vers l'infini si $n \geq 1$. Cela conduit à introduire le champ de vecteurs $\tilde{X}$ d'équation différentielle :

$$\dot{z} = \frac{P(z, \bar{z})}{1 + (z\bar{z})^n}. \tag{5.3}$$

Le champ de vecteurs $\tilde{X}$ est analytique en $z, \bar{z}$, c'est-à-dire réel analytique en x, y. La fonction $1 + (z\bar{z})^n$ étant strictement positive sur $\mathbb{C}$, le champ $\tilde{X}$ est $\mathcal{C}^\infty$-équivalent au champ X.

Opérons le changement de cartes $z \mapsto Z = \frac{1}{z}$ sur l'équation (5.3). Nous obtenons :

$$\dot{Z} = -\frac{Z^2(Z\bar{Z})^n}{1+(Z\bar{Z})^n} P\Big(\frac{1}{Z}, \frac{1}{\bar{Z}}\Big). \tag{5.4}$$

La fonction $Q(Z,\bar{Z}) = (Z\bar{Z})^n P\Big(\frac{1}{Z}, \frac{1}{\bar{Z}}\Big)$ est un polynôme de degré inférieur ou égal à $2n$ en $Z, \bar{Z}$. Le second membre de l'équation (5.4) est donc analytique en Z, pour tout Z, y compris en $Z = 0$. Le champ $\tilde{X}$ qui est défini par (5.3) dans la carte $(\mathbb{C}, z)$ et par (5.4) dans la carte $(\mathbb{C}, Z)$ est un champ de vecteurs réel analytique sur S^2.

On peut aussi prolonger le champ X en un champ défini sur le disque D^2 en identifiant $\mathbb{R}^2$ à l'intérieur de D^2. Cela peut se faire à partir de la compactification sur S^2 que nous venons de décrire, en faisant un éclatement du point ∞. Cette opération d'éclatement (ou de désingularisation) sera décrite dans le chapitre III-6. Cette compactification sur le disque est appelée *compactification de Poincaré* (voir [25] par exemple).

5.3.1. Préambule : le théorème de Jordan

La propriété topologique du plan qui sera utilisée est la suivante :

***Théorème 5.1* (Théorème de Jordan : [9], [7] pages 261 à 264).** *Soit γ une courbe fermée simple plongée topologiquement dans la sphère S^2 (c'est-à-dire un homéomorphisme de S^1 sur son image). Alors $S^2 - \gamma$ a exactement deux composantes connexes A et B. Ces deux composantes A et B ont γ pour frontière commune, et leurs fermetures $\overline{A}$ et $\overline{B}$ sont homéomorphes au disque fermé.*

Rappelons que la frontière ∂O d'un ouvert O est égale à $\overline{O} - O$.

Comme le plan $\mathbb{R}^2$ s'identifie à la sphère S^2 privée d'un point, le théorème précédent est complètement équivalent à la version suivante sur le plan.

Corollaire 5.1. *Soit γ une courbe fermée simple plongée topologiquement dans le plan $\mathbb{R}^2$. Alors $S^2 - \gamma$ a exactement deux composantes connexes : l'une bornée A et l'autre non bornée B. Ces deux composantes A et B ont γ pour frontière commune et $\overline{A}$ est homéomorphe au disque fermé.*

Remarque 5.7. Le théorème de Jordan est faux pour les surfaces autres que la sphère et le plan (autrement dit pour les surfaces de genre supérieur ou égal à 1, voir la définition 4.13 du genre). Par exemple, le cercle $\mathbb{R}/\mathbb{Z} \times \{0\} \subset \mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z}$ (le tore de dimension 2) a un complément connexe. On le voit immédiatement en remarquant que ce complémentaire est égal à $\pi^{-1}(]0,1[)$, où π est la projection de

$\mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z}$ sur son premier facteur et $]0,1[$ est considéré comme intervalle ouvert de $\mathbb{R}/\mathbb{Z}$.

La démonstration du théorème de Jordan, pour les courbes simples topologiques générales, fait appel à des méthodes indirectes de topologie algébrique (voir [9] par exemple). En effet une telle courbe peut être assez pathologique : pensez aux courbes fractales (flocon de Koch par exemple) qui peuvent avoir une longueur infinie. Par contre, le cas des courbes $\mathcal{C}^1$ par morceaux est beaucoup plus simple, quoiqu'il serait assez long de donner une preuve en détails. Cela étant, comme nous n'allons utiliser le théorème de Jordan que pour les courbes $\mathcal{C}^1$ par morceaux, il semble intéressant d'avoir une idée d'une preuve possible du théorème dans ce cas.

Esquisse d'une preuve du théorème de Jordan pour une courbe $\mathcal{C}^1$ par morceaux

(a) Considérons une courbe $\mathcal{C}^1$ par morceaux γ dans le plan. Cette courbe est formée d'arcs réguliers successifs : $\gamma_1, \ldots, \gamma_k$, de classe $\mathcal{C}^1$, numérotés de façon cyclique, les arcs γ_i et γ_{i+1} ayant le point p_i comme unique point commun, $(p_1, \ldots, p_k)$ sont les sommets de γ, ce sont les seuls points où le vecteur dérivé peut être discontinu en direction orientée : si l'angle entre les deux côtés adjacents d'un sommet est différent de π. Il peut même se faire que l'angle entre deux côtés adjacents en un sommet p soit nul, auquel cas le point p est un point cuspidal de γ (figure 5.8).

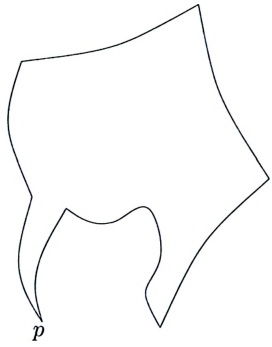


Figure 5.8. Exemple de sommet p cuspidal.

Dans ce cas, comme les arcs γ_i sont de classe $\mathcal{C}^1$, on peut trouver un système de coordonnées locales (x,y) au voisinage de p avec $p = (0,0)$, $(x,y) \in B = [-1,1] \times [-1,1]$, dans lequel les deux arcs adjacents en p, dont l'union est $B \cap \gamma$,

sont donnés par les graphes de deux fonctions $\mathcal{C}^1, \varphi, \psi : x \in [0,1] \mapsto [-1,1]$. Par hypothèse, on peut supposer que $\frac{d\varphi}{dx}(0) = \frac{d\psi}{dx}(0) = 0$, et que, pour tout $x \in [0,1]$, on ait $\varphi(x) < \psi(x)$. Il est alors très facile de contruire un homéomorphisme H du voisinage B dans lui-même, qui soit l'identité dans un voisinage du bord de B et qui envoie $B \cap \gamma$ sur l'union Γ des trois arcs suivants : le graphe de φ au-dessus de $[\frac{1}{2}, 1]$, le segment vertical $[\varphi(\frac{1}{2}), \psi(\frac{1}{2})] \times \{\frac{1}{2}\}$ et le graphe de ψ au-dessus de $[\frac{1}{2}, 1]$ (figure 5.9).

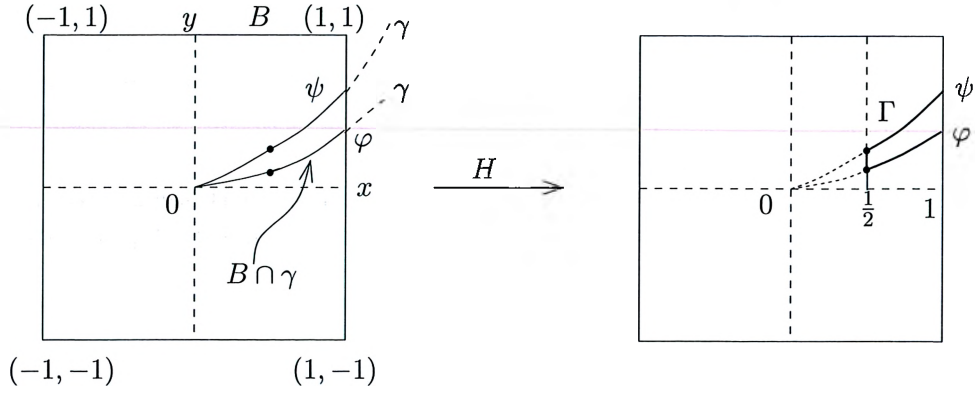


FIGURE 5.9. L'homéomorphisme H envoie $B \cap \gamma$ sur Γ. Γ est indiqué en traits plus épais sur la partie droite de la figure.

On peut appliquer cette construction à chaque sommet où l'angle est nul, c'est-à-dire en chaque sommet cuspidal. Les différents homéomorphismes locaux se prolongent par l'identité en un homéomorphisme global de $\mathbb{R}^2$, qui envoie la courbe γ sur une courbe $\mathcal{C}^1$ par morceaux, avec la propriété que les *angles à chaque sommet entre côtés adjacents sont non nuls.* Comme la propriété à démontrer est invariante par homéomorphisme, il suffit de la démontrer pour une telle courbe que l'on va encore appeler γ.

(b) *On peut approcher cette courbe γ par une courbe* Γ linéaire par morceaux, *telle qu'il existe un homéomorphisme du plan envoyant γ sur* Γ. Pour construire cet homéomorphisme, on commence par choisir des points suffisamment proches sur γ : $q_1, \ldots, q_l$ (en y incluant les sommets), de façon que les côtés correspondants $\tilde{\gamma}_1, \ldots, \tilde{\gamma}_l$ soient des graphes au-dessus des segments linéaires joignant les sommets ($\tilde{\gamma}_i$ joint q_i à q_{i+1}). Soit $\tilde{\gamma}$ la nouvelle courbe ainsi définie. Plus précisément, en utilisant à nouveau le théorème des fonctions implicites et la non-nullité des angles aux différents sommets de $\tilde{\gamma}$ (remarquez qu'aux points ajoutés en dehors des sommets de γ, l'angle vaut π), on peut trouver une suite de domaines rectangulaires $T_1, \ldots, T_l$ dont les intérieurs sont disjoints deux à deux, tels que,

dans T_i, l'arc $\tilde{\gamma}_i$ soit un graphe dans des coordonnées locales. Les côtés de T_i sont évidemment choisis parallèles aux axes de coordonnées locales et les extrémités de $\tilde{\gamma}_i$ sont portées par deux côtés opposés du bord de T_i (figure 5.10).

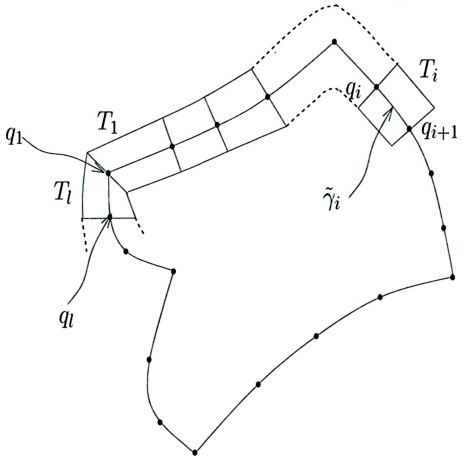


FIGURE 5.10

On peut alors construire dans chaque T_i un homéomorphisme H_i, égal à l'identité sur le bord de T_i et envoyant $\tilde{\gamma}_i$ sur le segment linéaire joignant les extrémités de ce côté. Ces homéomorphismes locaux se prolongent en un homéomorphisme global de $\mathbb{R}^2$ (à support compact, c'est-à-dire égal à l'identité en dehors d'un compact) envoyant la courbe γ sur la courbe linéaire par morceaux Γ, construite en reliant linéairement les points $q_1, \ldots, q_l$. Soit $\Sigma(\Gamma)$ l'ensemble fini de ces extrémités *(sommets)*, extrémités des segments linéaires *(arêtes)* constituant Γ.

On peut supposer que la *fonction hauteur* y n'ait que des extremums isolés sur Γ. Ces extremums appartiennent alors nécessairement à $\Sigma(\Gamma)$. On peut aussi supposer que les points de $\Sigma(\Gamma)$ sont à des hauteurs différentes (figure 5.11).

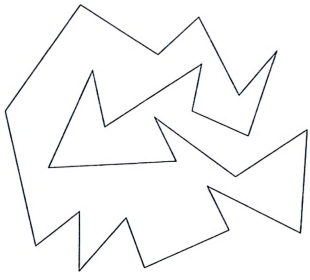


FIGURE 5.11. Une courbe Γ.

Une courbe linéaire par morceaux avec ces propriétés sera dite *générique*. Les maxima et les minima de la fonction hauteur y alternent le long d'une telle courbe Γ que l'on *supposera orientée dans le sens direct*. La fonction y est strictement croissante lorsqu'on va d'un minimum à un maximum, et strictement décroissante lorsqu'on va d'un maximum à un minimum. Si r, s sont deux points distincts de Γ, on notera par (r, s) le chemin linéaire par morceaux entre r et s, le long de Γ parcourue dans le sens direct. Comme la propriété à établir est invariante par homéomorphisme, *il suffira donc de l'établir dans le cas particulier des courbes linéaires par morceaux génériques.*

(c) *Pour commencer, nous allons établir le résultat lorsque la fonction y n'a que 2 extremums, un minimum y_m en (x_m, y_m) et un maximum y_M en (x_M, y_M).* En effet, dans ce cas, pour chaque $y \in]y_m, y_M[$, il existe exactement 2 points $(x_1(y), y)$ et $(x_2(y), y)$ appartenant à Γ_0. La réunion des intervalles $[x_1(y), x_2(y)] \times \{y\}$ avec les deux points $(x_m, y_m), (x_M, y_M)$ est un ensemble D homéomorphe au disque, avec Γ_0 pour bord (il est d'ailleurs facile de construire un homéomorphisme explicite entre le disque euclidien et D ; voir figure 5.12).

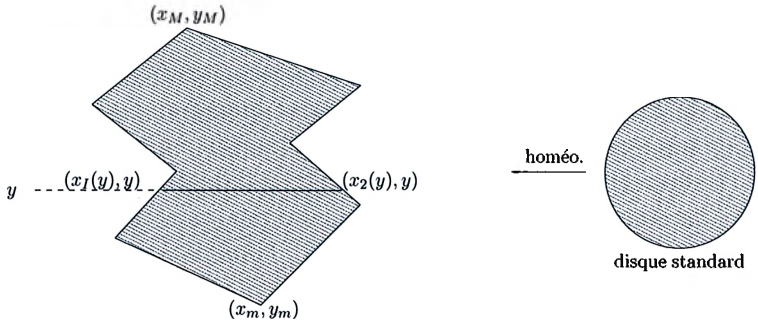


FIGURE 5.12. Cas à deux extremums. Le disque topologique D est à gauche sur la figure. Γ_0 est le bord du disque topologique D. Le disque standard est le disque euclidien usuel.

(d) Supposons maintenant que Γ soit une courbe linéaire par morceaux, générique, telle que la fonction hauteur ait strictement plus de deux extremums (ce nombre est un nombre pair $2s$ avec $s > 1$). *Alors, on peut construire un homéomorphisme h de $\mathbb{R}^2$, à support compact, tel que la courbe $h(\Gamma)$ soit linéaire par morceaux, générique, avec une fonction hauteur ayant deux extremums de moins que ceux de Γ.*

Pour ce faire, on choisit, en parcourant Γ dans le sens direct, un quadruplet de sommets p, m, M, q avec, par exemple, la configuration suivante :

1. m est un minimum, M est un maximum, le chemin (p, q) ne contient pas d'autre extremum dans son intérieur, et (p, m) est une arête de Γ (ce qui n'est pas supposé pour (m, M) et (M, q)).

2. La droite horizontale par m coupe (M, q) en un point intérieur r (ce point est nécessairement unique).

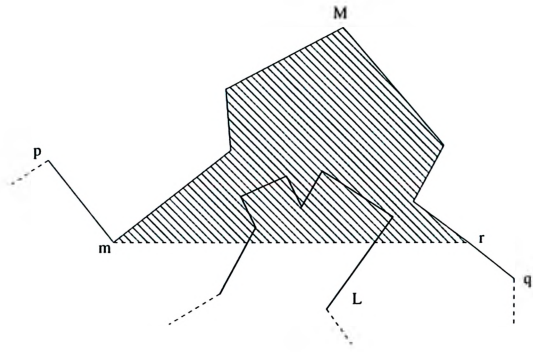


FIGURE 5.13. Le disque topologique D est hachuré. Le sous-ensemble L de Γ est indiqué en traits plus épais.

3. Si D est le disque topologique ayant pour bord l'union des arcs (m, M), (M, r) et le segment horizontal $[r, m]$, alors (p, m) n'est pas contenu dans D (figure 5.13).

Quoique assez intuitive, l'existence d'une telle configuration, ou bien d'une configuration équivalente (obtenue en permutant les rôles ou bien l'ordre du maximum et du minimum), est probablement le point le plus délicat de la preuve. Nous admettons cette existence. Nous allons maintenant passer à la construction de h.

Remarquez tout d'abord que le disque topologique D peut couper $L \subset \Gamma - (p, r)$. Cela suppose que L coupe le segment horizontal ouvert $]m, r[$ (figure 5.13). On peut, dans ce cas, trouver un homéomorphisme h_1, égal à l'identité sur un voisinage de (p, r), envoyant Γ sur une courbe Γ_1 linéaire par morceaux et générique, avec un même nombre d'extremums que Γ, tout en envoyant L dans le complémentaire de D (figure 5.14).

Le segment horizontal ouvert $]m, r[$ est maintenant disjoint de Γ_1. Soit r' un point sur (r, q), distinct de r. Si r' est choisi assez proche de r, on peut maintenant trouver un deuxième homéomorphisme h_2 qui laisse fixe Γ_1 en dehors de $(m, r') = (m, M) \cup (M, r') \subset \Gamma_1$, et qui envoie ce chemin sur le segment linéaire oblique $[m, r']$ (figure 5.15). Soit $h = h_2 \circ h_1$. La courbe $\Gamma_2 = h(\Gamma)$ est linéaire par morceaux et générique, avec deux extremums de moins, les points m et M : on a

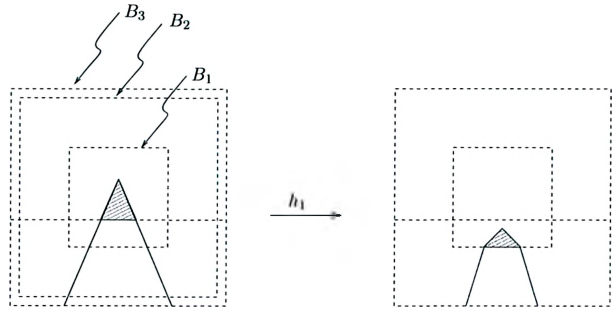


FIGURE 5.14. On prend h_1 dans B_1 comme indiqué dans la figure, avec $h_1 = Id$ dans $B_3 - B_2$ (en notant l'identité par Id).

remplacé l'oscillation verticale $(p,m) \cup (m,M) \cup (M,q)$ à deux extremums par le chemin $(p,m) \cup [m,r'] \cup (r',q) \subset \Gamma_2$, sur lequel la fonction hauteur est strictement décroissante et donc sans extremum (figure 5.15).

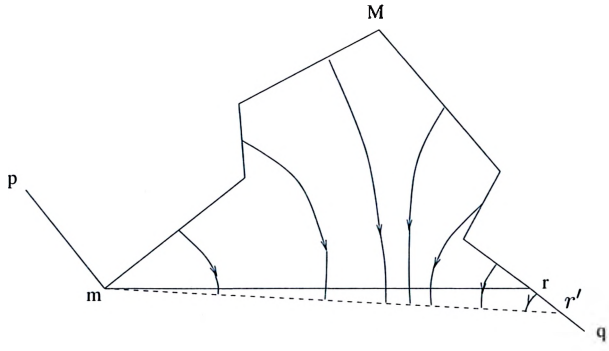


FIGURE 5.15. Deuxième étape de l'élimination. On supprime les extremums m et M en remplaçant « (p,m,M,q) » par « (p,m,r',q) ».

(e) Comme la propriété, pour une courbe fermée, d'être le bord d'un disque topologique est invariante par homéomorphisme, le résultat pour les courbes linéaires par morceaux, génériques, suit par récurrence des points (c) et (d). En effet, partant d'une courbe linéaire par morceaux, générique, Γ quelconque avec $2l$ extrema, on peut construire en appliquant (d) une suite d'homéomorphismes $h_1, \ldots, h_{l-1}$ qui, appliqués de façon récurrente, diminue de 2 le nombre d'extremums de la fonction hauteur à chaque étape. Si $h = h_{l-1} \circ \ldots \circ h_1$, la courbe $h(\Gamma)$ qui n'a plus que 2 extremums, est le bord d'un disque topologique d'après (c).

Grâce aux points (a), (b), il s'ensuit le théorème de Jordan pour les courbes $\mathcal{C}^1$ par morceaux. ∎

5.3.2. Théorème de Poincaré-Bendixson

On peut maintenant prouver le Théorème 5.2

***Théorème 5.2* (Théorème de Poincaré-Bendixson).** *Soit X un champ défini sur un ouvert $U \subset S^2$. Alors, X n'a pas de récurrence non triviale (si x est un point récurrent, alors $X(x) = 0$, ou bien γ_x est périodique).*

Démonstration. Prouvons ce théorème par l'absurde. Supposons donc que l'orbite par x soit une orbite récurrente non triviale. Cela suppose que $X(x) \neq 0$. On peut alors choisir un voisinage tubulaire B. Avec un léger abus de notation, on écrira $B = [-\tau, \tau] \times I$ où $x = (0,0)$, et si $(u,y) = [-\tau,\tau] \times I$ alors $X = \frac{\partial}{\partial u}$ sur B.

Comme γ_x est récurrente, la trajectoire $\varphi(t,x)$ revient indéfiniment dans B *et donc sur* I par une petite translation en temps ; c'est-à-dire que l'on peut choisir une suite de temps $t_1 < t_2 < \cdots < t_i < \ldots$, $(t_i)_i \longrightarrow \infty$ telle que $x_i = \varphi(t_i, x) \in I$. On peut supposer $(x_i)_i \longrightarrow x$ et comme γ_x *est supposée non périodique* les $x_i = \varphi(t_i,x)$ sont deux à deux disjoints. On peut, d'autre part, supposer que t_1 est le premier temps de retour sur I, c'est-à-dire que $\varphi(t,x) \notin I$, pour $t \in]0, t_1[$.

Considérons l'arc de trajectoire Γ entre $x = 0$ et $x_1 = \varphi(t_1, x)$. Soit de plus $\sigma = [x, x_1] \subset I$. Comme x_1 est le premier point de retour, $\sigma \cap \Gamma$ est composé des deux points x et x_1, et $\sigma \cup \Gamma = C$ est une courbe $\mathcal{C}^1$ par morceaux, homéomorphe au cercle, qui est incluse dans $U \subset S^2$. D'après le théorème 5.1, la région complémentaire de C dans S^2 est la réunion de deux composantes connexes : A_1 et A_2 (figure 5.16).

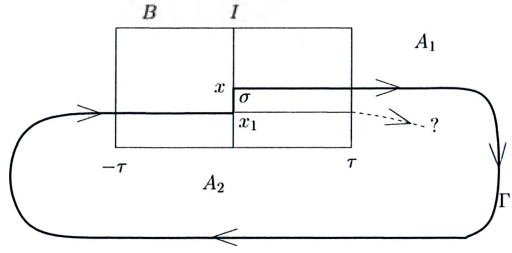


FIGURE 5.16. C est montrée en traits plus épais.

Supposons que A_2 soit la composante dans laquelle entre la trajectoire pour $t > t_1$ et $t - t_1$ petit. Montrons que, pour tout t tel que $t > t_1$, on a $\varphi(t,x) \in A_2$. En effet, pour entrer à nouveau dans A_1, par raison de connexité, la trajectoire devrait recouper la courbe C : soit Γ, soit σ. Cela est équivalent à dire qu'il existerait un $T > 0$ tel que $\varphi(t_1 + T, x) \in C$ et $\varphi(t,x) \in A_2$ pour $t_1 < t < t_1 + T$. Nous avons

les deux possibilités suivantes : $\varphi(t_1+T,x) \in \Gamma$ ou bien $\varphi(t_1+T,x) \in \sigma$. Nous allons montrer que ces deux possibilités sont impossibles.

Dans le premier cas, on aurait un t', $0 \leq t' \leq t_1$ tel que $\varphi(t',x) = \varphi(t_1+T,x)$ et donc $\varphi(T+t_1-t',x) = x$. Comme $T+t_1-t' > 0$, l'orbite γ_x serait périodique, ce qui n'est pas le cas, car nous avons supposé que la récurrence par x était non triviale.

La deuxième possibilité serait que $\varphi(t_1+T,x) \in \sigma$. Alors, pour tout $\eta > 0$ assez petit, $\varphi(t_1+T-\eta,x)$, appartenant à A_2, serait contenu dans le voisinage tubulaire B, à droite de I. Cela implique que $\frac{\partial\varphi}{\partial t}(t_1+T,x) = X(\varphi(t_1+T,x))$ aurait une composante négative ou nulle le long de $\frac{\partial}{\partial u}$, ce qui n'est évidemment pas le cas.

Nous venons de montrer que $\varphi(t,x) \in A_2$ pout $t > t_1$. Montrons maintenant que cela est en contradiction avec l'hypothèse $(x_i)_i \to x$. Cette hypothèse signifie que, si i est assez grand, $x_i = \varphi(t_i,x) \in I$ sera très proche de x. Pour un tel point x_i et si $\eta > 0$ est assez petit, on a $\varphi(-\eta,x_i) \in A_1$ et donc $\varphi(t_i-\eta,x) \in A_1$ (figure 5.17).

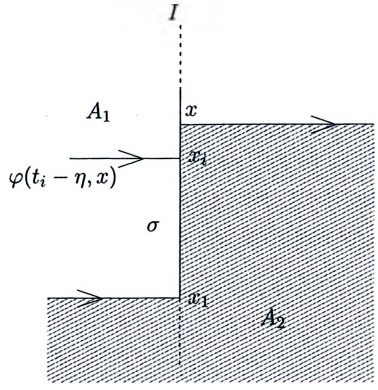


FIGURE 5.17. A_2 (resp. A_1) est la partie hachurée (resp. non hachurée).

Mais évidemment, pour un indice i assez grand, $(t_i-\eta) > t_1$, ce qui implique que $\varphi(t_i-\eta,x) \in A_2$ comme nous l'avons vu ; on arrive à une contradiction car A_1 et A_2 sont disjoints. Ce qui démontre le théorème. ∎

Remarque 5.8. 1. Un ouvert U de $\mathbb{R}^2$ peut être considéré comme un ouvert de $S^2 = \mathbb{R}^2 \cup \{\infty\}$. Le théorème 5.2 est donc vrai pour un champ X défini sur un ouvert $U \subset \mathbb{R}^2$, et en particulier pour $U = \mathbb{R}^2$.

2. La même démonstration, que celle utilisée dans le théorème 5.2, permet de montrer que, si I est un intervalle transverse au champ, et si une trajectoire

$\varphi(t,x)$ coupe I pour une suite d'instants successifs $t_1 < \ldots < t_i < \ldots$ en des points $x_i = \varphi(t_i, x)$ (éventuellement une infinité de fois), alors l'application $t_i \to x_i$ est monotone : la trajectoire doit nécessairement tourner en formant une *spirale* (figure 5.18).

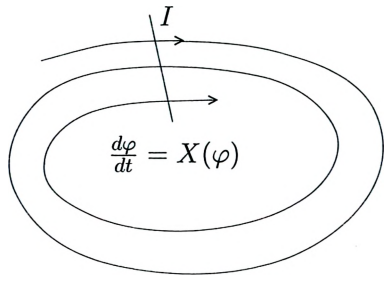


FIGURE 5.18. Illustration du théorème de Poincaré-Bendixson : comportement en spirale de l'orbite par un point x non récurrent du champ.

5.3.3. Applications du théorème de Poincaré-Bendixson

Le théorème de Poincaré-Bendixson est l'outil principal que l'on peut utiliser dans l'étude qualitative d'un champ de vecteurs sur un ouvert du plan. Nous allons en donner deux exemples.

Existence d'orbite périodique

Le première application concerne la recherche des orbites périodiques. Elle a été invoquée dans l'introduction de cet ouvrage.

Proposition 5.8 **(Piège à orbite périodique).** *Soit X un champ défini sur un voisinage d'un anneau A (c'est-à-dire d'une région compacte difféomorphe à $S^1 \times [0,1]$). Supposons que X soit sans singularités sur A et entrant dans A le long de ∂A. Alors, X a au moins une orbite périodique dans A (figure 5.19).*

Démonstration. Le champ X est défini sur un voisinage ouvert U de A. On peut supposer que X n'a pas de zéros sur U. Soit $x \in U$ et γ_x^+ la demi-trajectoire positive. Puisque le champ est entrant sur l'anneau, les trajectoires ne peuvent pas en sortir et donc $\gamma_x^+ \subset A \Rightarrow \overline{\gamma_x^+} \subset A$. D'après la proposition 5.7, $\overline{\gamma_x^+}$, qui est non vide et compact (comme fermé du compact A) contient des points récurrents y. Comme U est un ouvert de $\mathbb{R}^2$ et que, par hypothèse, $X(y) \neq 0$ pour

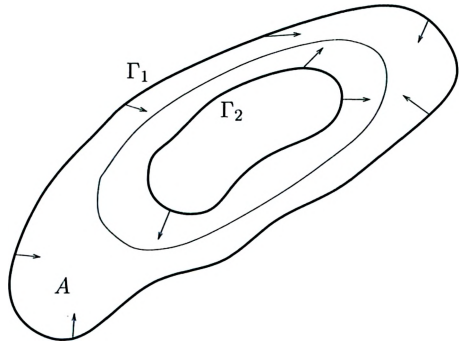


FIGURE 5.19

tout $y \in A$, il suit du théorème de Poincaré-Bendixson que γ_y est périodique dans A. ∎

Le type de preuve ci-dessus (de l'existence d'une solution périodique) est devenu classique dans la théorie des systèmes dynamiques : elle a été utilisée par exemple pour l'équation de Van der Pol. Le point difficile est de repérer un anneau adéquat pour le champ.

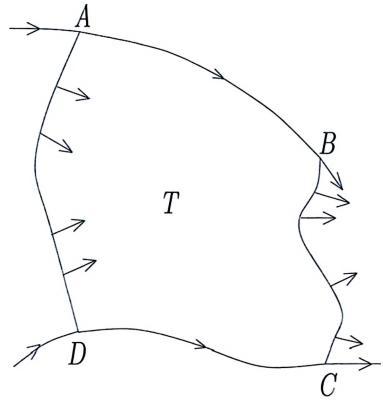


FIGURE 5.20

Théorème du grand voisinage tubulaire

Théorème 5.3. *Soit X un champ défini sur un voisinage d'un rectangle curviligne T de sommets A, B, C, D (c'est-à-dire que T est un domaine difféomorphe au carré $[0,1] \times [0,1]$). On suppose que les côtés AB et DC sont des arcs de trajectoires, que le champ est transverse et entrant dans T le long du côté AD et transverse*

et sortant de T le long de BC (figure 5.20). On suppose de plus que X n'a pas de zéros dans T. Alors T est, à équivalence près, un grand voisinage tubulaire pour X *au sens suivant : il existe un difféomorphisme Ψ du carré $[0,1]\times[0,1]$ de coordonnées (u,v) sur T qui envoie le champ constant $\frac{\partial}{\partial u}$ sur un champ fX (où f est une fonction de classe $\mathcal{C}^\infty$ et positive).*

Démonstration. Supposons que le côté entrant AD soit paramétré par une application $\nu : v \in [0,1] \to \mathbb{R}^2$ avec $\nu(0) = D$ et $\nu(1) = A$. Pour tout v, la trajectoire de X par $x = \nu(v)$ doit quitter T en un temps fini $t(v)$. En effet, dans le cas contraire, l'ensemble limite $\omega(x)$ serait un compact non vide contenu dans l'ouvert U, intérieur de T. Par la proposition 5.7, $\omega(x)$ doit contenir une orbite récurrente γ qui, d'après le théorème de Poincaré-Bendixson appliqué à U, est un zéro de X, ou bien une orbite périodique. Le premier cas est exclu par hypothèse. Si, d'autre part, γ était une orbite périodique, elle serait le bord d'un domaine D de U difféomorphe au disque (l'intérieur de D est la composante connexe bornée de $\mathbb{R}^2 - \gamma$, donnée par le théorème de Jordan). Le champ X, qui est tangent au bord $\partial D = \gamma$, devrait avoir au moins un zéro dans D (en effet, le disque D compact a une caractéristique d'Euler égale à 1, le champ X doit donc avoir un point singulier dans D en vertu du théorème 4.1 de Poincaré-Hopf).

La fonction $t : v \mapsto t(v)$ est de classe $\mathcal{C}^\infty$, comme il suit facilement du théorème des fonctions implicites (on reviendra sur ce point à propos de l'application de Poincaré, qui sera définie dans le chapitre 6). D'autre part $t(v) > 0$ pour tout $v \in [0,1]$ (remarquez aussi que $\varphi(t(v), \nu(v)) \in BC$). Considérons maintenant l'application $\Psi : (u,v) \in [0,1]\times[0,1] \mapsto \mathbb{R}^2$ définie par $\Psi(u,v) = \varphi(ut(v), \nu(v))$. Si $v \neq v'$, on aura pour tout u, u' que $\Psi(u,v) \neq \Psi(u',v')$. Il en résulte que Ψ est un difféomorphisme qui envoie le champ $\frac{\partial}{\partial u}$ sur le champ fX, où f est définie par $f(\Psi(u,v)) = t(v)$. ∎

Le théorème précédent permet d'établir la trivialité du portrait de phase sur de grands domaines : en effet, elle fournit un difféomorphisme du carré sur le rectangle curviligne T, en envoyant un champ constant sur un champ équivalent au champ X. Son intérêt est que les conditions du théorème sont de nature topologique, puisqu'elles portent : sur le fait que le domaine est homéomorphe au disque, sur le comportement du champ sur le bord du domaine, enfin sur l'absence de points singuliers dans le domaine. Il est relativement facile de l'étendre à d'autres situations similaires. Par exemple, si le domaine considéré est difféomorphe à un disque D et que le champ de vecteurs, sans singularités sur un voisinage de D, n'a que deux points de contact avec le bord ∂D, pour lesquels la trajectoire a un contact quadratique et est située localement à l'extérieur de D, comme il est montré sur la figure 5.21. On dit que deux courbes de $\mathbb{R}^2$ ont un contact

quadratique en p, s'il existe un difféomorphisme local de $\mathbb{R}^2$ en p envoyant : p sur l'origine, une des courbes sur l'axe des x et l'autre sur le graphe d'une fonction $y = \varphi(x)$, telle que $\varphi(0) = \frac{d\varphi}{dx}(0) = 0$ et $\frac{d^2\varphi}{dx^2}(0) \neq 0$. Une autre conséquence facile du théorème est que, si un champ X est défini dans un voisinage d'un anneau A, sans point singulier, transverse entrant à A sur une composante C_1 de ∂A, transverse sortant sur l'autre composante C_2 et avec un arc de trajectoire joignant C_1 avec C_2, alors il existe un difféomorphisme Ψ de $[0,1] \times S^1$, de coordonnées (u, v) avec $u \in [0,1]$ et $v \in S^1$, tel que X sur $A = \Psi([0,1] \times S^1)$ soit équivalent au champ constant $\frac{\partial}{\partial u}$. On trouvera un exemple d'application du théorème du grand voisinage tubulaire dans la preuve du théorème III-4.2.

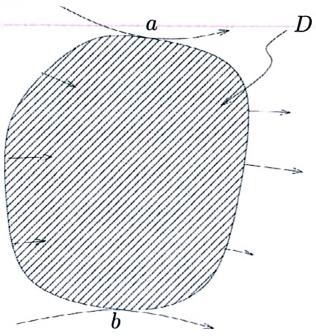


FIGURE 5.21. Aux points a et b la trajectoire du champ a un contact quadratique avec ∂D.

5.3.4. Vers la théorie de Poincaré-Bendixson

Le théorème 5.2 est le point de départ de la théorie de Poincaré-Bendixson, qui vise à la classification des ensembles limites des champs de vecteurs sur les ouverts du plan (voir [21] par exemple). Nous nous contenterons de montrer que, dans le cas où $\omega(x)$ contient une orbite périodique, on peut préciser davantage le comportement de la trajectoire par x.

Proposition 5.9. *Supposons que $\omega(x)$ contienne une orbite périodique γ. Alors $\omega(x) = \gamma$ ou bien, ce qui est équivalent : $\varphi(t, x) \to \gamma$ pour $t \to \infty$.*

Démonstration. Soit $y \in \gamma \subset \omega(x)$. On peut trouver un segment I transverse à X au point y (figure 5.22). Les intersections de γ_x avec I forment une suite ordonnée de points x_i, définis par $\varphi(t_i, x) = x_i$ pour des temps t_i, adhérente à y. En vertu de la remarque 5.8, la suite $(t_i)_i$ est monotone et tend en croissant vers l'infini (rappelons que cette remarque peut être établie par des arguments tout

à fait analogues à ceux utilisés dans la démonstration du théorème de Poincaré-Bendixson). L'orbite γ étant périodique est compacte. En utilisant cette propriété de compacité de γ, on peut alors établir que, si d_i est la distance entre le segment d'orbite Γ_i entre les points consécutifs x_i et x_{i+1} et l'orbite γ, la suite $(d_i)_i$ tend vers 0 quand i tend vers $+\infty$. Comme pour $t > 0$ assez grand le point $\varphi(t, x)$ appartient à la réunion des segments Γ_i, il s'ensuit que $\varphi(t, x) \to \gamma$ pour $t \to +\infty$. ∎

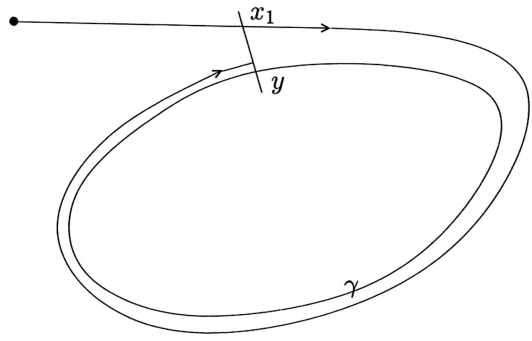


FIGURE 5.22

Remarque 5.9. Ce résultat est valable pour tout champ en dimension 2 sur une surface quelconque.

Par contre, si $\omega(x)$ contient un point critique et si $\omega(x)$ *est supposée compacte*, la situation est beaucoup plus compliquée (on sait que $\omega(x)$ ne peut pas contenir d'orbite périodique en vertu de la proposition 5.9). Par exemple, la théorie de Poincaré-Bendixson permet de montrer que si X *est analytique et que tous ses zéros sont isolés,* alors $\omega(x)$ est un *polycycle singulier* Γ (aussi appelé *graphique*) qui est une réunion finie de points singuliers connectés par des séparatrices qui sont des trajectoires tendant vers des points singuliers pour $t \to \pm\infty$. Évidemment, Γ peut aussi se réduire à un seul point singulier. Le paragraphe précédent et la proposition 5.9 sont les premiers pas dans cette théorie de Poincaré-Bendixson.

Ainsi, par exemple, si le champ X est entrant dans le disque et si X est analytique à zéros isolés, on connaît tous les ensembles ω possibles : *orbites périodiques* (si $\omega(x)$ ne contient pas de zéros isolés) ou *polycycles* (figure 5.23). Dans tous les cas, la trajectoire tend vers $\omega(x)$ comme il a été dit à la proposition 5.5. Ce résultat n'est pas facile à démontrer mais il donne toute la phénoménologie du comportement des solutions des champs de vecteurs des ouverts de la sphère.

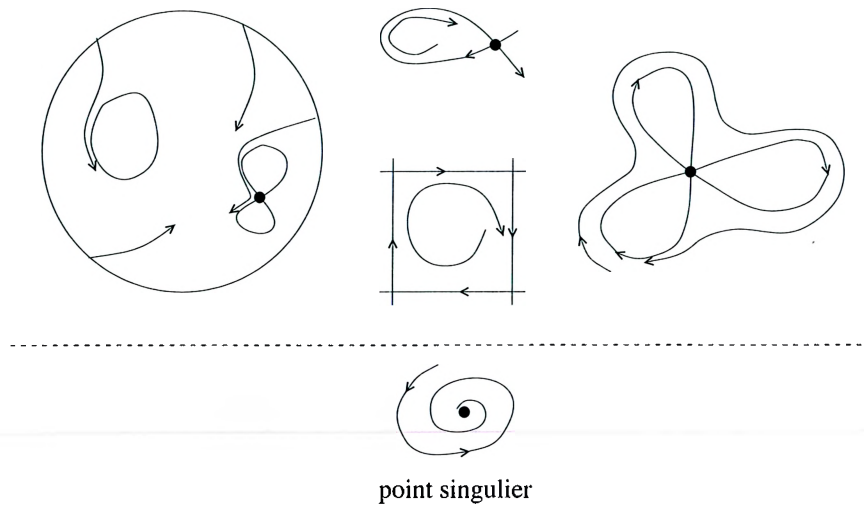


FIGURE 5.23. Exemples d'ensembles limites.

On trouvera plus de détails sur la théorie de Poincaré-Bendixson, ainsi que d'autres exemples de polycycles dans [21].

6

ORBITES ET CHAMPS PÉRIODIQUES

Dans ce chapitre nous étudions le comportement du flot d'un champ de vecteurs (autonome) au voisinage d'une orbite périodique γ_p. Une fois choisie une section Σ à l'orbite par le point p, telle qu'elle a été définie dans le chapitre 4, on peut définir l'application de retour h sur la section, aussi appelée application de Poincaré : cette application est le premier retour sur Σ des trajectoires issues de points de Σ. C'est un difféomorphisme défini localement sur Σ au voisinage du point p. On peut lire sur l'application de Poincaré toutes les propriétés locales du flot au voisinage de γ_p, à équivalence près.

Certains champs de vecteurs ont des sections globales (voir section 6.2), coupées par toutes les trajectoires. Une fois choisie une section globale Σ pour un tel champ, on peut définir l'application de Poincaré globale h qui est un difféomorphisme de la sous-variété Σ de codimension 1 sur elle-même. En fait, un difféomorphisme h d'une variété Σ définit une dynamique discrète par itérations (on introduit succintement cette idée dans le présent chapitre ; ce point de vue consistant à considérer un difféomorphisme comme une dynamique, sera amplement développé dans le tome 2). On peut lire sur le difféomorphisme h toutes les propriétés du flot du champ, examiné à équivalence près. Par exemple, les orbites périodiques du champ sont en correspondance biunivoque avec les orbites périodiques de h. La correspondance entre champ avec section globale en dimension n et difféomorphisme en dimension $n-1$, admet une correspondance inverse : on peut suspendre tout difféomorphisme en dimension $n-1$ en un champ de vecteurs avec section globale en dimension n.

À un champ de vecteur non autonome $X(x,t)$ dépendant périodiquement du temps, et défini « spatialement » sur une variété M, on associe un champ autonome $X+\frac{\partial}{\partial t}$, en considérant le temps t comme une variable supplémentaire (voir

chapitre 1). À cause de la périodicité, ce champ passe au quotient sur la variété $M \times S^1$ en un champ $\widetilde{X}$. Si X est complet, le champ $\widetilde{X}$ admet $M \times \{0\}$, par exemple, comme section globale. L'étude du champ non autonome X se ramène, de cette façon, à l'étude de la dynamique discrète sur M, obtenue par l'intégration de X sur une période.

Dans la suite, et comme dans les chapitres précédents, on appellera champ de vecteurs un champ de vecteurs supposé autonome, sauf mention expresse du contraire.

6.1. Orbites périodiques

Soit X un champ de vecteurs défini sur une variété M de dimension n. On le suppose $\mathcal{C}^\infty$ (quoique $\mathcal{C}^1$ suffise dans la suite). On rappelle que φ désigne le flot de X. Par commodité, on suppose aussi que le flot est complet, c'est-à-dire que $\varphi(t,x)$ est défini pour tout $(t,x) \in \mathbb{R} \times M$.

On suppose que γ_p est une orbite périodique du champ X. Considérons une section Σ au champ X, contenant le point p dans son intérieur (figure 6.1). La notion de section a été définie dans le chapitre 4 de cette partie. Rappelons qu'une section, pour un champ X dans une variété M de dimension n, est une sous-variété difféomorphe au disque ouvert de dimension $n-1$, transverse au champ en chacun de ses points. Pratiquement, on peut construire des sections par tout point $p \in M$ où $X(p) \neq 0$, en restreignant le champ à un ouvert de carte contenant le point p, et en choisissant pour section un disque dans un hyperplan $H \sim \mathbb{R}^{n-1}$ transverse en p au vecteur $X(p)$, dans l'image de la carte (un ouvert de $\mathbb{R}^n$).

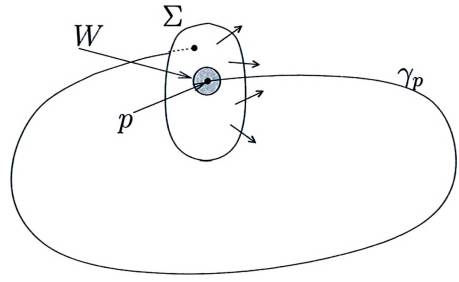


FIGURE 6.1. Le voisinage ouvert W est l'ensemble ombré. Dans cette figure, Σ est plongée dans M.

Pour étudier le voisinage de l'orbite périodique γ_p, on va lui associer une application d'un voisinage W de p dans Σ, à valeurs dans Σ. Cette application est dite application de Poincaré ou aussi *application de premier retour* pour une raison évidente de définition.

Nous établissons tout d'abord l'existence d'un *temps de premier retour* des trajectoires issues des points de Σ, assez proches du point p :

Proposition 6.1. *Soit γ une orbite périodique de période (minimale) T, $p \in \gamma$ et Σ une section locale par p. On suppose que*

$$\gamma \cap \overline{\Sigma} = \{p\}, \tag{6.1}$$

où $\overline{\Sigma}$ désigne la fermeture de Σ dans M. Alors, on peut trouver un voisinage ouvert W de p dans Σ et une fonction C^∞, $t : x \in W \mapsto \mathbb{R}^+$, telle que $t(x)$ soit le temps de premier retour *de la trajectoire $\varphi(t,x)$ sur Σ (c'est-à-dire que $\varphi(t(x),x) \in \Sigma$ et $\varphi(t,x) \notin \Sigma$ pour $0 < t < t(x)$). Évidemment on a $t(p) = T$.*

Commentaire. L'hypothèse γ périodique, $p \in \gamma$, interdit que $X(p) = 0$ (proposition 4.3), et donc il existe une section locale par p (proposition 4.4). La proposition 6.1 implique, par continuité, que $X(\varphi(t(x),x)) \neq 0$ pour $x \in W$, W suffisamment petit, et que la trajectoire $\varphi(t(x),x)$ est transverse à W pour $x \in W$.

Démonstration. Soit un voisinage tubulaire (B,Φ) autour du point p, où Φ est un difféomorphisme de $D^{n-1}\times]-\tau,\tau[$ sur B (voir figure 6.2). On peut supposer que $\Phi(D^{n-1}\times\{0\}) \subset \Sigma$ (voir la remarque 4.8 pour la discussion de la possibilité d'un tel choix) et que $\overline{\Sigma}\cap\gamma = \{p\}$, grâce à (6.1). Puisque $T > 0$ est la période de γ, il vient $\varphi(T,p) = p$ et

$$\varphi(]0,T'],p) \cap \overline{\Sigma} = \emptyset, \tag{6.2}$$

si $0 < T' < T$. Choisissons T' tel que $T - \tau < T' < T$. Le point $\varphi(T',p)$ est dans B à gauche de Σ.

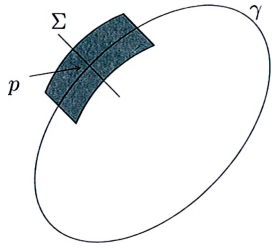


FIGURE 6.2. Le voisinage tubulaire B est le domaine ombré de la figure associée au cas unidimensionnel.

Par continuité, il existe un $\varepsilon > 0$ tel que pour $\|x - p\| < \varepsilon$ et $x \in \Sigma$, on ait $\varphi(T', x) \in \text{int } B$ et soit à gauche de Σ.

Soit $W = \{x \in \Sigma \cap B \mid \|x - p\| < \varepsilon\} \subset \Sigma$. Il suit de (6.2), que l'on peut choisir ε assez petit pour que

$$\varphi(]0, T'], x) \cap \overline{\Sigma} = \emptyset, \tag{6.3}$$

pour tout $x \in W$.

Soit $u : B \mapsto [-\tau, \tau]$ la fonction définie par $u(m) = \pi \circ \Psi(m)$ où $\Psi = \Phi^{-1}$ et π est la projection $\pi(x, u) = u$. Remarquez que l'on passe du point $m \in B$ à la section Σ en un temps $-u(m)$. On pose alors, pour $x \in W$,

$$t(x) = T' - u(\varphi(T', x)). \tag{6.4}$$

Cette fonction est évidemment $\mathcal{C}^\infty$. D'autre part, on a $\varphi(t(x), x) \in \Sigma$ (figure 6.3). En effet

$$\varphi(t(x), x) = \varphi(-u(\varphi(T', x)) + T', x) = \varphi(-u(\varphi(T', x)), \varphi(T', x))$$

appartient à Σ, comme il a été remarqué plus haut.

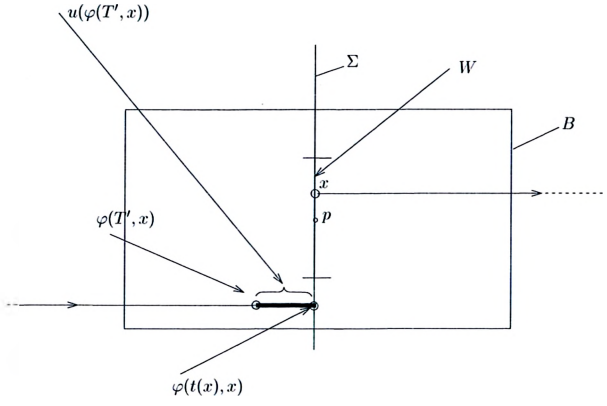


FIGURE 6.3. Illustration de la formule (6.4). En trait gras le temps $t(x)$.

Le temps $t(x)$ que l'on vient de définir est le temps de premier retour. En effet, comme $\varphi(T', x)$ est à gauche de Σ, on a $u(\varphi(T', x)) < 0$ et donc

$$\varphi([0, -u(\varphi(T', x))[, \varphi(T', x)) \cap \Sigma = \emptyset. \tag{6.5}$$

Comme $\varphi(]0, t(x)[, x) = \varphi(]0, T'], x) \cup \varphi([0, -u(\varphi(T', x))[, \varphi(T', x))$, on a donc que $\varphi(]0, t(x)[, x) \cap \Sigma = \emptyset$, en vertu de (6.3) et de (6.5). Le fait que $t(p) = T$ suit directement de la condition (6.1). ∎

Remarque 6.1. La condition (6.1) est là pour assurer que le temps de retour $t(p)$ soit égal à la période T, et aussi que le temps de retour soit continu en p. Cette condition peut se réaliser en choisissant une section en p assez petite. Remarquez que si cette condition n'est pas remplie, il peut se faire que $\varphi(]0,T[,p) \cap \partial\Sigma \neq \emptyset$, où $\partial\Sigma = \overline{\Sigma} - \Sigma$ est le bord de Σ. L'application $t : x \mapsto t(x)$ sera alors discontinue en p. En effet la condition ci-dessus signifie qu'il existe une valeur T_1, $0 < T_1 < T$ telle que $\varphi(T_1,p) \in \partial\Sigma$. On peut alors trouver un point p_1 de Σ, arbitrairement proche de p, pour lequel le temps de retour est proche de T, et, inversement, un point p_2 arbitrairement proche de p, pour lequel le temps de retour est proche de T_1, d'où la discontinuité du temps de retour. Cette pathologie peut arriver pour certaines sections trop grandes du champ (figure 6.4).

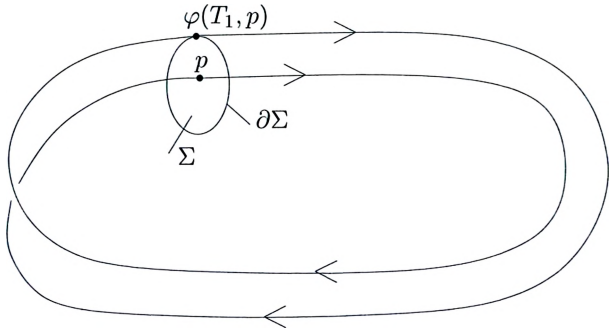


FIGURE 6.4. Illustration de la discontinuité éventuelle du temps de retour si la section est trop grande.

Nous allons maintenant considérer l'*application de premier retour* ou *application de Poincaré* associée au premier temps de retour d'une orbite périodique.

Proposition 6.2. *Soit W le voisinage de p dans Σ comme dans la proposition 6.1. On définit l'application $h : W \mapsto \Sigma$ par $h(x) = \varphi(t(x),x)$. Alors h est un difféomorphisme $\mathcal{C}^\infty$ (aussi différentiable que le champ X si ce champ est de classe de différentiabilité finie), de W sur son image $h(W) \subset \Sigma$ (voir figure 6.5).*

Démonstration. Comme la fonction temps de retour $t : x \mapsto t(x)$ ainsi que le flot $\varphi(t,x)$ sont $\mathcal{C}^\infty$ (voir la preuve de la proposition 6.1 et le théorème 3.4 car le champ X est supposé – dans ce chapitre – $\mathcal{C}^\infty$), l'application h est également de classe $\mathcal{C}^\infty$. Il reste à prouver qu'elle admet un inverse différentiable.

L'idée est de considérer le champ $-X$. Il est trivial de constater que si $\varphi(t,m)$ est le flot de X, alors $\psi(t,x) = \varphi(-t,m)$ est le flot du champ $-X$. Il en résulte que, partant de $y = h(x)$, on atteint x par le flot ψ au temps $-t(x)$ et que ce

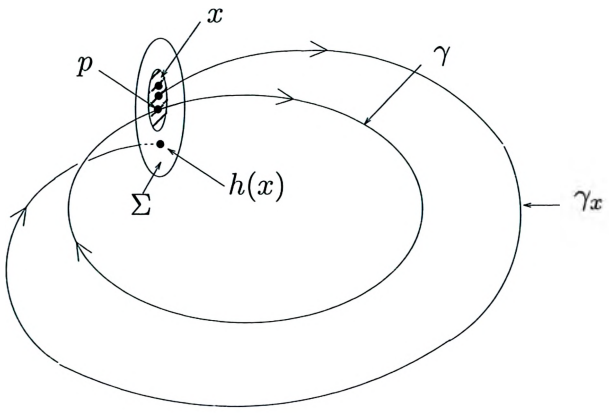


FIGURE 6.5. W est la partie hachurée.

temps est le temps de premier retour pour le champ $-X$, à partir du point y. Donc si $H(y)$ est l'application de Poincaré de $-X$ relative à Σ, H est définie sur $h(W)$ et

$$H(y) = H \circ h(x) = x, \quad \forall x \in W,$$

Autrement dit, H est donc l'inverse à gauche de h (figure 6.6).

En inversant l'argument, on obtient que $h(x) = h \circ H(y) = y$, pour tout $y \in h(W)$, ce qui signifie que H est aussi inverse de h à droite et donc que h et H sont des applications inverses l'une de l'autre : $H = h^{-1}$ sur W. L'application H est $\mathcal{C}^\infty$ en tant qu'application de Poincaré du champ $-X$ et donc h est un difféomorphisme $\mathcal{C}^\infty$ de W sur $h(W)$. Ce qui conclut la preuve de la proposition. ■

Définition 6.1. On appelle *application de premier retour* ou *application de Poincaré* relative à la section Σ, l'application de W dans Σ définie par $h(x) = \varphi(t(x), x)$. Pour tout $x \in W$ le point $h(x)$ est donc la première intersection pour les temps croissants de la trajectoire de x avec Σ (figure 6.5).

Remarque 6.2. On peut faire les observations suivantes à propos de l'application de Poincaré :

1. L'application h dépend des choix de Σ et de $W \subset \Sigma$, et W doit être choisi assez petit. On peut dire que l'application de Poincaré est une définition purement locale.

2. Généralement $h(W)$ n'est pas inclus dans W (cela pourra être, par exemple, le cas si l'orbite γ est asymptotiquement stable, au sens où on le définira dans

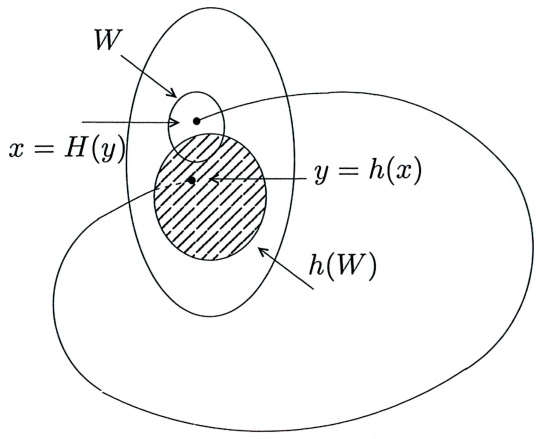


FIGURE 6.6

le prochain chapitre). Aussi, on ne peut pas prendre, en général, $W = \Sigma$, même si la section Σ est choisie petite.

3. L'application de Poincaré h ne dépend pas des *trajectoires mais seulement des orbites* (malgré le fait que la définition utilise le flot φ), car on peut justement définir h comme le premier point d'intersection, autre que x, de γ_x avec Σ, lorsque l'on parcourt γ_x dans le sens des t croissants. Ainsi, puisque X et fX (où $f(x) > 0$ et f est $\mathcal{C}^\infty$) ont les mêmes orbites (proposition 3.5), ces deux champs de vecteurs ont les mêmes applications de retour (définies pour les mêmes W et Σ).

On peut généraliser les notions de temps et d'application de retour, en définissant des temps et des applications de transitions entre deux sections locales à une orbite quelconque, non nécessairement périodique. Soit Σ_1 et Σ_2 deux sections locales, respectivement par deux points différents p_1 et p_2 choisis sur la même orbite γ. Il existe un plus petit temps $T_{12} > 0$ pour aller de p_1 à p_2. On suppose aussi que

$$\varphi([0, T_{12}], p_1) \cap \overline{\Sigma_2} = p_2.$$

De la même façon que plus haut, on peut montrer que si W_1 est un voisinage ouvert assez petit de p_1 dans Σ_1, il existe un temps de (première) transition $t_{12}(x)$ de classe $\mathcal{C}^\infty$, qui est le temps minimal pour aller de $x \in W_1$ à Σ_2 (on a évidemment $t_{12}(p_1) = T_{12}$). On définit ensuite une application de transition entre W_1 et Σ_2 par $h_{12}(x) = \varphi(t_{12}(x), x)$. Cette application de transition est un difféomorphisme de W_1 sur son image dans Σ_2.

Le point $h_{12}(x)$ est le premier point d'arrivée pour les temps positifs sur Σ_2, de la trajectoire par $x \in W_1$. L'application h_{12} ne dépend que des orbites, car c'est le passage d'une section à une autre le long d'une même orbite dans le sens des temps croissants. Si Σ_1, Σ_2, et Σ_3 sont trois sections consécutives, on aura évidemment l'égalité $h_{13}(x) = h_{23}(x) \circ h_{12}(x)$ sur un voisinage assez petit de p_1 dans Σ_1.

Supposons à nouveau que l'orbite γ soit périodique et soit $p_1, p_2 \in \gamma$. Si h_1 et h_2 sont des applications de Poincaré relatives à Σ_1 et Σ_2, h_{12} une transition de Σ_1 dans Σ_2, h_{21} une transition de Σ_2 dans Σ_1 (définies au voisinage de p_1 et p_2 respectivement), on a la figure 6.7 et les égalités vérifiées localement au voisinage de p_1, p_2 respectivement :

$$\begin{cases} h_1 = h_{21} \circ h_{12} \\ h_2 = h_{12} \circ h_{21} \end{cases},$$

d'où il suit

$$h_{12} \circ h_1 = h_2 \circ h_{12},$$

ce qui s'écrit aussi

$$h_2 = h_{12} \circ h_1 \circ h_{12}^{-1}. \tag{6.6}$$

La formule (6.6) correspond au diagramme de commutation (6.7).

$$\begin{array}{ccc} \Sigma_1 & \xrightarrow{h_1} & \Sigma_1 \\ {\scriptstyle h_{12}}\downarrow & & \downarrow{\scriptstyle h_{12}} \\ \Sigma_2 & \xrightarrow{h_2} & \Sigma_2 \end{array} \tag{6.7}$$

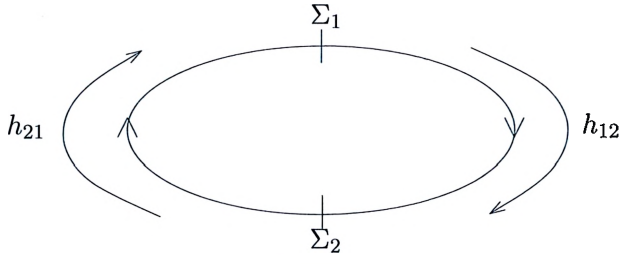


FIGURE 6.7

Le difféomorphisme h_{12} est une *conjugaison* entre h_1 et h_2, sur des voisinages assez petits de p_1 et p_2. Si l'on identifie localement les sections, le difféomorphisme de conjugaison h_{12} peut être interprété comme un changement de coordonnées qui transforme h_1 en h_2.

Lorsque deux difféomorphismes diffèrent par conjugaison, nous verrons qu'ils ont les mêmes propriétés dynamiques. Comme le choix de la section locale Σ

est arbitraire, les propriétés qui ont un sens, qui sont intéressantes, sont celles qui s'affranchissent de cet arbitraire ; c'est-à-dire que les *bonnes notions relatives à h* sont celles qui sont *invariantes par conjugaison*, et ce sont celles que l'on va exclusivement étudier. Par exemple, la stabilité d'une orbite périodique, les multiplicateurs de Floquet, sont invariants par conjugaison, c'est-à-dire que ce sont des propriétés ou quantités qui sont indépendantes du choix de la section locale.

Commentaire : utilité de l'application de Poincaré

L'application de Poincaré sert à l'étude des propriétés locales du champ de vecteurs au voisinage de l'orbite périodique γ. Naturellement toutes les propriétés locales ne seront pas obligatoirement détectables par cette application. Par exemple, notons que l'application de Poincaré h « ne voit pas » le temps de retour (puisque deux champs multiples l'un de l'autre par une fonction $f > 0$ et C^∞ ont même application de Poincaré h, mais ils n'ont pas nécessairement le même temps de retour). Par exemple encore, l'application de Poincaré ne pourra pas fournir la valeur des périodes des orbites périodiques coupant la section.

Par contre, toute propriété locale invariante par *équivalence* du champ est, par définition, détectable par h (locale signifie ici : ne dépendant que de la donnée du champ dans un voisinage arbitraire de l'orbite périodique). Les propriétés locales en question seront celles qui sont invariantes par conjugaison de l'application de Poincaré, car pour être intéressantes, elles doivent être indépendantes du choix de la section. Voici quelques exemples.

1. *La stabilité d'une orbite périodique* s'étudie grâce aux valeurs propres de h qui sont appelés les *multiplicateurs de Floquet* ; or ces multiplicateurs sont invariants par conjugaison, et donc indépendants du choix de la section locale. On reviendra sur cette question dans le prochain chapitre. Le lecteur peut s'en convaincre en considérant, par exemple, des systèmes linéaires. Les applications h_1, h_2 et h_{12} sont alors représentées respectivement par des matrices A, B et S, et la relation (6.6) se traduit par $A = SBS^{-1}$; les matrices A et B sont semblables et ont mêmes valeurs propres. Nous étudierons la stabilité des orbites périodiques dans le prochain chapitre.

2. *Certaines des orbites périodiques* voisines de l'orbite périodique considérée sont détectées grâce à h. On dit qu'un point x est *k-périodique* pour $h : W \mapsto \Sigma$ si $x, h(x), \ldots, h^{k-1}(x) \in W$ avec $h^k(x) = x$ et $h^j(x) \neq x$ pour

$1 \leq j \leq k-1$. Dans cette définition, on note par h^j l'itéré j fois de h, c'est-à-dire : $h^j = \underbrace{h \circ h \circ \ldots \circ h}_{j \text{ fois}}$.

Remarquez que l'on pourrait ne pas pouvoir itérer h indéfiniment et même une seule fois, car pour certains $x \in W$, on peut avoir $h(x) \notin W$. D'autre part, cette propriété de retour dépend du choix de W, choix qui a un caractère arbitraire. Il n'y a que le point p que l'on peut à coup sûr itérer indéfiniment par h, car $h(p) = p$.

Si x est un point k-périodique de h, l'orbite γ_x de X qui lui correspond est évidemment elle-même *périodique* de période

$$T_x = t(x) + t(h(x)) + \ldots + t(h^{k-1}(x)).$$

Évidemment T_x est approximativement de l'ordre de kT, car pour x voisin de p, $t(x) \simeq T$ (c'est un raisonnement par continuité : si $x \to p$, $t(x) \to T$ et l'on suppose que le nombre k est fixé).

Cependant, il y a des orbites périodiques de X que l'on pourrait ne pas repérer en ne considérant que les points périodiques de h. Ce sont les orbites pour lesquelles pour un certain itéré h^i, $h^i(x) \notin W$, et on ne peut plus poursuivre l'itération de h ! En quelque sorte, cette orbite s'écarte trop de γ. Il peut donc exister des orbites périodiques coupant W ne correspondant pas à des points périodiques de h : il n'y a donc pas correspondance entre les orbites périodiques de X et celles de h. Naturellement, plus on restreint W, plus on risque de perdre des orbites périodiques du champ ; en particulier, étant donné deux orbites périodiques, on ne peut pas *a priori* leur associer une section locale commune.

Les deux faits évoqués ci-dessus montrent que certaines propriétés dynamiques du champ de vecteurs, au voisinage d'une orbite périodique, s'étudient par des propriétés correspondantes de l'application de Poincaré. Cela a conduit à considérer les propriétés dynamiques de l'itération d'un difféomorphisme, comme on le fera dans le tome 2. Un inconvénient lié à l'utilisation de l'application de Poincaré tient à son caractère local, qui fait que l'on ne peut pas, en principe, l'itérer indéfiniment à partir de n'importe quel point x de son domaine de définition.

Cela ne sera pas le cas si le champ X possède une section globale, comme on va le voir dans le prochain paragraphe. On aura, dans ce cas, une correspondance parfaite entre les propriétés dynamiques du champ (propriétés définies à équivalence près, c'est-à-dire indépendantes du temps d'intégration), et les propriétés dynamiques de l'application de Poincaré (propriétés définies à conjugaison près). Dans le cas particulier des champs de vecteurs non autonomes périodiques, qui

seront considérés dans le dernier paragraphe, on pourra même considérer les propriétés liées au temps d'intégration, car le retour sur la section se fera, dans ce cas, en un temps constant : la période du champ de vecteurs.

6.2. Section globale, suspension

6.2.1. Section globale pour un champ de vecteurs

Définition 6.2. Soit X un champ de vecteurs $\mathcal{C}^\infty$ sur une variété M de dimension n. Soit Σ une sous-variété connexe et fermée de dimension $n-1$ (une hypersurface). $\Sigma \subset M$ est une *section globale* si les propriétés suivantes sont vérifiées :

1. Toute orbite coupe Σ.
2. X est transverse à Σ pour tout $x \in \Sigma$, ce qui s'écrit : $X \pitchfork_x \Sigma$ pour tout $x \in \Sigma$, ou bien simplement : $X \pitchfork \Sigma$.
3. La trajectoire par $x \in \Sigma$ revient sur Σ pour un temps de premier retour $t(x) > 0$. On exige également cette propriété de retour pour le champ $-X$ (c'est-à-dire pour les temps décroissants du flot de X).

Remarque 6.3. Cette définition implique que la variété M est égale à la réunion des orbites des points de la section Σ. On dit que M est obtenue par *saturation* de Σ par le flot de X. De plus, partant d'un point quelconque $x \in M$, on arrive sur Σ en un temps fini, aussi bien positif ou négatif. Ainsi Σ sera une section globale pour X si et seulement si elle est une section globale pour $-X$.

Propriétés.

1. On peut appliquer à ce temps de retour l'étude locale faite précédemment : $x \in \Sigma \mapsto t(x) \in \mathbb{R}^+$ est une fonction $\mathcal{C}^\infty$. En effet, si $p \in \Sigma$ et si $p' = \varphi(t(p), p)$ est le premier point de retour, *il suit du fait que* Σ *est fermée* que l'on peut trouver une section locale Σ' en p', contenue dans Σ, telle que $\varphi([0, t(p)], p) \cap \overline{\Sigma'} = p'$; comme dans le paragraphe précédent (proposition 6.1), on montre que le premier retour sur Σ se fera aussi sur Σ', pour les $x \in \Sigma$ suffisamment proches de p. On peut alors appliquer le résultat du paragraphe précédent pour obtenir que $t(x)$ est de classe $\mathcal{C}^\infty$.

La fonction $h(x) = \varphi(t(x), x)$ est appelée *application de premier retour* ou bien *application de Poincaré relative à* Σ.

2. Comme pour le cas local, la classe de différentiabilité de h est égale à celle du champ, donc ici de classe $\mathcal{C}^\infty$. L'application h est un difféomorphisme $\mathcal{C}^\infty$ *global* (et non plus local) de la section Σ sur elle-même. En conséquence, on pourra donc itérer h indéfiniment, sans restriction.

3. L'existence de Σ comme section globale et la définition de h ne dépendent que des orbites de X. Aussi ces notions ne dépendent de X qu'à équivalence $\mathcal{C}^\infty$ près (on peut remplacer X par fX, avec $f > 0$ et de classe $\mathcal{C}^\infty$).

4. Un champ de vecteurs avec section globale est assez particulier. Par exemple, chaque orbite est régulière, puisqu'elle doit couper la section transversalement (donc en un point où le champ n'est pas nul). *Il en résulte qu'un champ de vecteurs avec une section globale n'a pas de points singuliers* (voir le commentaire de la proposition 6.1 pour la justification de cette assertion).

6.2.2. Suspension d'un difféomorphisme

Soit X un champ de vecteurs sur une variété M, muni d'une section globale Σ. Rappelons que tout champ $\mathcal{C}^\infty$-équivalent à X admet aussi Σ comme section globale. Nous allons tout d'abord montrer que l'on peut toujours changer un tel champ à équivalence $\mathcal{C}^\infty$ près, de façon que le temps de retour sur Σ soit constant :

Proposition 6.3. *Soit X un champ de vecteurs sur une variété M, muni d'une section globale Σ. Alors, il existe une fonction $\mathcal{C}^\infty$ positive f sur M, telle que le temps de premier retour sur M pour le champ fX soit égal à 1, quel que soit le point $x \in \Sigma$.*

Démonstration. On rappelle que la fonction premier temps de retour $t(x)$ est de classe $\mathcal{C}^\infty$ et strictement positive pour tout x. Désignons par φ le flot de X. Soit $\Phi : \Sigma \times [0,1] \mapsto M$ l'application de classe $\mathcal{C}^\infty$ définie par $\Phi(x,u) = \varphi(ut(x), x)$. Cette application qui est localement inversible, est un difféomorphisme global sur $\Sigma\times]0,1[$ et envoie $\Sigma\times\{0\}$ et $\Sigma\times\{1\}$ difféomorphiquement sur Σ (on peut dire que l'on passe de M à $\Sigma\times[0,1]$ en *coupant* M le long de Σ). Le champ de vecteurs X se relève en un champ de vecteurs $\widetilde{X}$ sur $\Sigma \times [0,1]$, c'est-à-dire tel que $\Phi_*(\widetilde{X}) = X$, et il est donné explicitement par $\widetilde{X}(x,u) = (d\Phi(x,u))^{-1}[X(\Phi(x,u))]$. Les orbites de $\widetilde{X}$ sont les segments $\{x\} \times [0,1]$: autrement dit ce champ est tangent à la direction de la variable u et dirigé dans le sens des u croissants. Il s'écrit donc : $\widetilde{X}(x,u) = f(x,u)\frac{\partial}{\partial u}$ pour une fonction $f(x,u)$ de classe $\mathcal{C}^\infty$ et partout strictement positive.

Nous allons maintenant modifier le champ X à équivalence près pour avoir un temps de premier retour égal à 1. Pour ce faire, nous allons plutôt modifier le champ $\widetilde{X}$. Considérons tout d'abord une fonction $\psi : u \in [0,1] \mapsto \mathbb{R}$ de classe $\mathcal{C}^\infty$, qui soit nulle sur $[0,\eta]$ et $[1-\eta,1]$, pour un η, $0 < \eta < \frac{1}{2}$, et telle que $\psi(u) > 0$ pour tout $u \in]\eta, 1-\eta[$ (on peut par exemple choisir : $\psi(u) = \exp \frac{1}{(u-\eta)(u-1+\eta)}$ pour $u \in]\eta, 1-\eta[$).

Nous allons considérer sur $\Sigma \times [0,1]$ le champ de vecteurs

$$\widetilde{Y}(x,u) = \exp(-\lambda(x)\psi(u))\widetilde{X}(x,u),$$

où $\lambda(x)$ est une fonction de classe $\mathcal{C}^\infty$ définie sur Σ, à choisir. Rappelons que, par construction, le champ $\widetilde{X}$ passe au quotient par Φ, c'est-à-dire : $\Phi_*(\widetilde{X}) = X$. Comme les champs $\widetilde{X}$ et $\widetilde{Y}$ coïncident dans un voisinage des deux composantes du bord de $\Sigma \times [0,1]$ (grâce au choix de la fonction ψ), le champ $\widetilde{Y}$ passe aussi au quotient : il existe un champ de vecteurs Y sur M tel que $\Phi_*(\widetilde{Y}) = Y$. Comme $Y \equiv X$ sur tout un voisinage de $\Sigma \subset M$ et que Φ est un difféomorphisme en restriction à $\Sigma \times]0,1[$, le champ Y est aussi de classe $\mathcal{C}^\infty$. Les champs $\widetilde{X}$ et $\widetilde{Y}$ étant $\mathcal{C}^\infty$-équivalents, il en sera de même pour les champs X et Y.

L'équation de la trajectoire du champ $\widetilde{Y}$ issue du point $(x,0)$ se résume en l'équation unidimensionnelle :

$$\frac{du}{ds}(s,x) = \exp(-\lambda(x)\psi(u(s,x))f(u(s,x),x)$$

pour la u-composante $u(s,x)$, où s est le temps d'intégration pour le champ $\widetilde{Y}$ (la composante en x est nulle). Il en résulte que le temps $s(x)$ pour aller de $(x,0)$ à $(x,1)$ est égal à

$$s(x) = \int_0^1 \frac{\exp(\lambda(x)\psi(u))}{f(u,x)} du.$$

Remarquez que cette intégrale a un sens car le dénominateur $f(u,x)$ est une fonction strictement positive. Nous voulons trouver une fonction $\lambda(x)$, de classe $\mathcal{C}^\infty$, telle que $s(x) \equiv 1$. Pour ce faire, considérons la fonction

$$s(x,\lambda) = \int_0^1 \exp(\lambda\psi(u)) \frac{du}{f(u,x)}$$

définie pour $(x,\lambda) \in \Sigma \times \mathbb{R}$. Remarquons que

$$\frac{\partial s}{\partial \lambda}(x,\lambda) = \int_0^1 \psi(u) \exp(\lambda\psi(u)) \frac{du}{f(u,x)} > 0,$$

pour tout (x,λ). D'autre part, il est très facile de vérifier que, pour tout x fixé, on a : $s(x,\lambda) \to 0$ pour $\lambda \to -\infty$ et $s(x,\lambda) \to +\infty$ pour $\lambda \to +\infty$. Pour chaque

$x \in \Sigma$, la fonction $\lambda \mapsto s(x, \lambda)$ est donc un difféomorphisme de $\mathbb{R}$ sur $]0, +\infty)$: il en résulte qu'il existe une (unique) fonction $\lambda(x)$ telle que $s(x) = s(x, \lambda(x)) \equiv 1$. Il suit du théorème des fonctions implicites que cette fonction $\lambda(x)$ est de classe $\mathcal{C}^\infty$.

Le champ Y correspondant est $\mathcal{C}^\infty$-équivalent au champ X et a un temps de retour sur Σ constant et égal à 1. ∎

Remarque 6.4. La démonstration donnée ci-dessus a l'avantage d'être élémentaire. En utilisant un peu de topologie différentielle, il est possible de donner une preuve plus succinte. En effet, la première partie de la démonstration ci-dessus montre que M est l'espace total d'une fibration différentiable sur S^1, de fibre Σ. (Voir [12] par exemple pour la définition de fibré différentiel.) Le champ X est transverse à cette fibration. On peut alors relever le champ constant $\frac{\partial}{\partial u}$ du cercle (paramétré par u modulo $\mathbb{Z}$), en un champ Y parallèle à X. Ce champ est $\mathcal{C}^\infty$-équivalent au champ X et le temps de retour sur Σ est égal à 1, le temps pour parcourir S^1 avec le flot du champ constant. En effet si $\pi : M \mapsto S^1$ est l'application de fibration et si $\varphi(t, x)$ est le flot de Y, on a $\pi \circ \varphi(t, x) = \pi(x) + t$, puisque le flot de $\frac{\partial}{\partial u}$ est égal à $\varphi_0(t, u) = t + u$.

On peut remarquer que cet argument donnerait un champ analytiquement équivalent au champ X si ce dernier champ était analytique (réel), ce que ne permet pas d'obtenir la démonstration de la proposition 6.3.

À partir de maintenant, on va supposer que le temps de retour sur Σ est égal à 1. Dans ce cas, l'application de Poincaré est donnée par $h(x) = \varphi(1, x)$. Posons $\Sigma_t = \varphi(t, \Sigma)$. Σ_t est une hypersurface de M pour tout t avec évidemment $\Sigma_n = \Sigma$, pour tout $n \in \mathbb{Z}$. En fait, pour tout t, l'hypersurface Σ_t est une section globale. En particulier, on voit que Σ_t est transverse au champ en remarquant que le difféomorphisme de M, $\varphi_t : x \mapsto \varphi_t(x) = \varphi(t, x)$, envoie Σ sur Σ_t, en envoyant le vecteur $X(x)$ sur le vecteur $X(\varphi(t, x))$.

Nous allons montrer que le champ de vecteurs X provient d'un champ trivial de $\Sigma \times \mathbb{R}$ par un passage au quotient.

Proposition 6.4. *Soit X un champ de vecteurs sur une variété M, muni d'une section globale Σ, et une fonction temps de retour sur Σ constante et égale à 1. Soit $\widetilde{\Phi}$ l'application de $\Sigma \times \mathbb{R}$ sur M définie par :*

$$\widetilde{\Phi}(x, u) = \varphi(u, x).$$

Alors, $\widetilde{\Phi}$ est un difféomorphisme local (une application localement inversible) et, pour tout $u \in \mathbb{R}$:

$$\widetilde{\Phi}(\Sigma \times \{u\}) = \Sigma_u.$$

De plus, pour tout $(x,u) \in \Sigma \times \mathbb{R}$, *on a :*

$$d\widetilde{\Phi}(x,u)[\frac{\partial}{\partial u}] = X(\widetilde{\Phi}(x,u)). \tag{6.8}$$

Remarque 6.5. Dans la formule : $\widetilde{\Phi}(x,u) = \varphi(u,x)$, on joue sur le double statut de Σ qui est une sous-variété de M (terme de droite de la formule), et donc aussi une variété (terme de gauche de la formule).

Preuve de la proposition 6.4. On a $\frac{\partial \varphi}{\partial u}(u,x) = X(\varphi(u,x)) = X(\widetilde{\Phi}(x,u)) = d\widetilde{\Phi}(x,u)[\frac{\partial}{\partial u}]$, ce qui est l'égalité souhaitée. Remarquons que $\widetilde{\Phi}(x,u)|_{\Sigma\times\{u\}} = \varphi_u|_\Sigma$ est un difféomorphisme de Σ sur Σ_u (voir la remarque ci-dessus). Il en résulte que

$$d\widetilde{\Phi}(x,u)|_{\Sigma\times\{u\}} = d\varphi_u|_\Sigma$$

est un isomorphisme de $T_x\Sigma$ sur $T_{\widetilde{\Phi}(x,u)}\Sigma_u$. De plus on a vu que

$$d\widetilde{\Phi}(x,u)[\frac{\partial}{\partial u}] = X(\widetilde{\Phi}(x,u)).$$

Il en résulte que $\widetilde{\Phi}$ est un difféomorphisme local, car $X(\widetilde{\Phi}(x,u))$ est transverse à $T_{\widetilde{\Phi}(x,u)}\Sigma_u$. ∎

On peut définir un difféomorphisme F de $\Sigma \times \mathbb{R}$ sur lui-même par la formule

$$F(x,u) = (h^{-1}(x), u+1).$$

Comme $F(\Sigma\times\{u\}) = \Sigma\times\{u+1\}$, il est clair que ce difféomorphisme agit sur $\Sigma \times \mathbb{R}$ de façon propre et libre. Rappelons (voir l'item 3 du paragraphe I-2.1.3), que cela est équivalent à dire que tout point (x,u) admet un voisinage W tel que $W \cap F^{\circ n}(W) = \emptyset$ pour tout $n \in \mathbb{Z} - \{0\}$, où $F^{\circ n}$ désigne l'itéré n fois de F pour le produit de composition. Comme on l'a vu au paragraphe I-2.1.3, on peut alors définir une variété quotient $(\Sigma\times\mathbb{R})/F$ en identifiant (x,u) avec $F(x,u)$ pour tout $(x,u) \in \Sigma\times\mathbb{R}$. D'autre part, le champ de vecteurs constant $\frac{\partial}{\partial u}$ est invariant par F, c'est-à-dire $F_*(\frac{\partial}{\partial u}) = \frac{\partial}{\partial u}$. Il induit donc un champ de vecteurs sur la variété quotient, champ que l'on notera encore $\frac{\partial}{\partial u}$. Compte tenu de $h^{-1}(x) = \varphi(-1,x)$, on a

$$\widetilde{\Phi}(F(x,u)) = \widetilde{\Phi}(\varphi(-1,x),u+1) = \varphi(u+1,\varphi(-1,x)) = \varphi(u,x) = \widetilde{\Phi}(x,u).$$

Cela signifie que l'application $\widetilde{\Phi}$ est invariante par F, et donc induit une application Φ par passage au quotient (voir paragraphe I-2.1.2)

$$\Phi : (\Sigma\times\mathbb{R})/F \mapsto M.$$

Cette application est un difféomorphisme $\mathcal{C}^\infty$ de $(\Sigma \times \mathbb{R})/F$ sur M, car il est défini par passage au quotient d'un difféomorphisme $\widetilde{\Phi}$ local $\mathcal{C}^\infty$ (au quotient, ce qui est valide pour les homéomorphismes l'est pour les difféomorphismes), et deux points de $(\Sigma \times \mathbb{R})/F$ ont même image par Φ si et seulement s'ils diffèrent par un itéré de F. Par définition du quotient, on a $\widetilde{\Phi} = \rho \circ \Phi$, où ρ désigne la projection canonique de $\Sigma \times \mathbb{R}$ sur $(\Sigma \times \mathbb{R})/F$. Compte tenu de la formule (6.8), il en résulte que $\Phi_*(\frac{\partial}{\partial u}) = X$. Nous avons donc montré le résultat suivant :

Proposition 6.5. *Soit X un champ de vecteurs avec une section globale Σ. Soit h l'application de Poincaré relative à Σ et $F(x,u) = (h^{-1}(x), u+1)$ le difféomorphisme de $\Sigma \times \mathbb{R}$. Alors, il existe un difféomorphisme Φ de la variété quotient $\Sigma \times \mathbb{R}/F$ sur M qui envoie le champ constant $\frac{\partial}{\partial u}$ sur le champ X.*

Le sous-ensemble $D = \Sigma \times [0,1] \subset \Sigma \times \mathbb{R}$ est un domaine fondamental pour la variété quotient $(\Sigma \times \mathbb{R})/F$. On peut donc construire ce quotient en restreignant l'identification à D, comme nous l'avons expliqué au paragraphe I-2.1.2 :

$$(\Sigma \times \mathbb{R})/F = \Sigma \times [0,1]/(x,0) \sim (h^{-1}(x),1).$$

On peut dire que l'on a coupé M le long de Σ, pour obtenir une variété qui est difféomorphe à $\Sigma \times [0,1]$ (figure 6.8).

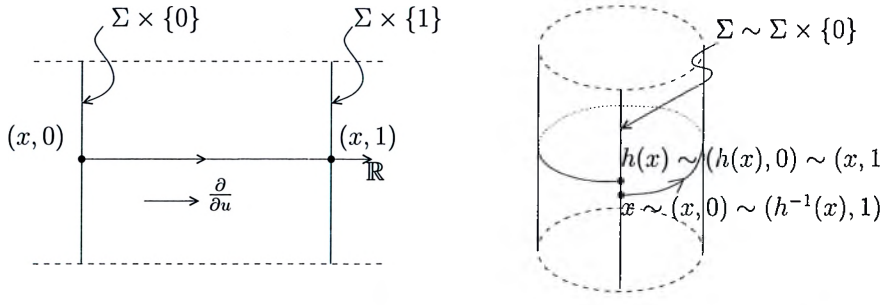


FIGURE 6.8. À gauche, $\Sigma \times [0,1]$ avec une trajectoire du champ constant $\frac{\partial}{\partial u}$. À droite, le quotient $\Sigma \times [0,1]/(x,0) \sim (h^{-1}(x),1)$.

On constate que dans cette identification de la variété M, coupée le long de Σ, avec $\Sigma \times [0,1]$, on a $(h(x),0) \sim (x,1)$ ou, d'une façon équivalente, $(x,0) \sim (h^{-1}(x),1)$. L'inconvénient de cette construction à partir de $\Sigma \times [0,1]$, est que l'on doit montrer ensuite comment munir le quotient d'une structure différentielle telle que le champ constant $\frac{\partial}{\partial u}$ induise un champ $\mathcal{C}^\infty$ sur le quotient. Ce problème est évidemment résolu par la construction du quotient de $\Sigma \times \mathbb{R}$ présentée ci-dessus.

Inversement, si h est un difféomorphisme d'une hypersurface Σ sur elle-même, on peut réaliser ce difféomorphisme comme application de Poincaré d'un champ X,

obtenu en considérant le quotient $M = (\Sigma \times \mathbb{R})/F$, avec $F(x,u) = (h^{-1}(x), u+1)$, sur lequel on considère le champ X induit par le champ constant $\frac{\partial}{\partial u}$. Cette opération s'appelle : *suspension du difféomorphisme h*. Il est à remarquer que la variété M, à difféomorphisme près (mais pas le champ X, même à équivalence $\mathcal{C}^\infty$ près !), ne dépend de h qu'*à isotopie près* (une isotopie entre deux difféomorphismes d'une variété Σ est un chemin continu qui les relie, dans l'espace des difféomorphismes). Par exemple si h est isotope à l'identité de Σ, la variété M est difféomorphe à $\Sigma \times S^1$.

Si Σ est un ouvert de R^{n-1}, la variété obtenue est alors difféomorphe à un ouvert de $\mathbb{R}^n$. On peut réaliser explicitement le plongement de $M \approx \Sigma \times S^1$ de la façon suivante. On considère $\Sigma \subset \mathbb{R}^{n-1} \approx \mathbb{R}^+ \times \mathbb{R}^{n-2}$ de coordonnées $(\rho > 0, x_3, \ldots, x_n)$ et l'on plonge $\Sigma \times S^1$ dans $\mathbb{R}^n$ en envoyant $((\rho > 0, x_3, \ldots, x_n), \theta) \in \Sigma \times S^1$ sur le point $(x_1 = \rho\cos\theta, x_2 = \rho\sin\theta, x_3, \ldots, x_n)$. (Autrement dit, on considère le produit d'une demi-droite tournant autour de l'origine de $\mathbb{R}^2$ par un facteur de $\mathbb{R}^{n-2}$.)

La variété quotient M n'est pas nécessairement un ouvert de $\mathbb{R}^n$; par exemple si le difféomorphisme h ne préserve pas l'orientation, ou bien si Σ n'est pas un ouvert de $\mathbb{R}^{n-1}$.

Considérons, par exemple, un difféomorphisme h de S^1. Si h préserve l'orientation, il est isotope à l'identité et la variété M est alors difféomorphe au tore T^2. Si, par exemple, h est une rotation $\theta \mapsto \theta + \alpha$, on retrouve les champs constants sur $\mathbb{T}^2$ déjà discutés dans la section 5.2. Rappelons que si l'angle α de la rotation est irrationnel, on obtient un flot irrationnel sur le tore, flot dont les orbites sont denses sur le tore. Inversement, on peut montrer que tout champ du tore T^2 sans orbite périodique, ni point singulier, admet une section globale difféomorphe au cercle S^1. Un tel champ est alors obtenu par la suspension d'un difféomorphisme de S^1 sans orbite périodique.

On voit donc qu'il y a identification complète entre l'étude des champs X avec une section globale Σ (considérés à équivalence $\mathcal{C}^\infty$ près) et l'étude des difféomorphismes sur Σ

$$X \text{ champ sur } M \underset{\text{suspension}}{\overset{\text{section globale } \Sigma}{\longleftrightarrow}} h \text{ difféomorphisme sur } \Sigma.$$

L'intérêt de cette identification vient du fait qu'il existe une correspondance parfaite entre *propriétés dynamiques de X et h*. On définit une dynamique par h en considérant l'itération $x \mapsto h^n(x) = h \circ \cdots \circ h(x)$. Les propriétés de cette dynamique (récurrence, stabilité, etc.) seront définies et étudiées dans le tome 2. Par exemple :

- x est périodique pour $X \iff x$ est périodique pour h ;

– les propriétés de stabilité de X correspondent biunivoquement à celles de h (ces propriétés seront définies et étudiées dans le prochain chapitre).

C'est pourquoi on considère l'étude des *systèmes dynamiques discrets* (itération d'un difféomorphisme) comme un cas particulier de l'étude des équations différentielles.

Remarque 6.6. À une section globale Σ, pour un champ X, dans une variété de dimension n, on sait maintenant associer un difféomorphisme de Poincaré de *dimension* $n-1$. Il faut distinguer ce difféomorphisme de celui obtenu en *discrétisant* le flot du champ X avec un temps $\delta > 0$, supposé fixé à une valeur petite (le *pas d'intégration*) : on obtient de cette façon une application $x \mapsto \varphi_\delta(x)$ qui est un difféomorphisme en *dimension n*. Cette discrétisation est le difféomorphisme que l'on construit numériquement pour simuler le champ : les trajectoires du difféomorphisme φ_δ sont contenues dans les orbites de X et, si le pas est très petit, elles approchent bien les orbites du champ. Par contre les propriétés de φ_δ sont très peu *génériques*, c'est-à-dire exceptionnelles en un certain sens (on verra dans le tome 2 une définition précise de la généricité). Cette non-généricité vient du fait que φ_δ possède beaucoup de courbes invariantes (les trajectoires de X, justement). Un difféomorphisme tel que φ_δ est dit parfois *intégrable* ou bien *plongeable dans un flot*. Les propriétés de φ_δ et celles de X sont très différentes en général (en ce qui concerne la stabilité structurelle, dont il sera question dans le chapitre III-4). Pour être plus précis, disons que l'on ne peut obtenir numériquement qu'une approximation ψ de la discrétisation φ_δ (plus ou moins bonne selon la méthode d'intégration utilisée). Pour ce difféomorphisme perturbé ψ, il peut apparaître des propriétés nouvelles (par exemple des *points homoclines transverses*, dont on donnera la définition dans le paragraphe III-4.7), très éloignées de celles des trajectoires du champ que l'on veut intégrer. La raison en est que l'approximation ψ a des chances d'être beaucoup plus générique que le difféomorphisme φ_δ. On reviendra sur ces questions dans le chapitre III-4.

Par contre, comme il a été dit plus haut, le flot d'un champ X avec une section globale et son application de Poincaré h ont les mêmes propriétés dynamiques.

6.3. Champs de vecteurs périodiques

Nous allons maintenant considérer un cas particulier de champ de vecteurs avec une section globale, associé aux champs de vecteurs non autonomes périodiques et complets dépendant périodiquement du temps.

Définition 6.3. Un champ de vecteurs dépendant périodiquement du temps, avec une période $T > 0$, sur une variété M est une famille $\mathcal{C}^\infty$ de champs de vecteurs de M, $X(x,t)$, telle que $X(x,t) = X(x,t+T)$. Pour chaque $(x,t) \in M \times \mathbb{R}$, on a $X(x,t) \in T_xM$, et le caractère $\mathcal{C}^\infty$ doit être vérifié pour chaque ouvert de carte U de M dans $M \times \mathbb{R}$.

Comme nous l'avons vu dans le paragraphe 1.4, l'équation différentielle du champ non autonome

$$\frac{d\varphi}{dt}(t) = X(\varphi(t), t), \tag{6.9}$$

est équivalente à l'équation différentielle autonome :

$$\begin{cases} \dfrac{d\varphi}{ds}(s) = X(\varphi(s), t(s)) \\ \dfrac{dt}{ds} = 1 \end{cases}, \tag{6.10}$$

où l'on désigne par s le temps d'intégration de (6.10).

Cette équation est celle du champ de vecteurs $\widetilde{X}(x,u) = X(x,u) + \frac{\partial}{\partial u}$, où la somme est faite dans $T_{(x,t)}M \times \mathbb{R} \sim T_xM \oplus \mathbb{R}$, pour chaque $(x,u) \in M \times \mathbb{R}$. On considère le champ $\widetilde{X}$ défini par passage au quotient sur $\widetilde{M} = (M \times \mathbb{R})/(x,u) \sim (x, u+T) = M \times S^1$ (on identifie $\mathbb{R}/\mathbb{Z}\{T\}$ avec S^1).

Rappelons la correspondance entre solution de (6.9) et trajectoire du champ $\widetilde{X}$, que nous avons décrite dans le paragraphe 1.4. Soit $(x,t_0) \in M \times \mathbb{R}$. Si $t \mapsto \varphi_t(x,t_0)$ est une solution de (6.9) vérifiant $\varphi_{t_0}(x,t_0) = x$, alors

$$\psi_s(x,t_0) = (\varphi_{t_0+s}(x,t_0), t_0 + s) \tag{6.11}$$

est la trajectoire de $\widetilde{X}$ par (x,t_0). Le théorème d'existence et d'unicité des trajectoires appliqué au champ $\widetilde{X}$ de classe $\mathcal{C}^\infty$ implique que la solution $\varphi_t(x,t_0)$ existe et est unique, du moins pour des valeurs de t proches de t_0. On voit que la solution $\varphi_t(x,t_0)$ constitue la composante spatiale de la trajectoire du flot de $\widetilde{X}$. L'autre composante du flot de $\widetilde{X}$ est la composante temporelle $t(s) = t_0 + s$. Cette composante, essentiellement triviale, traduit seulement le changement entre le temps s d'intégration de $\widetilde{X}$ et le temps initial t_0 de l'équation non autonome (6.9).

Les solutions $\varphi_t(x,t_0)$ de l'équation non autonome dépendent de deux valeurs du temps : la valeur initiale t_0 et la valeur finale t. Grâce à (6.11), la formule d'addition pour le flot de $\widetilde{X}$ (proposition 3.3) se traduit sur φ_t de la façon suivante :

Proposition 6.6. *Si t_0, t_1, t_2 sont trois temps quelconques (non nécessairement choisis en ordre croissant), on a :*

$$\varphi_{t_2}(x, t_0) = \varphi_{t_2}(\varphi_{t_1}(x, t_0), t_1), \tag{6.12}$$

chaque fois que les différents termes de la formule sont bien définis.

Démonstration. Posons $t_1 = t_0 + s$ et $t_2 = t_0 + s + \tau$. La formule d'addition pour le flot de $\widetilde{X}$ s'écrit :

$$\psi_{s+\tau}(x, t_0) = \psi_\tau(\psi_s(x, t_0)). \tag{6.13}$$

Par (6.11), le terme de gauche est égal à :

$$\psi_{s+\tau}(x, t_0) = (\varphi_{t_0+s+\tau}(x, t_0), t_0 + s + \tau). \tag{6.14}$$

Considérons maintenant le terme de droite de (6.13) :

$$\psi_\tau(\psi_s(x, t_0)) = \psi_\tau(\varphi_{t_0+s}(x, t_0), t_0 + s),$$

soit :

$$\psi_\tau(\psi_s(x, t_0)) = (\varphi_{t_0+s+\tau}(\varphi_{t_0+s}(x, t_0), t_0 + s), t_0 + s + \tau). \tag{6.15}$$

En égalant les premiers termes des membres de droite de (6.14) et (6.15), on obtient :

$$\varphi_{t_0+s+\tau}(\varphi_{t_0+s}(x, t_0), t_0 + s) = \varphi_{t_0+s+\tau}(x, t_0),$$

ce qui est la formule souhaitée. Cette formule est illustrée à la figure 6.9. ∎

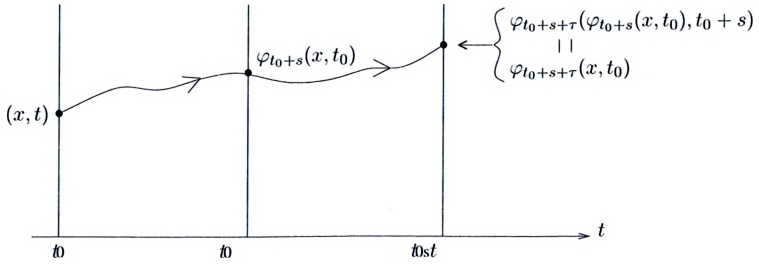


FIGURE 6.9

Définition 6.4. On dit qu'un champ non autonome défini sur $M \times \mathbb{R}$ est *complet* si, pour toute condition initiale (x, t_0), la solution $\varphi_t(x, t_0)$ est définie pour tout $t \in \mathbb{R}$.

Proposition 6.7. *Un champ périodique $X(x,t)$ est complet si et seulement si le champ autonome associé $\widetilde{X}$ est à flot complet.*

Démonstration. La preuve est immédiate d'après la formule (6.11). ∎

On suppose dorénavant que le champ périodique $X(x,t)$ est complet, ou, de façon équivalente, que le champ associé $\widetilde{X}$ est à flot complet.

Remarque 6.7.

1. Dire que le champ $\widetilde{X}$ n'est pas complet revient à dire qu'il existe une suite de temps $t_i > 0$, avec $(t_i)_i \to 0$, et des points $x_i \in M$, tels que la solution $\widetilde{\varphi}_t(x_i)$ ne soit définie que pour des $t < t_i$. Géométriquement, cela se traduit par le fait que la trajectoire de $\widetilde{X}$, par le point $(x_i, 0)$ et pour les temps positifs (c'est-à-dire le graphe de $\widetilde{\varphi}_t(x_i)$ dans $M \times \mathbb{R}^+$), est toute entière contenue dans l'anneau $M \times [0, t_i[$ et se rapproche asymptotiquement de $M \times \{t_i\}$ pour $t \to t_i$.

2. Nous avons vu que si $\widetilde{X}$ est complet, alors il admet la sous-variété $M \times \{0\}$ comme section globale. L'inverse est aussi vrai. En effet, un champ est complet si et seulement s'il peut s'intégrer sur un temps uniforme $\delta > 0$, temps qui peut être arbitrairement petit (uniforme signifie ici : indépendant de la condition initiale ; voir le lemme 3.3). Or ici, si le champ $\widetilde{X}$ admet $M \times \{0\}$ comme section globale, c'est qu'il s'intègre au moins sur un temps égal à la période T à partir des points de $M \times \{0\}$. Comme tous les points de $M \times S^1$ sont atteints à partir des points de $M \times \{0\}$, ce temps minimum d'intégration T est valide pour toutes les conditions initiales dans $M \times S^1$: le champ $\widetilde{X}$ est donc complet car on peut prendre $\delta = T$. Remarquez que l'on vient d'établir l'équivalence entre les trois affirmations suivantes : *$\widetilde{X}$ est complet* $\iff$ *$\widetilde{X}$ est complet à partir des points de $M \times \{0\}$* $\iff$ *$\widetilde{X}$ admet $M \times \{0\}$ comme section globale.*

3. Un champ *autonome* non complet, considéré comme champ périodique $\widetilde{X}$ de période T, sera aussi non complet en ce sens. Ainsi, le champ de $X(x) = x^2 \frac{\partial}{\partial x}$ intégré à la fin de la section 3.3.3 nous donne un exemple facile de cette assertion (les trajectoires de $\widetilde{X}$ sont les graphes des trajectoires de X dans le plan des (t,x) (voir la figure 3.5)). Remarquez que l'on considère ici, d'une manière qui peut sembler paradoxale, qu'*un champ autonome est un champ non autonome particulier* ; c'est un champ dépendant trivialement du temps, c'est-à-dire n'en dépendant pas ! Un champ autonome est ainsi considéré comme périodique dans le temps avec n'importe quelle période.

4. Pour que le champ $\widetilde{X}$ soit complet, il suffit que M soit compact ou bien, plus généralement, que $X(x,t)$ soit à support contenu dans une partie $K \times \mathbb{R}$, avec K partie compacte de M (comme X est périodique de période T, il suffira que $X_{|M\times[0,T]}$ soit à support dans $K \times [0,T]$).

Si $\widetilde{X}$ est complet, il est possible de préciser davantage le rapport entre le flot de $\widetilde{X}$ et les solutions $\varphi_t(x,t_0)$ de l'équation non autonome. Remarquons tout d'abord qu'en vertu de la formule (6.11), $\varphi_t : (x,t_0) \mapsto \varphi_t(x,t_0) = \varphi(t,x,t_0)$ définit une application $\mathcal{C}^\infty$ de $\mathbb{R}\times M\times\mathbb{R}$ à valeurs dans M. Toujours en vertu de (6.11), pour tout s fixé, l'application $(x,t_0) \mapsto (\varphi_{t_0+s}(x,t_0), t_0+s)$ est un difféomorpisme $\mathcal{C}^\infty$ de $M\times\mathbb{R}$. Ce difféomorphisme envoie $M\times\{t_0\}$ sur $M\times\{t_0+s\}$. Il en résulte que pour tout $(t_0,s) \in \mathbb{R}\times\mathbb{R}$, *l'application : $x \mapsto \varphi_{t_0+s}(x,t_0)$ est un difféomorphisme $\mathcal{C}^\infty$ de $M\times\{t_0\}$ sur $M\times\{t_0+s\}$.*

Nous allons dorénavant fixer $t_0 = 0$ et poser $\widetilde{\varphi}_t(x) = \varphi_t(x,0)$, autrement dit, on considère les solutions à partir du temps 0. Pour tout $t \in \mathbb{R}$, l'application $\widetilde{\varphi}_t : x \mapsto \widetilde{\varphi}_t(x)$ est un difféomorphime $\mathcal{C}^\infty$ de $M\times\{0\} \approx M$ sur $M\times\{t\} \approx M$. Cette application $\widetilde{\varphi}_t$ est appelée : *la famille de difféomorphismes $\mathcal{C}^\infty$ associée au champ de vecteurs non autonome $X(x,t)$*. Il est important de noter que cette famille $\widetilde{\varphi}_t$ *n'est pas un flot en général,* c'est-à-dire que l'on n'a pas en général l'identité $\widetilde{\varphi}_t \circ \widetilde{\varphi}_\tau \equiv \widetilde{\varphi}_{t+\tau}$, sauf si la dépendance de $X(x,t)$ en t est triviale, c'est-à-dire si le champ est autonome. Pour voir plus précisément pourquoi $\widetilde{\varphi}$ n'est pas un flot, utilisons la formule (6.12). On a, pour $s,\tau \in \mathbb{R}$:

$$\widetilde{\varphi}_{s+\tau}(x) = \varphi_{s+\tau}(\widetilde{\varphi}_\tau(x),\tau). \tag{6.16}$$

Or, en général $\varphi_{\tau+s}(z,\tau) \neq \widetilde{\varphi}_s(z)$ si $z \in M, \tau \neq 0$, si le champ X n'est pas autonome : autrement dit, la solution $s \mapsto \varphi_{s+\tau}(z,\tau)$, à partir du temps τ, ne se ramène pas en général à la solution $\widetilde{\varphi}_s(z)$ à partir du temps 0, car, précisément *le champ X a pu changer de valeurs entre les instants 0 et τ.*

Voici un petit exemple pour illustrer ce qui précède. Considérons le champ 2π-périodique de $\mathbb{R}$ suivant : $X(x,t) = \sin t\frac{\partial}{\partial x}$. Pour ce champ, on a $\varphi_t(x,t_0) = \cos t - \cos t_0 + x$, d'où : $\widetilde{\varphi}_t(x) = \cos t - 1 + x$ (d'après ce qui précède, on pose $\widetilde{\varphi}_t(x) = \varphi_t(x,0)$). Si l'on prend deux temps t,τ quelconques, on en déduit que $\widetilde{\varphi}_{t+\tau}(x) = \cos(t+\tau) - 1 + x$, alors que $\widetilde{\varphi}_t \circ \widetilde{\varphi}_\tau(x) = \cos t - 1 + \widetilde{\varphi}_\tau(x) = \cos t + \cos\tau - 2 + x$. Comme en général $\cos(t+\tau) \neq \cos t + \cos\tau - 1$, on aura aussi en général que $\widetilde{\varphi}_{t+\tau}(x) \neq \widetilde{\varphi}_t \circ \widetilde{\varphi}_\tau(x)$.

Inversement, toute famille $\mathcal{C}^\infty$ de difféomorphismes $\widetilde{\varphi}_t$ de M, avec $t \in \mathbb{R}$, est le flot d'un champ non autonome et complet de M. Explicitement, le champ est donné par la formule :

$$X(t,\widetilde{\varphi}_t(x)) = \frac{\partial\widetilde{\varphi}_t}{\partial t}(t,x). \tag{6.17}$$

Le champ non autonome $X(t,x)$ s'appelle le *générateur infinitésimal* de la famille $\widetilde{\varphi}_t(x)$. Dans beaucoup de problèmes où l'on veut construire une famille de difféomorphismes, il est plus facile de construire le générateur infinitésimal associé, puis d'obtenir la famille de difféomorphismes par intégration (voir [12] pour plus de détails et des exemples d'applications). Ici, de plus, le champ est supposé T-périodique : cela se traduit par l'identité :

$$\widetilde{\varphi}_{t+T}(x) \equiv \widetilde{\varphi}(t, \widetilde{\varphi}_T(x)) \text{ pour tout } (x,t) \in M \times \mathbb{R}. \tag{6.18}$$

Remarque 6.8. En chaque point $x \in M$, le champ $X(x,t)$ fluctue périodiquement dans le temps. L'allure des solutions dans M est donc fort différente de celle d'un champ de vecteurs autonome. Si l'on pense à l'écoulement d'un fluide dans une variété M, régi par une équation différentielle autonome, on observe des lignes d'écoulement immobiles, figées dans le temps : ces lignes sont les trajectoires du champ. On dit que l'écoulement est *laminaire* (pensez au cours d'une rivière calme). Si le fluide est régi par un champ non autonome, les lignes d'écoulement se recoupent sans cesse, ce qui se manifeste en chaque point x par une fluctuation de la direction de l'écoulement en fonction du temps. On dit que l'écoulement est *turbulent*. Pensez par exemple au cours d'une rivière à proximité d'une arche de pont, ou bien à l'écoulement du gaz dans une flamme turbulente, lorsque l'écoulement est périodique dans le temps.

Les courbes $t \mapsto \widetilde{\varphi}_t(x)$ ne sont donc pas en général les trajectoires d'un champ de vecteurs de M. Il peut se faire par exemple que ces courbes aient des auto-intersections, c'est-à-dire que l'on ait $x \in M$ et $t_1 \neq t_2$ tels que $x_1 = \widetilde{\varphi}_{t_1}(x) = \widetilde{\varphi}_{t_2}(x) = x_2$ avec des vecteurs dérivés $X(x_1,t_1), X(x_2,t_2)$ non colinéaires (si M est de dimension 2, en $x_1 = x_2$ la courbe a une auto-intersection transverse). Le champ X aura fluctué entre les deux passages successifs au même point $x_1 = x_2$. On revient évidemment à des trajectoires d'un champ de vecteurs (le champ $\widetilde{X}$), si l'on considère les graphes $t \mapsto (\widetilde{\varphi}_t(x), t)$ dans $M \times S^1$. Les courbes $t \mapsto \widetilde{\varphi}_t(x)$ peuvent être simplement interprétées comme les projections des orbites de $\widetilde{X}$ sur le facteur M du produit $M \times S^1$.

La sous-variété $\Sigma = M \times \{0\}$ (et plus généralement chaque $M \times \{t\}$) est une section globale du champ $\widetilde{X}$, avec un temps de retour constant égal à T, ce qui nous ramène dans le cadre du paragraphe précédent. Le difféomorphisme de M : $x \mapsto h(x) = \widetilde{\varphi}(T,x)$ est l'application de Poincaré h associée à $\widetilde{X}$, car le premier temps de retour est égal à T. L'application de Poincaré du champ autonome $\widetilde{X}$, associé au champ périodique X, s'obtient donc par *intégration de l'équation différentielle non autonome (6.9) sur une période, à partir du temps* $t = 0$. On a déjà dit que cette application permet d'étudier toutes les propriétés qualitatives du champ ne dépendant que des orbites. Ici, de plus, comme le retour a

lieu à intervalles de temps fixes (la période), on a aussi accès à certaines propriétés dépendantes du temps, par exemple les périodes d'orbites périodiques de $\widetilde{X}$: si l'application de Poincaré a un point fixe périodique $x \in M$ de période $k \in \mathbb{N}$, le champ $\widetilde{X}$ a une orbite périodique par chaque point (x,t), de période kT, et inversement.

Remarque 6.9.

1. Le calcul de l'application de Poincaré sur la section $M \times \{0\}$, qui permet une étude qualitative complète du champ $\widetilde{X}$ à équivalence près, ne demande que la connaissance des solutions de (6.9) à partir du temps $t = 0$, solutions notées $\widetilde{\varphi}_t(x)$. Remarquez que si l'on change la section $M \times \{0\}$ par la section $M \times \{t_0\}$, la nouvelle application de Poincaré est donnée par $h_1 : x \mapsto \varphi_{t_0+T}(x, t_0)$. Ce difféomorphisme h_1 est conjugué à h et a donc les mêmes propriétés dynamiques, comme il a été dit dans le paragraphe précédent.

   Rappelons, d'un autre côté, que la connaissance complète des solutions $\varphi_t(x, t_0)$ de (6.9) ne se réduit pas à la connaissance de la famille de difféomorphismes $\widetilde{\varphi}_t(x)$. Ce paradoxe vient du fait que toutes les propriétés du flot de $\widetilde{X}$ ne se traduisent pas à travers les applications de Poincaré (par exemple, la valeur maximale atteinte par $||\widetilde{\varphi}_t(x)||$ le long d'une orbite périodique, si $M = U$ ouvert de $\mathbb{R}^n$ et si $||\cdot||$ désigne la norme euclidienne).

2. Le champ de vecteurs $\widetilde{X}$ avec une section globale, que l'on considère ici, semble différer du champ qui serait obtenu par suspension de l'application de Poincaré, par le fait que $\widetilde{X}$ a une composante non triviale sur M, alors que le champ suspendu est le quotient du champ trivial (de composante nulle le long de M). Évidemment, ceci est illusoire, car les deux champs sont conjugués par un difféomorphisme et leur étude est complétement équivalente. Il faut retenir seulement que cette étude est ramenée à celle de l'application de Poincaré, qui se calcule par intégration du champ X sur une période, et que, malheureusement, il n'y a pas, en général, de méthode explicite pour obtenir cette application de Poincaré.

Exemple de champ périodique. Soit l'équation $\ddot{x} + g(x) = \xi(t)$, où $g(x)$ est une fonction de $x \in V$ (où $V = \mathbb{R}$ ou bien $V = S^1$), et $\xi(t)$ est une fonction périodique de période T. Si $m = (x, y) \in M = V \times \mathbb{R}$, le champ de vecteurs s'écrit :

$$X(t,m) = y\frac{\partial}{\partial x} + (-g(x) + \xi(t))\frac{\partial}{\partial y}, \tag{6.19}$$

et l'équation du second ordre s'écrit, de façon équivalente, comme l'équation différentielle du champ $X(t,m)$:

$$\begin{cases} \dot{x} = y \\ \dot{y} = -g(x) + \xi(t) \end{cases}.$$

Si $g(x) = \sin x$, on a l'équation d'un pendule soumis à une force extérieure périodique (par exemple, le pendule est constitué par une bille en fer attachée à un fil rigide et la bille oscille dans un plan vertical dans un champ magnétique périodique). On rappelle que l'intégration de l'équation du pendule non excité est très simple, puisque les orbites sont contenues dans les lignes de niveau de la fonction énergie $E(x,y) = \frac{1}{2}y^2 - \cos x$. Par contre, comme nous l'avons dit plus haut, l'étude du pendule excité va se ramener à celle de l'application de Poincaré, qui est une application du plan dans lui-même. Le comportement dynamique de cette application peut être très compliqué. Par exemple, elle peut avoir des point homoclines transverses et, en conséquence, une sous-dynamique de type « fer à cheval » que l'on étudiera dans le chapitre III-4. En particulier ce comportement peut exhiber différentes formes de chaos.

7

STABILITÉ DES TRAJECTOIRES

Dans ce chapitre nous envisageons deux notions de stabilité (*stabilité au sens de Lyapounov*, *stabilité au sens asymptotique*). Nous allons considérer ces notions pour les deux types de récurrence triviale d'un champ de vecteurs X : les points singuliers et les orbites périodiques, et aussi pour un point fixe d'un difféomorphisme. On veut que pour une petite perturbation *des conditions initiales*, le mouvement soit peu perturbé *à long terme*.

Dans le chapitre III-4 où parlera de la *stabilité structurelle* du champ, c'est-à-dire que l'on considérera *une perturbation de l'équation elle-même*. Un champ de vecteurs structurellement stable est un champ dont le *portrait de phase* reste *équivalent* à lui-même après une petite perturbation.

Ces notions de stabilité sont très différentes. Il est possible qu'une propriété soit instable au premier sens mais reste stable par perturbation de l'équation ; par exemple dans une dynamique chaotique, il peut se faire qu'aucune trajectoire ne soit stable alors que la dynamique peut être stable du point de vue structurel.

Nous supposerons que X est un champ $\mathcal{C}^\infty$ sur une variété M, quoique une grande partie de l'étude puisse se faire pour les champs de vecteurs de classe $\mathcal{C}^1$. Rappelons que, pour nous, champ signifie champ autonome. Un champ non autonome périodique X ne sera considéré que par l'intermédiaire de son champ autonome associé $\widetilde{X}$, comme il a été expliqué dans la section 1.4. Un tel champ $\widetilde{X}$ pourra avoir des orbites périodiques, mais n'aura pas de points singuliers. Si un champ non autonome sur une variété M est non périodique, on ne pourra lui associer qu'un champ autonome sur $M \times \mathbb{R}$ (sans possibilité de passer au quotient), et ce champ n'aura ni point singulier ni orbite périodique, ce qui implique que les résultats du présent chapitre lui seront inapplicables. Les notions particulières de stabilité, que l'on pourrait développer pour un tel champ, sortent du cadre de cet ouvrage (on trouvera une étude de la stabilité dans le cas non autonome dans [14, 24]).

Nous supposerons aussi que le champ est à flot complet, car nous considérons, dans la suite, la limite des trajectoires lorsque le temps t tend vers l'infini. Rappelons qu'un champ sur une variété compacte est toujours complet, et qu'un champ sur $\mathbb{R}^n$, par exemple, peut être rendu complet par multiplication par une fonction $\mathcal{C}^\infty$ positive, c'est-à-dire à équivalence $\mathcal{C}^\infty$ près. Si le champ provient d'un champ non-autonome de $\mathbb{R}^{n-1}$, cette multiplication va modifier la dépendance particulière du champ par rapport au temps, puisque l'équation $\dot{t} = 1$ va être remplacée par l'équation $\dot{t} = f(x,t)$, où f est la fonction multiplicative. Certaines questions dépendant explicitement de la variable temps ne peuvent plus alors être considérées, mais leur étude sort du cadre de cet ouvrage.

Dans le cas des points singuliers d'un champ de vecteurs X, les notions de stabilité que nous allons introduire sont locales, c'est-à-dire définissables dans une carte contenant le point singulier x, et indépendantes du choix de la carte, comme on le vérifiera facilement. Autrement dit, dans le cas des points singuliers, on peut toujours considérer que $M = \mathbb{R}^n$ et que la distance choisie est associée à la norme euclidienne $||\cdot||$ de $\mathbb{R}^n$, on peut donc prendre $\text{dist}(x,y) = ||x-y||$.

Par contre, une orbite périodique d'un champ de vecteurs sur une variété M n'est pas nécessairement contenue dans un seul domaine de carte. L'utilisation d'une distance sur M, telle que nous l'avons introduite dans le chapitre I-2, est alors très utile pour définir les notions de stabilité relatives à une telle orbite. Cependant, nous verrons que ces notions se ramènent aux notions relatives à une application de Poincaré associée à l'orbite périodique. On est alors ramené à une étude locale : celle d'un difféomorphisme d'une section Σ au voisinage d'un point fixe. Pour cette étude, on peut supposer à nouveau que le difféomorphisme est défini sur un ouvert de $\mathbb{R}^{n-1}$ et travailler avec la norme euclidienne. C'est évidemment ce que l'on fera.

7.1. Stabilité d'un point singulier de champ de vecteurs

On a vu que si $p \in M$ est un point singulier, c'est-à-dire si $X(p) = 0$, alors $\varphi(t,p) \equiv p$: on reste au point p. C'est pourquoi on appelle les points singuliers : *points d'équilibre* ou *points stationnaires*. Si, maintenant, x est une position voisine, on voudrait savoir si $\varphi(t,x)$ reste voisin de p, pour $t \to +\infty$.

L'étude de la stabilité est locale. Dire que l'on s'intéresse aux propriétés locales signifie, précisément, que l'on s'intéresse aux propriétés indépendantes du choix de la carte et de sa taille. Nous allons choisir une carte (U,ψ) telle que $p \in U$, avec $\psi(p) = 0 \in \mathbb{R}^n$. Dans U, le champ de vecteurs X est représenté par un champ de vecteurs encore noté X, défini sur le voisinage ouvert $\Omega = \psi(U)$ de $0 \in \mathbb{R}^n$. Pour ce champ, on a $X(0) = 0$. Le lecteur pourra aisément se convaincre que les définitions et l'étude peuvent se transcrire en termes intrinsèques (indépendants

du choix de la carte choisie au voisinage de $p \in M$, ainsi que de la distance utilisée pour définir la topologie).

7.1.1. Différents types de stabilité

Stabilité au sens de Lyapounov

Définition 7.1. Soit X un champ de vecteurs $\mathcal{C}^\infty$ sur le voisinage Ω de $\{0\}$ dans $\mathbb{R}^n$. On suppose que $X(0) = 0$. On dit que $\{0\}$ est un point singulier stable (au sens de Lyapounov), pour les temps positifs, si l'on a la condition suivante :

$$\forall \varepsilon > 0 \text{ assez petit, } \exists \delta(\varepsilon) > 0 \text{ tel que, } \forall x \in \Omega \text{ avec } ||x|| \leq \delta(\varepsilon),$$

on a

$$\varphi(t,x) \in \Omega \text{ et } ||\varphi(t,x)|| \leq \varepsilon\,,\ \forall t \geq 0.$$

L'orbite $\{\varphi(t,x)\,|\,t \geq 0\}$ reste donc dans la boule $B_\varepsilon(0)$, pourvu que l'on parte d'un point x situé à moins de $\delta(\varepsilon)$ de l'origine. C'est une propriété de *confinement de la trajectoire*. Évidemment, on doit avoir $\delta(\varepsilon) \leq \varepsilon$, car la condition $||\varphi(t,x)|| \leq \varepsilon\,,\ \forall t \geq 0$ doit être vrai en particulier pour $t = 0$ et que $\varphi(0,x) = x$.

Comme il a été dit plus haut, la notion précédente est invariante par changement de coordonnées local en $\{0\} \in \mathbb{R}^n$, et donc permet de définir la stabilité au sens de Lyapounov d'un point singulier p d'un champ sur une variété quelconque M.

Pour ce faire, on rappelle que $\{W_i\}_{i\in I}$ est un système fondamental de voisinages, si chaque W_i est un voisinage de x, et si, pour tout voisinage W de x, il existe un W_i tel que $W_i \subset W$. On rappelle aussi que W est invariant par le flot φ_t pour les temps positifs si $\varphi(t,W) \subset W$ pour tout $t \geq 0$ (voir définition 5.2).

Voici une formulation purement topologique et intrinsèque de cette notion (c'est-à-dire n'utilisant pas explicitement la distance) :

Lemme 7.1. *Soit X un champ de vecteur sur une variété M. Le point singulier p est stable au sens de Lyapounov si et seulement s'il possède un* système fondamental de voisinages invariants par φ_t pour les temps positifs.

Démonstration. Par choix d'une carte au voisinage de p, tout revient à démontrer le lemme pour un champ X défini sur un voisinage Ω de l'origine avec $X(0) = 0$. Supposons que le point 0 soit stable au sens de Lyapounov. Pour tout $\varepsilon > 0$ assez petit, l'ensemble $W_{\delta(\epsilon)} = \bigcup_{t\geq 0} \varphi(t, B_{\delta(\epsilon)}(x))$ est invariant par le flot pour tout

$t \geq 0$. En effet, grâce à la formule de translation (3.13) de la proposition 3.4, pour tout $\tau \geq 0$ on a

$$\varphi(\tau, W_{\delta(\epsilon)}) = \bigcup_{t \geq 0} \varphi(t + \tau, B_{\delta(\epsilon)}(x)) \subseteq W_{\delta(\epsilon)}.$$

D'autre part, $W_{\delta(\epsilon)}$ est un voisinage car il contient $B_{\delta(\epsilon)}(x) = \varphi(0, B_{\delta(\epsilon)}(x))$. Comme par hypothèse $\varphi(t, B_{\delta(\epsilon)}(x)) \subseteq B_\varepsilon(x)$ pour tout $t \geq 0$, on a aussi $W_{\delta(\epsilon)} \subseteq B_\varepsilon(x)$, ce qui montre que la collection $\{W_{\delta(\epsilon)}\}_{\varepsilon > 0}$ est un système fondamental de voisinages.

Inversement, supposons que l'on a un système fondamental $\mathcal{W}$ de voisinages invariants. Soit $\varepsilon > 0$ quelconque. Il existe un élément $W \in \mathcal{W}$ tel que $W \subseteq B_\varepsilon(x)$. Comme W est un voisinage de x, il existe $\delta(\varepsilon) > 0$ tel que $B_{\delta(\epsilon)}(x) \subseteq W$. Maintenant, si $y \in B_{\delta(\epsilon)}(x)$, on a $y \in W$, et donc $\varphi(t, y) \in W \subseteq B_\varepsilon(x)$ pour tout $t \geq 0$. Cet argument, appliqué à tout $\varepsilon > 0$, montre que x est stable au sens de Lyapounov. ∎

Remarque 7.1. Il existe des notions similaires pour $t \to -\infty$ ou pour $|t| \to +\infty$. Cette dernière forme de la stabilité au sens de Lyapounov est clairement indépendante du sens de parcours du temps. Elle est équivalente à l'existence d'un système fondamental de voisinages du point singulier, invariants par le flot (*pour tous les temps* et pas seulement pour les temps positifs). C'est cette notion qui est naturelle en mécanique pour les systèmes conservatifs.

Stabilité asymptotique

La stabilité au sens de Lyapounov d'un point singulier p implique le confinement des orbites issues de points voisins au voisinage de p, mais pas que les trajectoires issues de ces points tendent vers p pour $t \to +\infty$. Nous allons introduire maintenant une notion plus forte qui implique cette convergence. Comme la notion est locale, nous allons de nouveau supposer que X est un champ sur un ouvert Ω, voisinage de l'origine $0 \in \mathbb{R}^n$, et que $X(0) = 0$.

Définition 7.2 **(Stabilité asymptotique d'un point singulier).** Soit X un champ de vecteur sur un ouvert Ω voisinage de l'origine $0 \in \mathbb{R}^n$, on suppose (toujours) $X(0) = 0$. On dit que $\{0\}$ est *asymptotiquement stable* s'il existe $\rho_0 > 0$ tel que, pour tout $\varepsilon > 0$, il existe un $t(\varepsilon)$ tel que :

$$||x|| \leq \rho_0 \text{ et } t \geq t(\varepsilon) \Longrightarrow ||\varphi(t, x)|| \leq \varepsilon.$$

***Remarque* 7.2.** Pour tout $x \in \overline{B_{\rho_0}(0)}$ (boule fermée) on a donc : $\varphi(t,x) \to 0$. De plus, d'après la définition 7.2, cette convergence est *uniforme* dans $\overline{B_{\rho_0}(0)}$ (le $t(\epsilon)$ ne dépend pas de x). En fait un argument simple, utilisant la compacité de $\overline{B_{\rho_0}(0)}$ comme dans le lemme 7.2 ci-dessous, montre que l'uniformité suit de la propriété de convergence pour chaque point : $\varphi(t,x) \to 0$ pour tout $x \in \overline{B_{\rho_0}(0)}$.

On dira, évidemment, qu'un point singulier p d'un champ de vecteur X sur une variété M est asymptotiquement stable, si son représentant dans une carte (U,ψ) avec $\psi(p) = 0$, et donc dans toute telle carte, est stable au sens de la définition 7.2.

Relation entre les deux formes de stabilité

La stabilité asymptotique implique la stabilité au sens de Lyapounov :

***Lemme* 7.2.** *Soit un point singulier p d'un champ X. Alors, si p est asymptotiquement stable, il est aussi stable au sens de Lyapounov.*

Démonstration. On peut supposer que le champ est défini sur un ouvert Ω de $\mathbb{R}^n$, comme plus haut, avec $p = 0$. Soit $\varepsilon > 0$. Par hypothèse, il existe un temps $t(\varepsilon)$ et un $\rho_0 > 0$ tels que, pour tout $x \in \overline{B_{\rho_0}(0)}$ et tout $t \geq t(\varepsilon)$, on ait $\|\varphi(t,0)\| \leq \varepsilon$, ce qui peut se traduire par :

$$\varphi([t(\varepsilon), +\infty) \times \overline{B_{\rho_0}(0)}) \subset B_\varepsilon(0). \tag{7.1}$$

Nous allons maintenant nous occuper des points $\varphi(t,x)$, pour $t \in [0, t(\varepsilon)]$. En fait, $\varepsilon > 0$ *étant fixé,* on va montrer qu'il existe $\delta(\varepsilon)$ assez petit pour que $\|\varphi(t,x)\| \leq \varepsilon$ dès que : $\|x\| \leq \delta(\varepsilon)$ et $t \in [0, t(\varepsilon)]$. Ceci se traduit par :

$$\varphi([0, t(\varepsilon)] \times B_{\delta(\varepsilon)}(0)) \subset B_\varepsilon(0). \tag{7.2}$$

Pour ce faire, on utilise la *compacité* de l'intervalle $[0, t(\varepsilon)]$. Comme l'origine est un point singulier, on a $\varphi(t,0) = 0$ pour chaque $t \in \mathbb{R}$. Soit un $t \in [0, t(\varepsilon)]$ quelconque. L'application φ étant continue en (t,x), il existe $\alpha_t > 0$ et $\delta_t > 0$ tels que $\varphi(]t - \alpha_t, t + \alpha_t[\times B_{\delta_t}(0)) \subset B_\varepsilon(0)$. *Comme l'intervalle* $[0, t(\varepsilon)]$ *est compact,* on peut extraire un sous-recouvrement fini du recouvrement ouvert $\{]t - \alpha_t, t + \alpha_t[\}_{t \in [0, t(\varepsilon)]}$. Autrement dit, il existe : $t_1, \ldots, t_k \in [0, t(\varepsilon)]$ tels que $[0, t(\varepsilon)] \subset \cup_{i=1}^k]t_i - \alpha_{t_i}, t_i + \alpha_{t_i}[$. Si maintenant on pose

$$\delta(\varepsilon) = \mathrm{Inf}\{\delta_{t_i} \mid i = 1, \ldots, k\},$$

on a $\delta(\varepsilon) > 0$ et on obtient l'inclusion (7.2) souhaitée (figure 7.1).

Le résultat suit maintenant des deux inclusions (7.1) et (7.2) en prenant $\delta(\varepsilon) = \min(\rho_0, \delta(\varepsilon))$. ∎

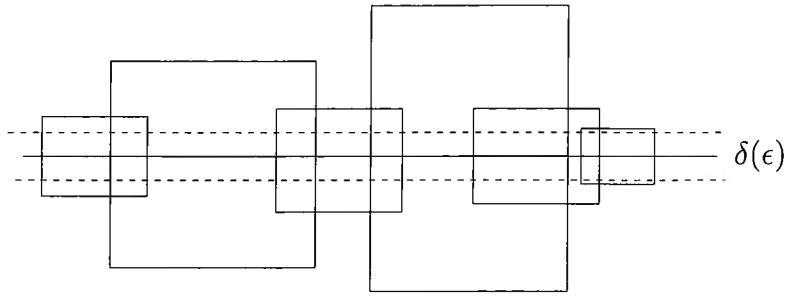


FIGURE 7.1. Chaque rectangle représente un domaine du type $]t_i - \alpha_{t_i}, t_i + \alpha_{t_i}[\times B_{\delta_{t_i}}(0)$.

Le lemme 7.2 n'admet pas de réciproque : la stabilité de Lyapounov n'entraîne pas la stabilité asymptotique. Considérons par exemple le champ linéaire à coefficients constants et réels :

$$X = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} + \alpha\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)$$

les deux premiers termes correspondent à des termes de rotation, et les deux derniers à des termes de contraction, ou de dilatation, selon le signe de α. L'origine $0 \in \mathbb{R}^2$ est l'unique point singulier. Le système associé à ce champ de vecteurs est de la forme $\frac{dY}{dt} = AY$, où A est la matrice de *similitude* : $\begin{pmatrix} \alpha & 1 \\ -1 & \alpha \end{pmatrix}$.

On a expliqué, dans le chapitre 2, comment intégrer simplement ce système. Il suffit pour cela d'introduire la variable complexe $z = x + iy$. Le système se réduit à l'équation différentielle linéaire de $\mathbb{C}$:

$$\dot{z} = (\alpha - i)z,$$

qui s'intègre par la formule :

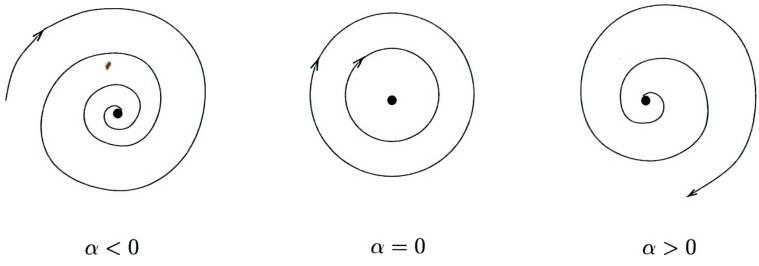
$$\varphi(t, z) = ze^{(\alpha - i)t}.$$

Remarquez que $|\varphi(t, z)| = |z|e^{\alpha t}$. Il en résulte immédiatement la discussion suivante de la stabilité de l'origine :

1. Si $\alpha = 0$, les orbites sont des cercles concentriques à l'origine (figure 7.2). On dit que l'origine est un *centre*. On a évidemment la stabilité de Lyapounov sans la stabilité asymptotique (les trajectoires ne tendent pas vers l'origine lorsque $t \to +\infty$).

2. Si $\alpha < 0$, l'origine est un *foyer attractif* : la trajectoire par $z \neq 0 \in \mathbb{C}$, quel que soit ce point, tend vers l'origine, lorsque $t \to +\infty$, en spiralant (figure 7.2). On a la stabilité asymptotique.

3. Si $\alpha > 0$, l'origine est un *foyer répulsif* : la trajectoire par $z \neq 0 \in \mathbb{C}$, quel que soit ce point, s'écarte, en tournant en forme de spirale, de l'origine, lorsque $t \to +\infty$ (figure 7.2). On peut dire que l'origine est asymptotiquement stable pour $t \to -\infty$ (ou bien que l'origine est asymptotiquement stable pour le champ $-X$).



$\alpha < 0$ $\alpha = 0$ $\alpha > 0$

FIGURE 7.2

7.1.2. Théorèmes de stabilité

Fonction de Lyapounov

Soit p un point singulier d'un champ de vecteurs X, défini sur une variété M. La définition de la fonction de Lyapounov fait appel à la notion de dérivée de Lie : $X \cdot f = L_X f$, d'une fonction f par un champ de vecteurs X (voir la définition 1.1).

Définition 7.3. Une fonction de Lyapounov F, pour le champ X, au voisinage du point singulier p, est une fonction différentiable définie sur un voisinage ouvert U de p, telle que

1. Le point p est un minimum absolu strict de F : $F(x) > F(p)$ pour tout $x \in U - \{p\}$.

2. La dérivée de Lie : $X \cdot F = L_X F$ est inférieure ou égale à 0 : $X \cdot F(x) \leq 0$ en tout point $x \in U$.

Remarque 7.3. 1. Si F est une fonction de Lyapounov, cette fonction décroît le long de chaque trajectoire de X dans U. Cela suit de la formule de dérivation : $X \cdot F(\varphi(t,x)) = \frac{\partial}{\partial t}(F \circ \varphi)(t,x)$ établie au lemme 1.1. Il en résulte,

d'après l'item 2 de la définition 7.3, que pour tout $x \in U$, la fonction à une variable :

$$t \mapsto (F \circ \varphi)(t, x)$$

est à dérivée négative et est donc décroissante.

2. Soit F une fonction $\mathcal{C}^\infty$ sur un ouvert de $\mathbb{R}^n$. Son gradient ∇F admet la fonction $-F$ comme fonction de Lyapounov, car

$$\nabla F \cdot (-F) = -||\nabla F||^2.$$

3. Il n'y a pas de méthode générale pour exhiber une fonction de Lyapounov. C'est pour chaque problème une question de savoir-faire. Pour les systèmes mécaniques ou électriques, l'énergie est souvent une fonction de Lyapounov. Par exemple, dans un système mécanique non conservatif, tel un pendule avec frottement, l'énergie totale du système est une fonction de Lyapounov, au voisinage de tout minimum local strict du potentiel.

Les théorèmes suivants sont dûs à Lyapounov :

Théorème 7.1. *Supposons qu'il existe une fonction de Lyapounov F, pour le champ X, au voisinage du point singulier $p \in \Omega$. Alors le point p est stable au sens de Lyapounov.*

Démonstration. Si F est une fonction de Lyapounov, il en est de même pour $F - F(p)$. On peut donc supposer que la fonction de Lyapounov F vérifie que $F(p) = 0$ et $F(x) > 0$ si $x \in U - \{p\}$. On peut aussi supposer que F est définie sur le domaine d'une carte (U, ψ) contenant le point p, avec $\Omega = \psi(U) = B_{\rho_0}(0)$, pour un certain $\rho_0 > 0$. Sur la boule ouverte B_{ρ_0}, le champ représentatif de X admet alors la fonction $F \circ \psi$ comme fonction de Lyapounov. Comme le champ X est défini sur l'ouvert Ω, pour ne pas alourdir les notations, on peut supposer que la fonction F est aussi définie sur Ω et qu'elle admet en $\psi(p) = 0 \in \Omega$ un minimum absolu strict, avec de plus $X \cdot F(x) \leq 0$ pour tout $x \in \Omega$. Pour tout $\varepsilon : 0 < \varepsilon < \rho_0$, on pose

$$\mu(\varepsilon) = \mathrm{Inf}\{F(x) \mid ||x|| = \varepsilon\}.$$

Comme $F(x) > 0$ pour $x \neq 0$, on a $\mu(\varepsilon) > 0$ pour $0 < \varepsilon < \rho_0$. Ceci est une conséquence du théorème du minimum atteint, appliqué à la fonction *continue* F restreinte au *compact* : $\partial B_\varepsilon(0)) = \{||x|| = \varepsilon\}$ (F est continue car elle est supposée différentiable).

Pour tout $\varepsilon : 0 < \varepsilon < \rho_0$, on considère l'ensemble ouvert

$$W_\varepsilon = \left\{ x \in B_\varepsilon(0) \mid F(x) < \frac{\mu(\varepsilon)}{2} \right\}.$$

Comme $F(0) = 0$, cet ensemble contient l'origine et est donc un voisinage ouvert de l'origine. Choisissons un $\delta(\varepsilon)$, $0 < \delta(\varepsilon) < \varepsilon$, tel que $B_{\delta(\varepsilon)}(0) \subset W_\varepsilon$. Alors, si $||x|| < \delta(\varepsilon)$ on aura que $||\varphi(t,x)|| < \varepsilon$ pour tout $t \leq 0$, ce qui est le résultat souhaité. En effet, comme on l'a vu à la remarque 7.3, la valeur de F diminue le long de la trajectoire de x. On a donc que, pour tout $t \leq 0$, $F(\varphi(t,x)) \leq F(x) < \frac{\mu(\varepsilon)}{2}$. Comme la valeur de F, en tout point du bord de la boule $B_\varepsilon(0)$, est supérieure ou égale à $\mu(\varepsilon)$, la trajectoire partant du point x situé dans la boule ne peut pas atteindre ce bord. Il s'en suit bien que $\varphi(t,x) \in B_\varepsilon(0)$ pour tout $t \geq 0$. ∎

Théorème 7.2. *Sous les hypothèses du théorème précédent, supposons que F soit de classe $\mathcal{C}^1$ et que l'on ait : $X \cdot F(x) < 0$ pour tout $x \in U - \{p\}$. Alors le point p est asymptotiquement stable.*

Démonstration. Comme dans la preuve du théorème précédent, on peut supposer que le champ X est donné sur la boule ouverte $\Omega = B_{\rho_0}(0)$ et que le point singulier est situé à l'origine. La fonction de Lyapounov est $\mathcal{C}^1$, admet $F(0) = 0$ comme valeur minimale stricte et vérifie $X \cdot F(x) < 0$ pour tout $x \in U - \{p\}$. Choisissons un ρ quelconque, tel que $0 < \rho < \rho_0$, que nous supposerons dorénavant fixé. On considère le voisinage ouvert de l'origine : $W_\rho \subset \Omega$, défini comme l'ouvert W_ε dans la preuve du théorème 7.1.

D'après la remarque 7.2, il suffit de prouver que pour tout $x \in W_\rho - \{0\}$, $\varphi(t,x) \to 0$ pour $t \to +\infty$. Cette demi-orbite ne passant pas par 0 (qui est un point singulier), on a l'inégalité stricte : $X \cdot F(\varphi(t,x)) < 0$ pour tout $t \geq 0$, et donc, encore par la remarque 7.3, la fonction $t \mapsto F(\varphi(t,x))$ est strictement décroissante. D'autre part, toute la demi-trajectoire positive γ_x^+ est contenue dans le compact $\overline{B_\rho(0)} \subset \Omega$. L'ensemble ω-limite de x, noté ω_x, est donc un compact non vide dans $\overline{B_\rho(0)}$, car il est un ensemble fermé dans le compact $\overline{B_\rho(0)}$ (voir proposition 5.2). Nous allons montrer que cet ensemble ne peut pas contenir de point de $\Omega - \{0\}$. Comme il est non vide, il en résultera que $\omega_x = \{0\}$ et donc que $\varphi(t,x) \to 0$ lorsque $t \to +\infty$, d'après la proposition 5.5.

Supposons donc, par l'absurde, qu'il existe un point $x_0 \in \omega_x - \{0\}$. En ce point x_0 on a $X \cdot F(x_0) \neq 0$. Il en résulte que $dF(x_0) \neq 0$ et $X(x_0) \neq 0$. Par le théorème des fonctions implicites, l'équation $\{F(x) = F(x_0)\}$ définit localement une hypersurface de classe $\mathcal{C}^1$, passant par le point x_0, et transverse en ce point au vecteur $X(x_0)$, comme il suit aussi de la condition : $X \cdot F(x_0) \neq 0$. On a vu dans la proposition 4.4 que l'on pouvait choisir une section locale Σ par le point x_0, contenue dans cette hypersurface, puis construire un voisinage tubulaire B tel que $B \cap \{F(x) = F(x_0)\} = \Sigma$ (voir proposition 4.5 où le rôle de M est joué, ici, par l'hypersurface $\{x | F(x) = F(x_0)\}$ fermé de $\mathbb{R}^n$; comme la fonction F et donc Σ sont

seulement de classe $\mathcal{C}^1$, il en est de même pour le difféomorphisme de trivialisation du champ X dans B, mais cela n'a pas d'importance pour la suite de la preuve). Le point x_0 appartenant à ω_x, il existe une suite strictement croissante de temps : $(t_i)_i \to +\infty$, telle que $\varphi(t_i, x)$ appartienne à B pour tout i et $(\varphi(t_i, x))_i \to x_0$. De plus, comme on l'a fait par exemple dans la preuve du théorème 5.2 de Poincaré-Bendixson, on peut choisir les t_i pour que $\varphi(t_i, x) \in \Sigma$ pour tout i. On a maintenant une contradiction. En effet, pour deux indices quelconques $i < j$ on aura $t_i < t_j$ et donc $F(\varphi(t_i, x)) > F(\varphi(t_j, x))$ puisque la fonction $t \mapsto F(\varphi(t, x))$ est strictement décroissante. Or, comme à la fois $\varphi(t_i, x)$ et $\varphi(t_j, x)$ appartiennent à $\Sigma \subset \{F(x) = F(x_0)\}$, on a aussi $F(\varphi(t_i, x)) = F(\varphi(t_j, x)) = F(x_0)$, ce qui est absurde. ∎

Voici un exemple d'application du théorème 7.1. On considère l'équation du pendule avec frottement : $\ddot{x} + k\dot{x}^{2p+1} + \sin x = 0$, où $p \in \mathbb{N}$ et $k > 0$ (le frottement est linéaire lorsque $p = 0$). On considère le champ de vecteurs associé :

$$X = y\frac{\partial}{\partial x} - (ky^{2p+1} + \sin x)\frac{\partial}{\partial y},$$

défini dans l'espace de phase des $(x, y) \in S^1 \times R$ (la transformation d'une équation du second ordre en un champ de vecteurs dans l'espace de phase est expliquée dans le paragraphe 1.4). La fonction d'énergie totale $F(x, y) = \frac{1}{2}y^2 - \cos x$, du système conservatif correspondant à $k = 0$, est une fonction de Lyapounov au voisinage du point d'équilibre $(0, 0)$ pour les $k > 0$. En effet, le point $(0, 0)$ est un minimum local de Morse pour F (voir le paragraphe I-3.2) et :

$$\begin{aligned} X \cdot F(x, y) &= y\frac{\partial F}{\partial x} - (ky^{2p+1} + \sin x)\frac{\partial F}{\partial y} \\ &= y \sin x - (ky^{2p+1} + \sin x)y = -ky^{2(p+1)} \leq 0. \end{aligned}$$

Comme la fonction $-ky^{2(p+1)}$ s'annule sur tout l'axe des x, on ne peut appliquer le théorème 7.2. Malgré cela, il est très facile de montrer que la fonction d'énergie F est strictement décroissante au voisinage de $(0, 0)$, le long de chaque trajectoire distincte de ce point. Comme dans la preuve du théorème 7.2, on en déduit alors que ce point d'équilibre est asymptotiquement stable : chaque trajectoire du pendule, avec une condition initiale assez proche de $(0, 0)$, va tendre vers la position d'équilibre.

Un critère de stabilité

Soit X un champ de vecteurs sur une variété M de dimension n et $p \in M$ un point singulier de X. Par choix d'une carte, X est représenté par un champ de

vecteurs défini sur un voisinage ouvert de $0 \in \mathbb{R}^n$ et p est situé à l'origine. On notera encore ce champ par X. Développons-le à l'ordre un, à savoir : $X(x) = Ax + o(\| x \|)$ où A est une matrice de $\mathbb{R}^n$.

Dans la section 4.4, le *spectre des valeurs propres de X en p* est défini, par la définition 4.9, comme la donnée des valeurs propres de A (dans $\mathbb{C}$), comptées avec leur ordre de multiplicité. Nous avons montré par le lemme 4.1 que ce spectre est indépendant du choix de la carte.

Voici maintenant un théorème, donnant un critère de stabilité asymptotique, basé sur les propriétés du spectre des valeurs propres. Ce théorème est une conséquence du théorème 7.2.

***Théorème* 7.3 (Critère de stabilité d'un point singulier pour les champs de vecteurs).** *Si p est un point singulier de X et si toutes les valeurs propres du spectre en p ont une* partie réelle strictement négative, *alors p est asymptotiquement stable.*

Démonstration. On choisit une carte locale, au voisinage du point singulier $p = 0$, dans laquelle le champ s'écrit $X(x) = X_1(x) + O(||x||^2)$ avec une partie linéaire, $X_1(x) = Ax$, donnée par une matrice A. On va établir le théorème par une succession de lemmes simples.

La preuve va consister à vérifier, tout d'abord, qu'après un changement linéaire de coordonnées bien choisi, la forme quadratique $Q(x) = \frac{1}{2}(x_1^2 + \ldots + x_n^2)$ est une fonction de Lyapounov pour le champ linéaire X_1, avec $X_1 \cdot Q(x) < 0$ pour $x \neq 0$. Le résultat est direct si la matrice A est diagonale sur $\mathbb{R}$. En effet, si $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{R}$ est le spectre des valeurs propres, cela suit de l'observation que

$$X_1 \cdot Q = \langle Ax, x \rangle = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2 < -\mu ||x||^2 \tag{7.3}$$

avec $-\mu = \mathrm{Sup}\{\lambda_1, \ldots, \lambda_n\} < 0$ car, par hypothèse, $\lambda_i < 0$ pour tout i. On va tout d'abord montrer que l'inégalité (7.3) reste vraie, dans un bon système de coordonnées, pour une matrice quelconque, à un $\delta > 0$ arbitraire près, si on prend pour $-\mu$ la borne supérieure des parties réelles des valeurs propres. Pour ce faire, nous allons utiliser le résultat suivant de triangulation sur $\mathbb{C}$ d'une matrice quelconque, qui est une forme faible de la mise sous forme de Jordan (voir [16] par exemple). ∎

***Lemme* 7.3.** *Soit $A \in \mathrm{Mat}(n, n)$ une matrice réelle de spectre :*

$$\{\lambda_1, \ldots, \lambda_l, \alpha_1 \pm i\beta_1, \ldots, \alpha_s \pm i\beta_s\}$$

où $\lambda_i, \alpha_j, \beta_j \in \mathbb{R}$ et $n = l + 2s$. Alors, A est conjuguée dans $\mathrm{Gl}(n, \mathbb{R})$ à une matrice $S + N$ avec les propriétés suivantes :

1. $\mathbb{R}^n$ *se décompose en somme directe :* $\mathbb{R}^n = \oplus_{i=1}^{l+s} E_i$ *avec* $E_i = \mathbb{R}$ *si* $i \leq l$ *et* $E_i = \mathbb{R}^2$ *pour* $i = l+1, \ldots, l+s$.

2. S *laisse invariant cette décomposition :* $S(E_i) \subset E_i$ *pour tout* $i = 1, \ldots, l+s$.

3. *La restriction de* S *au facteur* E_i *est la multiplication par* λ_i *pour* $i \leq l$, *et est la similitude* $\begin{pmatrix} \alpha_{i-l} & -\beta_{i-l} \\ \beta_{i-l} & \alpha_{i-l} \end{pmatrix}$ *pour* $i = l+1, \ldots, l+s$.

4. *La matrice* N *est triangulaire supérieure au sens suivant : si* $v \in E_1$, *alors* $N(v) = 0$, *et si* $v \in E_k$ *avec* $k \geq 2$, *alors* $N(v) \in \oplus_{i=1}^{k-1} E_i$.

Remarque 7.4. Les conditions 1 et 2 du lemme 7.3 déterminent complètement la matrice S. Cette matrice contient l termes diagonaux correspondant aux valeurs propres réelles de A, puis s mineurs diagonaux de dimension 2 correspondant aux paires de valeurs propres complexes conjuguées de A. Tous les autres termes de S sont nuls. cette matrice est diagonalisable sur $\mathbb{C}$.

Remarque 7.5. La condition 3 implique clairement que la matrice N est une matrice triangulaire supérieure stricte (à éléments diagonaux nuls). Cette condition est plus générale que celle portant sur une matrice nilpotente élémentaire, qui est une matrice triangulaire supérieure stricte composée uniquement de zéros à l'exception de la première sur-diagonale composée uniquement de 1. C'est ce type de matrice nilpotente élémentaire qui est exhibée dans le théorème classique de Jordan (voir [16] par exemple). Dans notre cas, la matrice nilpotente N n'est pas élémentaire, c'est ce qui nous fait appeler le lemme 7.3 la forme faible de Jordan.

On désigne par $||B||$ la norme d'opérateur d'une matrice associée à la norme euclidienne de $\mathbb{R}^n$ (voir la définition 2.2). Alors on peut, dans cette norme, rendre aussi petite que voulu la norme de la matrice N du lemme ci-dessus :

Lemme 7.4. *Soit* $\delta > 0$ *un nombre positif arbitraire. On peut choisir la conjugaison dans le lemme précédent pour que* $||N|| < \delta$. *Autrement dit,* A *est conjuguée dans* $\mathrm{Gl}(n, \mathbb{R})$ *à* $S + N$, *comme dans le lemme 7.3, avec* $||N|| < \delta$.

Démonstration. On conjugue tout d'abord A à une matrice $S + N$ comme dans le lemme 7.3. Soit un ν, $0 < \nu < 1$, nombre que nous allons choisir dans la suite. Considérons l'élément $P \in \mathrm{Gl}(n, \mathbb{R})$ égal à l'homothétie de rapport ν^{n-i} sur chaque facteur E_i pour $i = 1, \ldots, l+s$. La conjugaison $B \mapsto PBP^{-1}$ laisse la matrice S invariante. Soit N_{ij} avec $1 \leq i < j \leq n$, les coefficients éventuellement non nuls de N. En utilisant la condition 3 du lemme 7.3, on vérifie sans peine que chaque coefficient de $\widetilde{N} = PNP^{-1}$, éventuellement non nul, est de la forme

$\tilde{N}_{ij} = \nu^{k(i,j)} N_{ij}$, pour $1 \leq i < j \leq n$, où $k(i,j)$ est un entier strictement positif. Remarquez que :

$$||\tilde{N}|| = \operatorname*{Sup}_u \left\{ \left(\sum_i \left(\sum_j \tilde{N}_{ij} u_j \right)^2 \right)^{\frac{1}{2}} \mid \sum_j u_j^2 = 1 \right\} \leq \left(\sum_i \left(\sum_j |\tilde{N}_{ij}| \right)^2 \right)^{\frac{1}{2}},$$

car $|u_j| \leq 1$ pour tout j. Comme $\nu^{k(i,j)} \leq \nu$ pour tout (i,j) (puisqu'on a choisi ν tel que $0 < \nu < 1$), on a $|\tilde{N}_{ij}| \leq \nu |N_{ij}|$ pour tout (i,j), d'où il résulte que :

$$||\tilde{N}|| \leq \nu \left(\sum_i \left(\sum_j |N_{ij}| \right)^2 \right)^{\frac{1}{2}}.$$

Il suffit alors de choisir $\nu > 0$ assez petit pour que $\nu (\sum_i (\sum_j |N_{ij}|)^2)^{\frac{1}{2}} \leq \delta$. ∎

On suppose maintenant, comme dans l'énoncé du théorème, que les valeurs propres de la matrice A ont des parties réelles strictement négatives. Soit $-\mu < 0$ la borne supérieure de ces parties réelles, c'est-à-dire :

$$-\mu = \operatorname{Sup} \{ \lambda_i,\ i = 1, \ldots, l,\ \ \alpha_j,\ j = 1, \ldots, s \}. \tag{7.4}$$

On suppose choisi le système de coordonnées linéaires $(x_1, \ldots, x_n)$ dans lequel A est égale à $S + N$, comme dans le lemme 7.4, et on désigne par $\langle \cdot, \cdot \rangle$ le produit scalaire euclidien de $\mathbb{R}^n$.

Lemme 7.5. *Le nombre μ étant défini par (7.4), supposons que $||N|| < \delta$ pour un certain $\delta > 0$. Alors :*

$$\langle Ax, x \rangle \leq -(\mu - \delta) ||x||^2. \tag{7.5}$$

Démonstration. Soit $u \in \mathbb{R}^n$. On a :

$$\langle Au, u \rangle = \langle Su, u \rangle + \langle Nu, u \rangle. \tag{7.6}$$

Les facteurs E_i sont deux à deux orthogonaux. Si on décompose $u = \sum_{i=1}^{l+s} u_i$ avec $u_i \in E_i$, on peut écrire que $\langle Su, u \rangle = \sum_{i=1}^{l+s} \langle Su_i, u_i \rangle$.

Maintenant si $i \leq l$, on a $\langle Su_i, u_i \rangle = \langle \lambda_i u_i, u_i \rangle = \lambda_i ||u_i||^2$. Si $j \geq l+1$, on a $u_j = (u_j^1, u_j^2) \in \mathbb{R}^2$ et donc $Su_j = (\alpha_{j-l} u_j^1 - \beta_{j-l} u_j^2, \beta_{j-l} u_j^1 + \alpha_{j-l} u_j^2)$. Il en résulte que $\langle Su_j, u_j \rangle = \alpha_{j-l} ||u_j||^2$, si $j \geq l+1$.

On obtient donc :

$$\langle Su, u \rangle = \sum_{i=1}^{l} \lambda_i ||u_i||^2 + \sum_{j=l+1}^{l+s} \alpha_{j-l} ||u_j||^2 \leq -\mu ||u||^2 \tag{7.7}$$

car $||u||^2 = \sum_{i=1}^{l+s} ||u_i||^2$, par le théorème de Pythagore ((7.7) est une inégalité stricte si $||u|| \neq 0$).

Par le lemme de Schwarz, on a $\langle Nu, u\rangle \leq ||Nu|| \cdot ||u||$. Comme $||N||$ est la norme d'opérateur associée à la norme euclidienne de $\mathbb{R}^n$, on a $||Nu|| \leq ||N|| \cdot ||u||$ et donc

$$\langle Nu, u\rangle \leq ||N|| \cdot ||u||^2 \leq \delta ||u||^2. \tag{7.8}$$

Le résultat suit des inégalités (7.7) et (7.8). ∎

Le théorème est une conséquence du théorème 7.2 et du lemme 7.6.

Lemme 7.6. *Soit $\delta > 0$ tel que $\mu - 2\delta > 0$. Plaçons-nous dans le système de coordonnées $x = (x_1, \ldots, x_n)$ fourni par le lemme 7.4. Alors il existe un voisinage ouvert Ω de l'origine sur lequel la fonction quadratique $Q(x) = \frac{1}{2}(x_1^2 + \ldots + x_n^2)$ vérifie*

$$X \cdot Q(x) \leq -(\mu - 2\delta)||x||^2. \tag{7.9}$$

Démonstration. La dérivée de la fonction Q par rapport au champ $X(x) = \sum_{i=1}^n X_i(x)\frac{\partial}{\partial x_i}$ est égale à

$$X \cdot Q(x) = \sum_{i=1}^{n} x_i X_i(x) = \langle X(x), x\rangle. \tag{7.10}$$

Dans les coordonnées choisies, le champ X s'écrit

$$X(x) = Ax + O(||x||^2).$$

Il en résulte que $\langle X(x), x\rangle = \langle Ax, x\rangle + O(||x||^3)$. Compte tenu de l'inégalité (7.5), cela implique qu'il existe une fonction $\psi(x)$ qui tend vers 0 quand x tend vers l'origine, telle que $\langle X(x), x\rangle \leq -(\mu - \delta - \psi(x))||x||^2$. Si l'on choisit un voisinage ouvert Ω de l'origine sur lequel $|\psi(x)| \leq \delta$, on a l'inégalité

$$\langle X(x), x\rangle \leq -(\mu - 2\delta)||x||^2 \tag{7.11}$$

dans tout Ω, qui se réduit à l'inégalité (7.9) compte tenu de (7.10). ∎

Fin de la Preuve du Théorème 7.3. La fonction Q possède un minimum absolu strict à l'origine $0 \in \mathbb{R}^n$ et elle est évidemment de classe $\mathcal{C}^1$! L'inégalité (7.9) implique que Q est une fonction de Lyapounov sur Ω, telle que $X \cdot Q(x) < 0$ pour tout $x \in \Omega - \{0\}$. La stabilité asymptotique suit alors du théorème 7.2. ∎

***Remarque* 7.6.** Vérifier ou prouver que toutes les valeurs propres ont une partie réelle négative n'est pas évident. Cependant, le lecteur intéressé trouvera, par exemple dans le chapitre 2 de [23], une série de résultats concernant les signes des parties réelles des zéros du polynôme caractéristique en fonction de critères portant sur les coefficients de ce polynôme.

On peut tirer, de l'inégalité (7.9), une estimation quantitative de la vitesse de convergence des trajectoires vers l'origine. Soit un $\delta > 0$ comme dans le lemme 7.6. Choisissons un $\rho = \rho(\delta) > 0$ tel que $\overline{B_\rho(0)} \subset \Omega$. Remarquons que l'on peut choisir $\delta > 0$ aussi petit que voulu, mais que l'on doit choisir $\rho(\delta) \to 0$ pour $\delta \to 0$. On désigne par φ le flot de X. Dans $\overline{B_\rho(0)}$ on a l'inégalité différentielle suivante.

***Lemme* 7.7.** *Soit $\rho = \rho(\delta)$ comme ci-dessus. Soit $t \geq 0$ et x tels que $\varphi(t,x) \in \overline{B_\rho(0)}$. Alors*

$$\frac{\partial}{\partial t}||\varphi(t,x)||^2 \leq -2(\mu - 2\delta)||\varphi(t,x)||^2. \tag{7.12}$$

Démonstration. On a $||\varphi(t,x)||^2 = \langle \varphi(t,x), \varphi(t,x)\rangle$. En utilisant la bilinéarité et la symétrie du produit scalaire euclidien, on obtient :

$$\frac{\partial}{\partial t}\langle \varphi(t,x), \varphi(t,x)\rangle = 2\langle \varphi(t,x), \frac{\partial \varphi}{\partial t}(t,x)\rangle. \tag{7.13}$$

Comme $\frac{\partial \varphi}{\partial t}(t,x) = X(\varphi(t,x))$, on a

$$\langle \varphi(t,x), \frac{\partial \varphi}{\partial t}(t,x)\rangle = \langle X(\varphi(t,x)), \varphi(t,x)\rangle. \tag{7.14}$$

Le point $\varphi(t,x)$ appartenant à $\overline{B_\rho(0)}$ par hypothèse, il suit de (7.11) que

$$\langle X(\varphi(t,x)), \varphi(t,x)\rangle \leq -(\mu - 2\delta)||\varphi(t,x)||^2,$$

et donc l'inégalité (7.12) souhaitée d'après (7.14) et (7.13). ∎

Montrons que si $x \in \overline{B_\rho(0)}$, alors toute la demi-trajectoire positive de x est aussi contenue dans cette boule :

***Lemme* 7.8.** *Soit $x \in \overline{B_\rho(0)}$, avec $\rho = \rho(\delta)$. alors pour tout $t \geq 0$ on a $\varphi(t,x) \in \overline{B_\rho(0)}$.*

Démonstration. Soit $F = \{\tau \in \mathbb{R}^+ \mid \varphi(\tau, x) \in \overline{B_\rho(0)}\}$. C'est un fermé non vide, car il contient $0 \in \mathbb{R}^+$. Supposons que $F \neq \mathbb{R}^+$. Soit alors

$$\tau^0 = \mathrm{Inf}\{\tau \in \mathbb{R}^+ \mid \tau \notin F\}.$$

Par définition, on a $\varphi(\tau, x) \in \overline{B_\rho(0)}$ pour tout $\tau \leq \tau^0$ et il existe une suite $(\tau_i)_i \to \tau^0$, tels que $\tau_i \geq \tau^0$ et $\varphi(\tau_i, x) \notin \overline{B_\rho(0)}$ pour tout i. Par continuité, cela implique que $\varphi(\tau_0, x) \in \partial\overline{B_\rho(0)} = S_\rho$, la sphère de rayon ρ. Or, le champ de vecteurs X est *transverse et rentrant* en chaque point de S_ρ. Cela suit de l'inégalité $< X(x), x >< 0$ vérifiée en tout $x \in S_\rho$ (cette inégalité signifie que l'angle entre le vecteur sortant extérieur x et le vecteur $X(x)$ est strictement supérieur à $\frac{\pi}{2}$). Donc, il existe $\eta > 0$ tel que, pour tout $\tau \in]\tau^0 - \eta, \tau^0[$, le point $\varphi(\tau, x)$ est situé à l'extérieur de $\overline{B_\rho(0)}$, ce qui est en contradiction avec la définition de τ^0. ∎

Nous pouvons maintenant intégrer l'inéquation différentielle (7.12) tout au long d'une orbite partant de $x \in \overline{B_\rho(0)}$, pour obtenir l'estimation suivante de la convergence de $\varphi(t, x)$ vers l'origine.

Proposition 7.1. *Soit $\rho = \rho(\delta)$ comme plus haut et $x \in \overline{B_\rho(0)}$, alors pour tout $t \in \mathbb{R}^+$:*

$$||\varphi(t, x)|| \leq ||x|| e^{-(\mu - 2\delta)t}. \tag{7.15}$$

Démonstration. Le point x étant fixé, posons $f(t) = ||\varphi(t, x)||^2$. D'après le lemme 7.8, la fonction $f(t)$ est définie pour tout $t \geq 0$, et d'après le lemme 7.7, cette fonction (qui est C^∞) vérifie l'inéquation différentielle

$$\frac{df}{dt}(t) \leq -2(\mu - 2\delta) f(t)$$

pour tout $t \geq 0$, avec, comme condition initiale, $f(0) = ||x||^2$. Posons $g(t) = f(t)e^{2(\mu - 2\delta)t}$, on voit que $g \geq 0$. En dérivant $f(t) = g(t)e^{-2(\mu - 2\delta)t}$, on obtient

$$\frac{dg}{dt}(t)e^{-2(\mu - 2\delta)t} - 2(\mu - 2\delta)g(t)e^{-2(\mu - 2\delta)t} \leq -2(\mu - 2\delta)g(t)e^{-2(\mu - 2\delta)t},$$

ce qui implique que $\frac{dg}{dt}(t) \leq 0$, et donc que $g(t) \leq g(0) = ||x||^2$. En revenant à la fonction f, on obtient $f(t) \leq ||x||^2 e^{-2(\mu - 2\delta)t}$, et donc l'inégalité (7.15) souhaitée. ∎

Remarque 7.7. La formule (7.15), qui donne une forme forte et quantitative de convergence asymptotique, est appelée *convergence exponentielle*. Remarquez que le coefficient $\mu - 2\delta$, qui contrôle la convergence, peut être choisi (par restriction du domaine de validité) aussi proche que voulu de μ, constante ne dépendant que du spectre des valeurs propres du champ X au point singulier.

7.2. Stabilité d'une orbite périodique

7.2.1. Différents types de stabilité pour une orbite périodique

Les définitions de stabilité, pour une orbite périodique γ d'un champ X sur une variété M, vont reprendre celles introduites plus haut pour les points singuliers, l'orbite jouant le rôle du point. Cependant, comme l'orbite périodique n'est pas nécessairement contenue dans une seule carte, on ne peut pas considérer la situation d'un champ de vecteurs sur un ouvert de $\mathbb{R}^n$. Pour contourner cette petite difficulté, nous allons faire appel au choix d'une distance $\mathrm{dist}(x, y)$ pour définir la topologie de M. Rappelons qu'une telle distance peut être, par exemple, obtenue à partir d'une métrique riemannienne choisie sur M, comme il a été expliqué au paragraphe I-2.4.3.

Lyapounov a introduit sa notion de stabilité pour *toute trajectoire* (et non pas seulement pour les points singuliers) :

$\varphi(t, x)$ est stable si et seulement si pour tout y assez proche de x, $\varphi(t, y)$ reste proche de $\varphi(t, x)$ pour tout $t \geq 0$. Autrement dit, pour tout $\varepsilon > 0$ et tout $t \geq 0$, il existe un $\delta(\varepsilon) > 0$, tel que :

$$\mathrm{dist}(x, y) \leq \delta(\epsilon) \Rightarrow \mathrm{dist}(\varphi(t, x), \varphi(t, y)) \leq \epsilon\,.$$

C'est cette notion de stabilité qui est utilisée pour les champs non autonomes et non périodiques. Cette notion pourrait être utilisée pour les orbites périodiques des champs autonomes, mais elle est très restrictive, car elle exige que les points $\varphi(t, x)$ et $\varphi(t, y)$ restent proches, ceci pour tout temps $t \geq 0$. Aussi, on a intérêt à affaiblir cette exigence en demandant seulement que pour y voisin de l'orbite γ le point $\varphi(t, y)$ reste également proche de γ pour les temps $t \geq 0$. On rappelle que la distance d'un point x à un fermé non vide F est définie par $\mathrm{dist}(x, F) = \mathrm{Inf}\{\mathrm{dist}(x, y) \mid y \in F\}$. On a alors la définition suivante de la stabilité de Lyapounov pour les orbites périodiques :

***Définition 7.4* (Stabilité de Lyapounov d'une orbite périodique).** Supposons que γ soit périodique. On dit que γ est *stable* (au sens de Lyapounov) si et seulement si pour tout $\epsilon > 0$, il existe $\delta(\epsilon)$ tel que :

$$\mathrm{dist}(y, \gamma) \leq \delta(\epsilon) \Rightarrow \ \forall t \geq 0 : \mathrm{dist}(\varphi(t, y), \gamma) \leq \epsilon. \qquad (7.16)$$

La stabilité au sens de Lyapounov d'une orbite périodique γ par le point x_0 n'implique pas que

$$\text{dist}(x,y) \leq \delta(\epsilon) \Rightarrow \text{dist}(\varphi(t,x),\varphi(t,y)) \leq \epsilon\,. \tag{7.17}$$

Ceci est illustré dans la figure 7.3 : on y voit l'orbite périodique en trait continu avec les points x et $\varphi(t_1,x)$ d'une part, l'orbite périodique perturbée en trait pointillé avec les points y et $\varphi(t_1,y)$ d'autre part. On reste près de l'orbite périodique mais on ne contrôle pas ce qui se passe, les points $\varphi(t_1,x)$ et $\varphi(t_1,y)$ ne sont pas nécessairement proches même si les points x et y le sont. L'orbite périodique γ est stable, au sens de Lyapounov, alors que l'implication (7.17) est fausse.

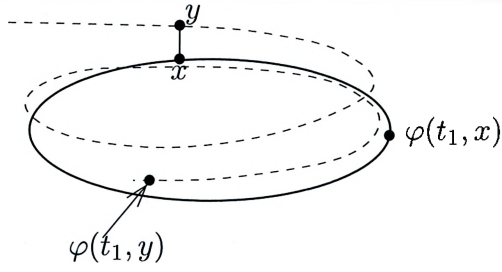


FIGURE 7.3. L'orbite périodique γ est en trait continu.

Voici un exemple explicite de ce phénomène. Sur l'anneau ouvert $A = \mathbb{R}/\mathbb{Z}\times]0,1[$, paramétré par $(\theta,r) \in \mathbb{R}/\mathbb{Z}\times]0,1[$, on considère le champ de vecteur $r\frac{\partial}{\partial\theta}$ (en plongeant A dans $\mathbb{R}^2$, par exemple par le difféomorphisme $(\theta,r) \mapsto (x = e^r\cos 2\pi\theta, y = e^r \sin 2\pi\theta)$, on peut considérer A comme un ouvert de $\mathbb{R}^2$). Clairement, toutes les orbites sont périodiques et stables au sens de Lyapounov. Par contre, elles ont des périodes différentes, la période par le point $m_r = (0,r)$ étant égale à $\frac{1}{r}$. Si l'on prend deux points quelconques m_{r_1}, m_{r_2}, avec $0 < r_1 < r_2 < 1$, on voit que les trajectoires par ces deux points auront un écart angulaire de la variable θ *maximal,* c'est-à-dire égal à $\frac{1}{2}$ pour la suite des temps $t_n = \frac{1}{2}n(\frac{1}{r_1} - \frac{1}{r_2})$, avec $n \in \mathbb{Z}$. Si l'on plonge l'anneau dans $\mathbb{R}^2$ comme ci-dessus, il est clair que $m_{r_2} \to m_{r_1}$ si $r_2 \to r_1$, alors que la distance euclidienne entre les deux points atteints aux temps t_n est égale à $e^{r_1} + e^{r_2}$ et tend vers $2e^{r_1} \neq 0$ si $r_2 \to r_1$.

Comme dans le cas des points singuliers, on peut donner une formulation équivalente de la stabilité, au sens de Lyapounov, d'une orbite périodique, en termes de l'existence d'un système fondamental de voisinages invariants de l'orbite. La démonstration est complètement similaire à celle donnée pour un point singulier (voir lemme 7.1), il suffit de remplacer la norme de $\mathbb{R}^n$ par la distance riemannienne $\text{dist}(x,y)$ (voir le paragraphe I-2.4.3).

Lemme 7.9. *Une orbite périodique γ est stable au sens de Lyapounov si et seulement si γ possède un système fondamental de voisinages invariants par φ_t pour les temps positifs.*

On définit de la même manière la *stabilité asymptotique des orbites périodiques* :

Définition 7.5. On dit qu'une orbite périodique γ est *asymptotiquement stable* s'il existe un voisinage Ω de γ tel que, pour tout $\varepsilon > 0$, il existe un $t(\varepsilon)$ avec la propriété suivante :

$$\forall x \in \Omega,\ \forall t \geq t(\varepsilon) \Longrightarrow \mathrm{dist}(\varphi(t,x),\gamma) \leq \varepsilon.$$

Comme dans le cas des points singuliers, pour qu'une orbite soit asymptotiquement stable, il faut et suffit que, pour tout $x \in \Omega$, on ait $\varphi(t,x) \to \gamma$. L'uniformité de la convergence suit de la compacité de l'orbite γ. De même, la stabilité asymptotique entraîne la stabilité de Lyapounov.

La stabilité de Lyapounov d'une orbite critique γ, point singulier ou orbite périodique, est préservée par équivalence $\mathcal{C}^\infty$ entre champs de vecteurs. Cela suit trivialement de la forme équivalente de la définition, en termes de l'existence d'un système fondamental de voisinages invariants de l'orbite. La stabilité asymptotique est aussi préservée par équivalence $\mathcal{C}^\infty$, car cette stabilité revient à dire que $\varphi(t,x) \to \gamma$ pour les points x voisins de γ, condition qui est aussi clairement invariante par équivalence $\mathcal{C}^\infty$. Ces remarques suggèrent que les notions de stabilité doivent pouvoir se traduire et s'étudier à travers les applications de Poincaré de γ. C'est ce que nous allons examiner dans la suite de ce chapitre, après avoir défini, dans le prochain paragraphe, les notions de stabilité au voisinage d'un point fixe de difféomorphisme.

7.2.2. Différents types de stabilité pour un point fixe de difféomorphisme

Soit Σ une variété de dimension m. On suppose que f est un difféomorphisme défini au voisinage d'un point fixe p de f (point tel que : $f(p) = p$). Dans la suite, f sera une application de Poincaré associée à une orbite périodique, mais il est intéressant de donner une définition indépendante de ce contexte, valable pour toute dynamique discrète locale.

Dire que f est définie au voisinage du point fixe p signifie que f est définie sur un voisinage ouvert arbitraire W de $p \in \Sigma$. On peut prendre pour W un domaine

de carte, et donc supposer que f est définie sur un ouvert de $\mathbb{R}^m$. C'est ce que l'on va supposer dorénavant. Cela va nous permettre d'utiliser la norme euclidienne $||\cdot||$ de $\mathbb{R}^m$. On désigne par B_{ρ_0} la boule de centre p et de rayon ρ_0 dans $\mathbb{R}^m$.

Définition 7.6.

1. On dit que le point fixe p est *stable, au sens de Lyapounov,* si, pour tout $\epsilon > 0$, il existe $\delta(\epsilon) > 0$ tel que :
$$\| x - p \| \leq \delta(\epsilon) \Rightarrow \forall n \in \mathbb{N}, \| f^n(x) - p \| \leq \epsilon.$$
2. On dit que p est *asymptotiquement stable* s'il existe $\rho_0 > 0$ tel que $B_{\rho_0} \subset W$, avec la propriété suivante. Pour tout $\epsilon > 0$, il existe $t(\epsilon) \in \mathbb{R}^+$ tel que :
$$\| x - p \| \leq \rho_0 \text{ et } t \geq t(\epsilon) \Rightarrow \| f^n(x) - p \| \leq \epsilon.$$

Dans les définitions ci-dessus, f^n est le difféomorphisme obtenu en itérant f n fois : $f^n = f \circ \ldots \circ f$, n fois.

Remarque 7.8. On a une analogie manifeste entre les notions de stabilité au voisinage d'un point fixe de difféomorphisme et celles données pour un flot de champ de vecteurs, au voisinage d'un point singulier : le temps continu t du flot est simplement remplacé par le « temps discret » n. Les remarques et propriétés valables pour les champs sont encore vraies ici, à savoir que la stabilité de Lyapounov est équivalente à l'existence d'un système fondamental de voisinages invariants par le difféomorphisme f. De même que la stabilité asymptotique implique la stabilité de Lyapounov. De plus, la stabilité asymptotique est impliquée par la convergence $f^n(x) \to p$ pour $n \to +\infty$, pour les points x assez voisins de p, l'uniformité s'en déduisant par un simple argument de compacité.

7.2.3. Relation entre la stabilité d'une orbite périodique et celle de ses applications de Poincaré

Soit X un champ de vecteurs, γ_p une orbite périodique par un point $p \in M$. On choisit une section locale Σ par p. Soit $h : W \mapsto \Sigma$ $(W \subset \Sigma)$ une application de Poincaré associée. D'après sa définition 4.7, Σ est difféomorphe au disque unité de $\mathbb{R}^{n-1}$, muni de la distance euclidienne. Les diamètres ci-dessous sont définis par la norme euclidienne associée à cette distance. On suppose que $p = 0$. On a alors la proposition suivante :

Proposition 7.2. *Le point p est un point fixe stable au sens de Lyapounov (respectivement au sens asymptotique) pour le difféomorphisme h si et seulement si γ_p est une orbite stable au sens de Lyapounov (respectivement au sens asymptotique) du champ de vecteurs X.*

Démonstration. On utilise la même idée de compacité, considérée précédemment dans le lemme 7.2 : on utilise le fait que le temps de retour de h est borné.

Comme les preuves des différentes implications sont similaires, nous allons nous contenter de faire l'une d'entre elles, à savoir que si h est stable au sens de Lyapounov, alors γ_p est stable au sens de Lyapounov. Le cas asymptotique se traite de la même façon, ainsi que les réciproques.

Supposons donc que h soit stable au sens de Lyapounov en p. Alors il existe un système fondamental de voisinages compacts $U_i \subset W$ de p, invariants par h (voir remarque 7.8) ; c'est-à-dire que $h(U_i) \subset U_i$ et $\operatorname{diam}(U_i) \to 0$ pour $i \to \infty$.

Pour tout i, on considère le *saturé* W_i *de* U_i par les trajectoires, obtenu en prenant toutes les trajectoires partant de U_i, soit

$$W_i = \bigcup_{t \geq 0} \{\varphi_t(x) \mid x \in U_i\}. \tag{7.18}$$

Les W_i sont des *voisinages invariants* de γ_p pour les temps positifs du flot car, pour tout $\tau \geq 0$, on a

$$\varphi_\tau(W_i) = \bigcup_{t \geq 0} \{\varphi_{t+\tau}(x) \mid x \in U_i\} \subset W_i.$$

Soit $x \in U_i$ et $t(x)$ le temps de premier retour. Comme $h(U_i) \subset U_i$ (figure 7.4), on repassera par les points déjà atteints de W_i après le temps de retour $t(x)$, car on repart d'un ensemble $h(U_i)$ plus petit que U_i. Dans la définition du saturé par (7.18), il suffit donc de se restreindre aux temps $0 \leq t \leq t(x)$, soit

$$W_i = \bigcup_{t \leq t(x)} \{\varphi_t(x) \mid x \in U_i\}.$$

On peut choisir W compact. Le temps $t(x)$ est alors borné pour $x \in W$. Soit $T_1 = \operatorname{Sup}\{t(x) \mid x \in W\}$. On peut alors remplacer, dans la formule ci-dessus, le temps variable $t(x)$ par le temps T_1, et écrire pour tout i

$$W_i = \bigcup_{t \leq T_1} \{\varphi_t(x) \mid x \in U_i\} = \varphi([0, T_1] \times U_i).$$

On définit $\operatorname{diam}(W_i)$ par

$$\operatorname{diam}(W_i) = \sup\{\operatorname{dist}(\varphi(t, y), \gamma_p) \mid (t, y) \in [0, T_1] \times U_i\}.$$

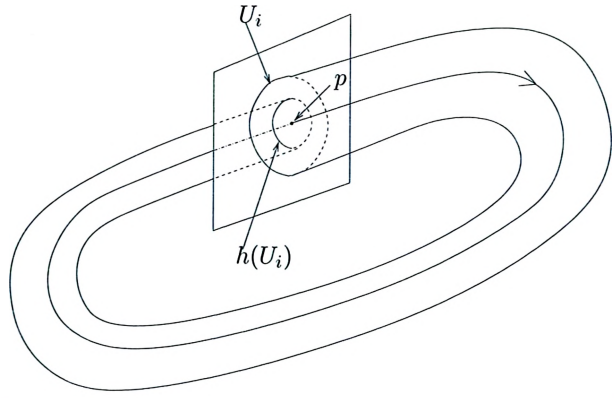


FIGURE 7.4

Alors on a $\text{diam}(W_i) \to 0$ lorsque $i \to \infty$ (c'est-à-dire que $W_i \to \gamma_p$ au sens de la topologie de Hausdorff, lorsque $i \to \infty$).

Pour démontrer ce point, on va utiliser un argument simple de *compacité,* argument déjà utilisé par ailleurs comme il a été dit. Voici cet argument. La fonction à deux variables $f(t,y) = \text{dist}(\varphi(t,y), \gamma_p)$ est continue et positive. De plus, $f(t,0) = 0$, puisqu'on est sur γ_p pour $y = 0$. Soit maintenant un $\varepsilon > 0$ quelconque. Par continuité de f, il existe, pour tout $t \in [0,T_1]$, un $\eta_t > 0$ et un $\delta_t > 0$ tels que $f(]t-\eta_t, t+\eta_t[\times B_{\delta_t}(0)) \leq \varepsilon$. Comme $[0,T_1]$ est *compact,* il existe un ensemble fini $\{t_1,\ldots,t_k\}$ tel que les intervalles $]t_1-\eta_{t_1}, t_1+\eta_{t_1}[, \ldots,]t_k-\eta_{t_k}, t_k+\eta_{t_k}[$ recouvrent $[0,T_1]$. Si $\delta = \text{Inf}\{\delta_{t_i} \mid i = 1,\ldots,k\}$, on a $f([0,T_1] \times B_\delta(0)) \leq \varepsilon$ et donc $\text{diam}(W_i) \leq \varepsilon$. Cela sera réalisé si $\text{diam}(U_i) \leq \delta$, c'est-à-dire si i est assez grand.

Cela montre que les W_i forment un système fondamental de voisinages de γ_p, invariants par φ_t par définition, et donc que γ_p est Lyapounov-stable (voir remarque 7.8). ■

Remarque 7.9. Ainsi, les stabilités de γ_p se réduisent à celles du difféomorphisme de Poincaré h associé à une *section quelconque.* Si on change de section, on change h à conjugaison près, ce qui ne modifie pas les propriétés de stabilité du difféomorphisme. Ceci est évidemment cohérent avec la proposition ci-dessus.

7.2.4. Théorèmes de stabilité

Fonction de Lyapounov pour un difféomorphisme

On peut étendre la notion de fonction de Lyapounov pour un difféomorphisme local h, au voisinage d'un point fixe.

Définition 7.7. Une fonction de Lyapounov F pour un difféomorphisme local h, au voisinage du point fixe p, est une fonction différentiable définie sur un voisinage ouvert U de p telle que :

1. Le point p est un minimum absolu strict de F : $F(x) > F(p)$ pour tout $x \in U - \{p\}$.
2. En tout point $x \in U$, on a $F(h(x)) \leq F(x)$: la valeur de F diminue lorsque l'on itère h.

On a des versions similaires des théorèmes 7.1 et 7.2 pour les difféomorphismes. Les démonstrations sont très comparables et on ne les répétera pas.

Théorème 7.4. *Supposons qu'il existe une fonction de Lyapounov F pour le difféomorphisme local h au voisinage du point fixe p. Alors le point p est stable au sens de Lyapounov. Si de plus on a une inégalité stricte $F(h(x)) < F(x)$ pour tout $x \in U - \{p\}$, alors p est asymptotiquement stable.*

Critère de stabilité

Choisissons une carte locale au voisinage de p, telle que $p = 0$. Soit

$$h(x) = Hx + o(\| x \|)$$

où $H \in \mathrm{Gl}(m, \mathbb{R})$, partie linéaire de h en 0, est une matrice inversible. Comme dans le cas d'un point singulier de champ de vecteurs (définition 4.9), le *spectre des valeurs propres* de h en p est l'ensemble des valeurs propres de H dans $\mathbb{C}$, comptées avec leur ordre de multiplicité. Ici, comme H est inversible, toutes les valeurs propres sont non nulles. Ce spectre est indépendant du choix de la carte : par changement de carte, H est modifiée par conjugaison en une matrice semblable PHP^{-1} ayant même spectre que H.

On a le critère suivant de stabilité asymptotique :

Théorème 7.5 **(Critère de stabilité pour les difféomorphismes).** *Supposons que toutes les valeurs propres du spectre de h au point fixe p aient un module strictement inférieur à 1 (toutes les valeurs propres sont à l'intérieur du cercle unité de $\mathbb{C}$). Alors p est un point asymptotiquement stable de h. Plus précisément, il existe un $K < 1$ et un $\rho > 0$ tels que, pour une certaine norme euclidienne $\| \cdot \|$ de $\mathbb{R}^m$, on ait*

$$\|h^n(x)\| \leq K^n \|x\|$$

pour tout $n \in \mathbb{N}$ et pour tout $x \in B_\rho(0)$, boule à fermeture contenue dans le domaine de définition de h : on a une convergence géométrique *uniforme des itérés de h vers 0, à partir de tous les points de $B_\rho(0)$ (la convergence est uniforme puisque $||h^n(x)|| \leq K^n\rho$).*

Démonstration. La preuve est en fait plus facile que dans le cas des champs de vecteurs. Comme elle est assez simililaire, nous allons en donner seulement les grandes lignes.

On a : $h(x) = Hx + o(||x||)$ sur un domaine de définition U, que l'on peut supposer être un ouvert de $\mathbb{R}^m$. Soit $\alpha = \sup_i |\lambda_i|$, les λ_i étant les valeurs propres de H. L'hypothèse est que $0 < \alpha < 1$. Comme dans le cas des champs de vecteurs, on a le résultat algébrique suivant, obtenu en choisissant une réduction triangulaire convenable de H : *pour tout $\delta > 0$, on peut trouver des coordonnées linéaires de $\mathbb{R}^m$ telles que dans ce système de coordonnées, encore noté : $x = (x_1, \ldots, x_m)$, la norme d'opérateur de $x \mapsto Hx$ soit bornée par $\alpha + \delta$: $||H|| < \alpha + \delta$.*

Comme dans le cas des champs, on en déduit l'existence d'un $\rho = \rho(\delta)$, tel que dans la boule $\Omega = B_\rho(0)$, supposée contenue dans U, on ait :

$$||h(x)|| \leq (\alpha + 2\delta)||x||.$$

On choisit $\delta > 0$ assez petit pour que $K = \alpha + 2\delta < 1$. Pour cette valeur K, on a donc :

$$||h(x)|| \leq K||x|| \tag{7.19}$$

si $x \in \Omega$. Cela implique en particulier que $h(\Omega) \subset \Omega$ et que l'on peut itérer indéfiniment h dans Ω. On peut aussi itérer l'inégalité (7.19), ce qui donne

$$||h^n(x)|| \leq K^n||x|| \leq K^n\rho,$$

pour tout $n \in \mathbb{N}$. Les itérés $h^n(x)$ convergent géométriquement en norme vers le point fixe 0. ∎

Application aux orbites périodiques

***Définition 7.8* (Multiplicateurs de Floquet).** Soit $h : W \mapsto \Sigma$ l'application de Poincaré relative à une section locale Σ par p. Les multiplicateurs de Floquet de γ_p sont les éléments du spectre des valeurs propres de h au point fixe p. Autrement dit, si $H = dh(p)$, partie linéaire de h en p dans une carte de W, les *multiplicateurs de Floquet* sont les valeurs propres de H dans $\mathbb{C}$, comptées avec leur multiplicité.

Les multiplicateurs de Floquet sont aussi appelés les exposants de Floquet. Dans la littérature on les trouve aussi sous les noms de multiplicateurs (ou facteurs) caractéristiques, qui sont à distinguer des exposants caractéristiques désignant, on le rappelle, les exposants des exponentielles solutions des systèmes différentiels linéaires.

Le spectre des valeurs propres est indépendant du choix de la carte (des coordonnées de W) mais aussi du choix de Σ : en effet si Σ_1 et Σ_2 sont deux sections, on passe de h_1 à h_2 par une conjugaison (voir le chapitre 6) ; c'est pour cela que l'on peut parler des multiplicateurs de l'orbite : ils sont indépendants de la section considérée.

Il suit alors de la proposition 7.2 et du théorème 7.5 qu'une orbite périodique est asymptotiquement stable *si* ses multiplicateurs de Floquet sont de module strictement inférieur à 1.

Remarque 7.10. Soit X un champ de vecteurs de flot φ. Supposons que $X(p) = 0$. Alors p est un point fixe du difféomorphisme φ_t pour tout t. Si p est un point fixe asymptotiquement (resp. Lyapounov) stable pour X alors, pour tout $t > 0$, p est un point fixe asymptotiquement (resp. Lyapounov) stable du difféomorphisme φ_t. Cela suit immédiatement des définitions.

Remarquons, par ailleurs, que si A est la partie linéaire de X en $p = 0$, alors e^{tA} est celle de φ_t (en effet si $X(x) = Ax + o(||x||)$ alors, pour tout $t \neq 0$, $\varphi_t(x) = e^{tA}.x + o(||x||)$). Donc si $\{\lambda_1, \ldots, \lambda_n\}$ est le spectre de X en 0, $\{e^{\lambda_1 t}, \ldots, e^{\lambda_n t}\}$ est celui de φ_t. La condition $Re(\lambda_i) < 0$ implique que $|e^{\lambda_i t}| < 1$ (car $e^{\lambda t} = e^{Re(\lambda)t} e^{iIm(\lambda)t}$). Tout ceci est cohérent avec les théorèmes précédents.

Remarque 7.11. L'application du théorème de Lyapounov aux orbites périodiques nécessite le *calcul pratique* des exposants de Floquet, par exemple par intégration numérique d'une *équation variationnelle au premier ordre* le long de l'orbite périodique (cette équation permet de contrôler la partie linéaire du flot transversalement à l'orbite).

Petit aperçu sur la stabililité dans le cas conservatif

Le critère de stabilité du théorème 7.5 suppose que le difféomorphisme est dissipatif et utilise l'existence d'un « effet dissipatif linéaire ». Dans le cas conservatif, c'est-à-dire pour un difféomorphisme symplectique qui, par définition, préserve une 2-forme symplectique (voir la définition dans [3] par exemple), l'étude de la stabilité est beaucoup plus difficile. Il a fallu attendre les années 1960 pour avoir un premier résultat de stabilité en dimension 2 (dans ce cas un difféomorphisme symplectique est un difféomorphisme préservant l'aire de $\mathbb{R}^2$) :

***Théorème 7.6* (Kolmogorov-Arnold-Moser).** *Soit h un difféomorphisme local en $0 \in \mathbb{R}^2$, ayant ce point comme point fixe. Supposons que h préserve l'aire (*$\text{aire}(h(V)) = \text{aire}(V)$*, pour tout ouvert $V \subset \mathbb{R}^2$). Supposons que le développement limité de h s'écrive en coordonnées polaires (r, θ) comme suit :*

$$h = \begin{cases} \theta_1 = \theta + \mu_0 + \mu_1 r^2 + o(\| r \|^2) \\ r_1 = r + \dots \end{cases}, \tag{7.20}$$

alors, si $\mu_1 \neq 0$, l'origine est stable pour h au sens de Lyapounov.

La preuve de ce théorème est très difficile. On pourra en trouver une approche dans le petit livre de Moser [20]. Il admet des versions analytique (Arnold) ou bien différentiable (Moser). Nous les admettrons évidemment.

Remarque 7.12. Le théorème de Kolmogorov-Arnold-Moser concerne les difféomorphismes de $\mathbb{R}^2$. Nous en avons donné ci-dessus une version faible. En fait, Arnold dans le cas analytique, et Moser dans le cas différentiable, ont montré qu'une infinité de courbes invariantes entouraient l'origine en s'y accumulant : c'est l'existence de ces courbes invariantes qui implique la stabilité au sens de Lyapounov. Il n'existe pas de critère de stabilité comparable pour les difféomorphismes symplectiques en dimension $2n$, avec $n > 1$. Il existe seulement des critères pour l'existence générique de tores invariants de dimension n, dont la dimension est trop petite pour assurer la stabilité de Lyapounov : des trajectoires « peuvent s'échapper » en passant entre les tores. Ce phénomène est appelé : *diffusion d'Arnold.*

Remarque 7.13. Les hypothèses du théorème 7.6 sont génériques pour les difféomorphismes symplectiques (elles sont vérifiées pour presque tous les h symplectiques au voisinage d'un point fixe elliptique, c'est-à-dire dont la partie linéaire est une rotation). Un tel difféomorphisme conservatif ne peut pas avoir de partie radiale non triviale (en conséquence, on peut supposer que $r_1 - r$ est une fonction plate dans (7.20)). S'il existe une fonction de Lyapounov, elle doit être conservée, ce qui est une situation hautement non générique pour les difféomorphismes généraux. Cela fait que le théorème 7.2 perd de son intérêt dans le cas conservatif, et fait au contraire tout l'intérêt du théorème 7.6 de Kolmogorov-Arnold-Moser dans le cas symplectique.

Remarque 7.14. Si l'on tronque la formule (7.20) à l'ordre 2, on obtient un difféomorphisme h_2 qui laisse invariants les cercles et y induit une *rotation différentielle* $\theta \mapsto \theta + \mu_0 + \mu_1 r^2$, où μ_0 est une rotation à l'origine (on constate que cette rotation dépend de r par le terme en μ_1). C'est cette *rotation différentielle* qui est le mécanisme de la stabilité dans la démonstration du théorème 7.6.

BIBLIOGRAPHIE

[1] R. Abraham, J. Robbins, *Transversal mappings and flows*, W.A. Benjamin, New-York, 1967.

[2] V. Arnold, *Équations différentielles ordinaires,* Éditions Mir, Moscou, 1974.

[3] V. Arnold, *Méthodes mathématiques de la Mécanique classique,* Éditions Mir, Moscou, 1976.

[4] M. Berger et B. Gostiaux, *Géométrie différentielle : variétés, courbes et surfaces,* PUF, Paris, 1987.

[5] J. Bochnak, M. Coste, M.-F. Roy, *Géométrie algébrique réelle,* Springer, Berlin, 1987.

[6] H. Cartan, *Cours de Calcul Différentiel,* Hermann, Paris, 1977.

[7] J. Dieudonné, *Éléments d'analyse, tome 1,* Gauthier-Villars, Paris, 1968.

[8] J. Dixmier, *Cours de Mathématiques,* Dunod, Paris,1974.

[9] M. Greenberg, *Lectures on Algebraic Topology,* Benjamin, New-York, 1967.

[10] P. Hartman, *Ordinary differential equations,* John Wiley & Sons, New-York, 1964.

[11] N. Jacobson, *Basic Algebra I,* 2nd edition, W. H. Freeman and Co, New-York, 1985.

[12] J. Lafontaine, *Introduction aux variétés différentiables,* PUG, Grenoble, 1996.

[13] S. Lang, *Analysis,* vol. 1, Addison-Wesley Publishing Company, Reading, Massachusetts,1968.

[14] J. La Salle and S. Lefschetz, *Stability by Liapounov's Method,* Academic Press, New-York, 1961.

[15] S. Lefschetz, *Differential equations: Geometric theory,* 2nd edition, Dover, New-York, 1977.

[16] J. Lelong-Ferrand, J. M. Arnaudiès, *Cours de mathématiques-Tome 1, Algèbre,* Dunod Université 511, Paris, 1971.

[17] W. Massey, *Algebraic topology: an Introduction,* Harcourt-Brace-World, New-York, 1967.

[18] J. Milnor, *Topology from the Differential Viewpoint,* The University Press of Virginia, 1965.

[19] S. Mac Lane, G. Birkhoff, *Algèbre, 2/ Les grands théorèmes,* Cahiers scientifiques fasc. XXXVI, Gauthier-Villars, Paris, 1971.

[20] J. Moser, *Stable and random motions in dynamical systems, with spcial emphasis on celestial mechanics,* Annals of Mathematical Studies 77, Princeton University Press, 1973.

[21] J. Palis, W. de Melo, *Geometric Theory of Dynamical Systems, An Introduction,* Springer-Verlag, Berlin, Heidelberg, New-York, 1982.

[22] H. Poincaré, *Leçons de mécanique céleste,* Gauthier-Villars, Paris, 1907 (réédité chez Jacques Gabay, Paris, 2003).

[23] L. S. Pontriaguine, *Équations différentielles ordinaires,* Éditions Mir, Moscou, 1969.

[24] N. Rouche, P. Habets, M. Laloy, *Stability Theory by Liapounov's Direct Method,* Springer-Verlag, Berlin, Heidelberg, New-York, 1977.

[25] R. Roussarie, *Bifurcations of Planar Vector Fields and Hilbert's Sixteenth Problem,* Progress in Mathematics Vol. 164, Birkhauser ed., Basel, 1998.

[26] B.L. van der Waerden, *Modern Algebra,* 2 vol., Springer-Verlag, Berlin, Heidelberg, New-York, 1955.

INDEX

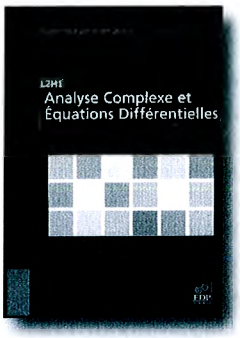
R

S

T

V

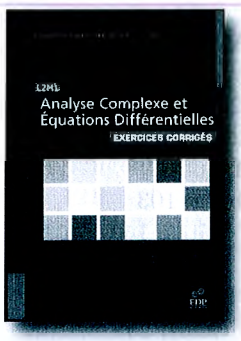
Dans la même collection
Enseignements SUP / Mathématiques



Analyse complexe et équations différentielles

L. Barreira

• 2011 • Public : Étudiants L2-M1 • 978-2-7598-0616-4 • 229 pages • 25 €

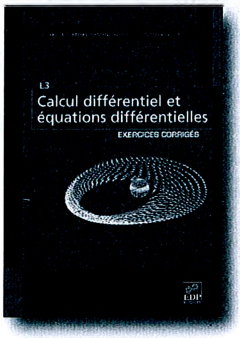


Analyse complexe et équations différentielles

Exercices corrigés

L. Barreira et C. Valls

• 2011 • Public : Étudiants L2-M1 • 978-2-7598-0615-7 • 197 pages • 19 €

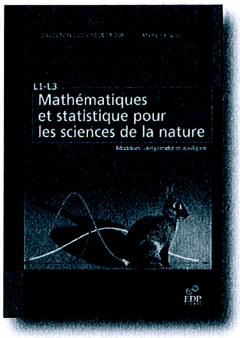


Calcul différentiel et équations différentielles

Exercices corrigés

D. Azé, G. Constans et J.- B. Hiriart-Urruty

• 2010 • Public : Étudiants L3 • 978-2-7598-0413-9 • 240 pages • 24 €



Mathématiques et statistique pour les sciences de la nature

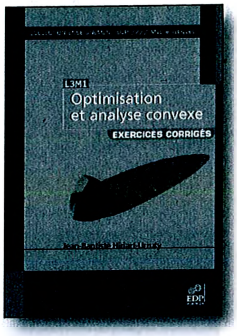
Modéliser, comprendre et appliquer

G. Biau, J. Droniou et M. Herzlich

• 2010 • Public : Étudiants L1-L3 • 978-2-7598-0481-8 • 530 pages • 45 €

Retrouvez ces ouvrages sur www.edition-sciences.com

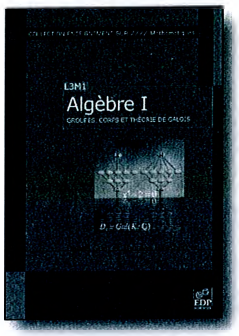
Dans la même collection
Enseignements SUP / Mathématiques



Optimisation et analyse convexe
Exercices corrigés

J.- B. Hiriart-Urruty

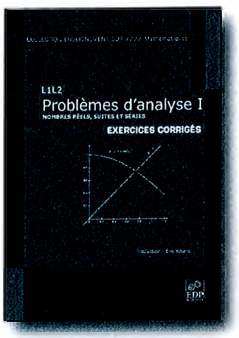
• 2009 • Public : Étudiants L3-M1 • 978-2-7598-0373-6 • 344 pages • 29 €



Algèbre T1
Groupe, corps et théorie de Galois

D. Guin et T. Hausberger

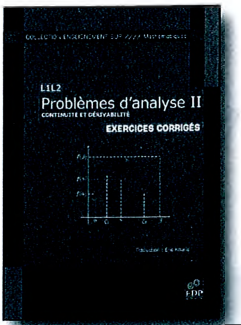
• 2008 • Public : Étudiants L3-M1 • 978-2-86883-974-9 • 478 pages • 37 €



Problèmes d'analyse I - Exercices corrigés
Nombres réels, suites et séries

W.J. Kaczor et M.T. Nowak

• 2008 • Public : Étudiants L1-L2 • 978-2-7598-0058-2 • 380 pages • 30 €



Problèmes d'analyse II - Exercices corrigés
Continuité et dérivabilité

W.J. Kaczor et M.T. Nowak

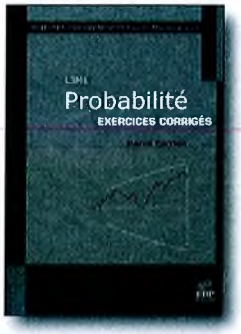
• 2008 • Public : Étudiants L1-L2 • 978-2-7598-0086-5 • 388 pages • 30 €

Dans la même collection
Enseignements SUP / Mathématiques

Problèmes d'analyse III - Exercices corrigés
Intégration

W.J. Kaczor et M.T. Nowak

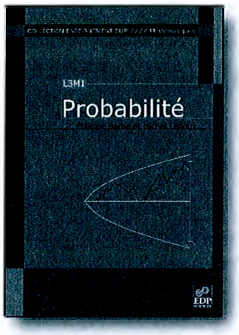
• 2008 • Public : Étudiants L2-L3 • 978-2-7598-0087-2 • 376 pages • 30 €



Probabilité
Exercices corrigés

H. Carrieu

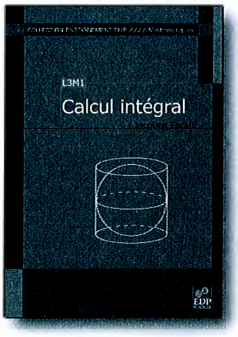
• 2008 • Public : Étudiants L3-M1 • 978-2-7598-0006-3 • 152 pages • 16 €



Probabilité

P. Barbe et M. Ledoux

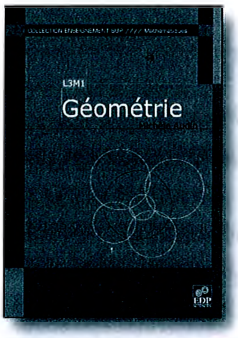
• 2007 • Public : Étudiants L3-M1 • 978-2-86883-931-2 • 256 pages • 26 €



Calcul intégral

J. Faraut

• 2006 • Public : Étudiants L3-M1 • 2-86883-912-6 • 208 pages • 21 €



Géométrie

M. Audin

• 2006 • Public : Étudiants L3-M1 • 2-86883-883-9 • 428 pages • 35 €

Retrouvez ces ouvrages sur www.edition-sciences.com

Des équations différentielles aux systèmes dynamiques I

Théorie élémentaire des équations différentielles
avec éléments de topologie différentielle

Robert Roussarie et Jean Roux

Cet ouvrage est une introduction élémentaire à la théorie des équations différentielles. Il est destiné à illustrer un cours classique sur les équations différentielles dans le cadre d'une licence de mathématiques, mais il peut également servir d'initiation aux notions de base indispensables aux applications.

Une première partie est consacrée à des pré-requis de calcul différentiel et de topologie différentielle : définition des termes et notions de base utilisées par la suite, concernant aussi bien le calcul différentiel dans un espace euclidien que la topologie différentielle.

La deuxième partie est la matière d'un cours classique sur les équations différentielles. Les champs linéaires et les propriétés générales des trajectoires sont donc évidemment exposés. Mais, dans la tradition initiée par Henri Poincaré, on insiste aussi sur les aspects qualitatifs du comportement des solutions, avec l'introduction de la notion de flot d'un champ de vecteurs, qui joue un rôle fondamental car elle sert de base à l'étude essentielle des propriétés de récurrence et de stabilité des orbites. La notion d'application de Poincaré d'une orbite périodique est développée et quelques résultats importants de la théorie qualitative sont démontrés.

Les lecteurs trouveront un développement de cet ouvrage dans le tome II, publié dans la même collection (*Vers la théorie des systèmes dynamiques*).

Robert Roussarie, ancien élève de l'École Polytechnique, a soutenu une thèse en mathématiques sur la théorie des feuilletages. Il a été chercheur au CNRS puis professeur à l'Université de Bourgogne. Il est un spécialiste des équations différentielles (bifurcations des champs de vecteurs du plan, 16e problème de Hilbert, systèmes lents-rapides en dimension 2).

Jean Roux a soutenu une thèse en mathématiques à l'Université de Paris. Il a été ingénieur-chercheur aux Études et Recherches de l'EDF et maître de conférences en analyse numérique aux Ponts et Chaussées. Il est actuellement enseignant en mathématiques appliquées au département Géosciences de l'ENS.

COLLECTION ENSEIGNEMENT SUP //// Mathématiques

9 782759 805129

ISBN : 978-2-7598-0512-9

25 euros

EDP SCIENCES