

M 403 - GEOMETRIE ET EQUATIONS DIFFERENTIELLES

F. BALACHEFF

1. VISION GÉOMÉTRIQUE DU CALCUL DIFFÉRENTIEL

Soit p, n deux entiers non nuls. Dans la suite, I désigne un intervalle ouvert de \mathbb{R} et U un ouvert de \mathbb{R}^p .

1.1. La notion de différentiabilité. Nous introduisons la notion de différentiabilité en généralisant la notion de dérivabilité des fonctions $\gamma : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ aux fonctions $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$.

Rappelons tout d'abord la notion de dérivabilité et son interprétation géométrique. Une fonction $\gamma : I \rightarrow \mathbb{R}^n$ s'appelle une courbe (paramétrée) de \mathbb{R}^n .

Definition 1.1. Une courbe $\gamma : I \rightarrow \mathbb{R}^n$ est dite dérivable en $t \in I$ si la limite

$$\lim_{h \rightarrow 0} \frac{\gamma(t+h) - \gamma(t)}{h}$$

existe. Dans ce cas, on note cette limite $\gamma'(t)$ qui est un vecteur de \mathbb{R}^n appelé vecteur tangent à la courbe au temps t .

La courbe γ est dérivable en t si

$$\frac{\gamma(t+h) - \gamma(t)}{h} = \gamma'(t) + o(1)$$

c'est-à-dire

$$(1.1) \quad \gamma(t+h) = \gamma(t) + \gamma'(t)h + o(h).$$

Observons en particulier que la dérivabilité en t implique la continuité en t .

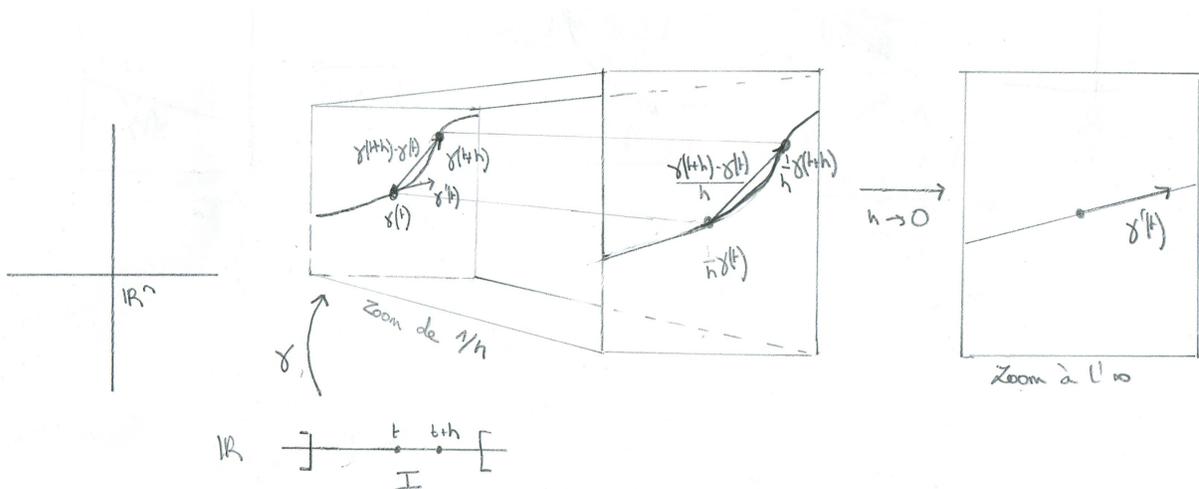


FIGURE 1. Interprétation géométrique de la dérivée.

Autrement dit, *infinitésimalement au voisinage de $\gamma(t)$, la courbe γ est modélisée sur l'application linéaire $h \mapsto \gamma'(t) \cdot h$ dont l'image est une droite si $\gamma'(t) \neq 0$. Une autre manière de le formuler est la suivante : si on zoome à l'infini sur le point $\gamma(t)$, la figure limite que l'on observe est une droite de \mathbb{R}^n de vecteur directeur $\gamma'(t)$, voir la figure 1.*

La dérivabilité en un point est donc une condition forte. Mais on imagine bien que pour avoir des propriétés raisonnables, il faut imposer que, d'une part, ce modèle infinitésimal soit vérifié en tout point et que, d'autre part, ces modèles infinitésimaux varient de manière continue d'un point à un autre (voir figure 2).

Definition 1.2. *La courbe γ est dite de classe C^1 sur I si elle est dérivable sur I (cad en tout point de l'intervalle I) et l'application vecteur tangent $t \in I \mapsto \gamma'(t) \in \mathbb{R}^n$ varie de manière continue sur I .*



FIGURE 2. Interprétation géométrique de la notion C^1 .

Remarquons qu'une courbe C^1 ne subit ainsi ni déchirure, ni pincement, ni d'oscillation condensée.

Ce point de vue est généralisable au cas d'une fonction de plusieurs variables et débouche sur la notion de différentiabilité.

Definition 1.3. *Une fonction $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ est différentiable en $x \in U$ si il existe une application linéaire $df_x \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ telle que*

$$f(x+h) = f(x) + df_x(h) + o(\|h\|)$$

pour tout $h \in \mathbb{R}^p$ tel que $x+h \in U$.

La fonction est dite différentiable sur U si elle l'est en tout point de U .

Ici, $\|\cdot\|$ désigne la norme euclidienne usuelle de \mathbb{R}^p :

$$\|h\| = \sqrt{\sum_{i=1}^p h_i^2},$$

mais les normes sur \mathbb{R}^p étant toutes équivalentes, ce choix n'est pas significatif. Par ailleurs, la notation $o(\|h\|^k)$ avec $k \in \mathbb{N}$ désigne une fonction à valeurs dans \mathbb{R}^n dont la norme puisse s'écrire

$$\|o(\|h\|^k)\| = \|h\|^k \cdot \varepsilon(h)$$

où $\varepsilon(h) \rightarrow 0$ lorsque $h \rightarrow 0$.

Cette formule généralise bien la notion de dérivabilité en comparant avec l'équation (1.1).

On remarque à nouveau que la différentiabilité en x implique la continuité en ce point. Les deux règles classiques de différentiation suivantes découlent de la définition :

- $d(f + \lambda g)_x = df_x + \lambda dg_x$,
- $d(f \circ g)_x = df_{g(x)} \circ dg_x$.

Autrement dit, *infinitésimalement au voisinage de $f(x)$, l'application f est modélisée sur une application linéaire que l'on note df_x* . Mais au fait, à quoi ressemble une application linéaire ?

Proposition 1.1. *Soit L une application linéaire de \mathbb{R}^p dans \mathbb{R}^n . Alors il existe une base B de \mathbb{R}^p et B' de \mathbb{R}^n telle que*

$$\text{mat}_{B,B'}(L) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

où $r \leq \min\{p, n\}$ est le rang de L .

Démonstration. En effet, nous pouvons commencer par choisir une base $(f(e_1), \dots, f(e_r))$ de l'image de f . Ainsi la famille (e_1, \dots, e_r) de \mathbb{R}^p est automatiquement libre. Cette famille libre se complète en une base $(e_1, \dots, e_r, e_{r+1}, \dots, e_p)$ de \mathbb{R}^p . Cette base peut ensuite être modifiée de la manière suivante. Pour chaque $i = r + 1, \dots, p$, il existe des uniques coefficients $\lambda_{i,1}, \dots, \lambda_{i,r} \in \mathbb{R}$ tels que $f(e_i) = \sum_{k=1}^r \lambda_{i,k} f(e_k)$ et on pose alors

$$e'_i = e_i - \sum_{k=1}^r \lambda_{i,k} e_k.$$

Chaque $e'_i \in \ker L$, et la famille $B = (e_1, \dots, e_r, e'_{r+1}, \dots, e'_p)$ est encore une base. On complète alors la famille $(f(e_1), \dots, f(e_r))$ en une base B' de \mathbb{R}^n . □

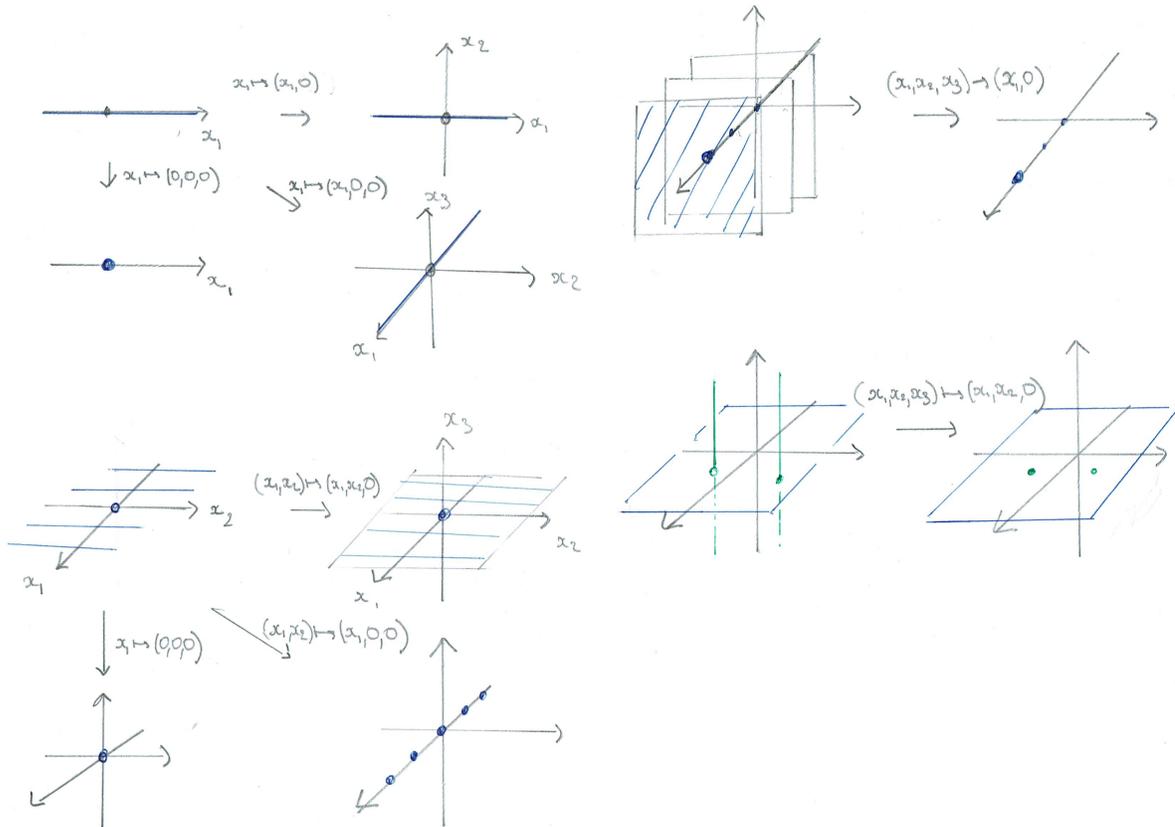


FIGURE 3. Exemples d'applications linéaires.

A un changement de coordonnées près dans l'espace de départ et d'arrivée, L est donc donnée par l'application

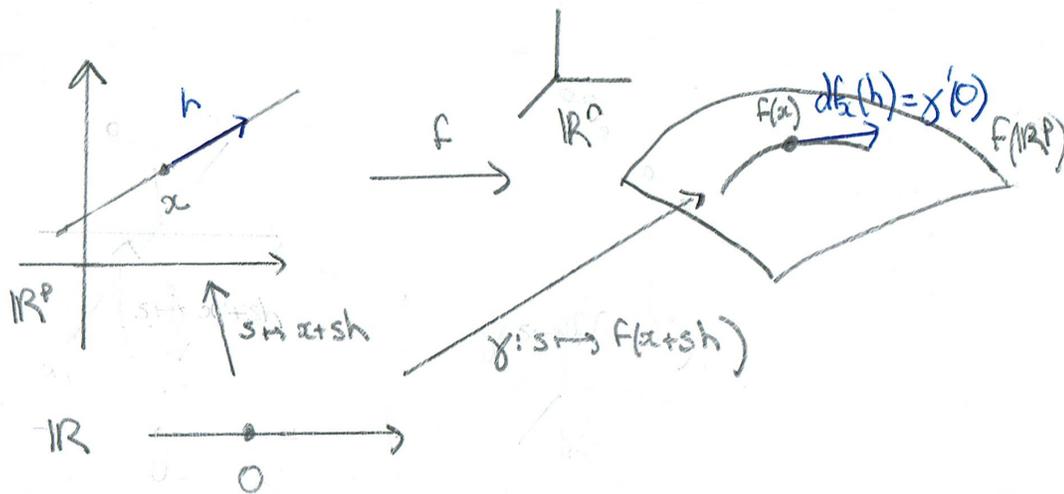
$$(x_1, \dots, x_p) \mapsto (x_1, \dots, x_r, 0, \dots, 0).$$

Observons que cela revient à inclure et contracter des sous-espaces affines les uns dans les autres, voir figure 3.

Si maintenant ce modèle infinitésimal est satisfait en chaque point du domaine de f et varie de manière raisonnable (*i.e.* continue), la fonction f sera dite de classe C^1 et aura de bonnes propriétés globales. L'application ne subira alors ni déchirure, ni pincement ou autre catastrophe, mais certaines compressions sont possibles et elles correspondront à la situation où le modèle linéaire infinitésimal contracte certains sous-espaces. Mais avant d'expliquer plus précisément la notion de classe C^1 dans le cas de plusieurs variables, expliquons une implication géométrique de la notion de différentiabilité.

Observation 1.1. Si f est différentiable en x , alors pour tout $h \neq 0 \in \mathbb{R}^p$, l'image de la droite $s \mapsto x + s \cdot h$ de \mathbb{R}^p passant par x et de vecteur tangent h en 0 est une courbe $s \in \mathbb{R} \mapsto f(x + s \cdot h)$ de \mathbb{R}^n passant par $f(x)$ et dérivable en 0 de vecteur tangent $df_x(h)$.

Voici un schéma de la situation.



Démonstration. C'est équivalent de montrer que

$$df_x(h) = \lim_{s \rightarrow 0} \frac{f(x + s \cdot h) - f(x)}{s}$$

où $s \in \mathbb{R}$.

En effet, d'après la définition de différentiabilité

$$\lim_{s \rightarrow 0} \left\| \frac{f(x + s \cdot h) - f(x) - df_x(s \cdot h)}{s \cdot h} \right\| = \lim_{s \rightarrow 0} \frac{o(\|s \cdot h\|)}{\|s \cdot h\|} = 0.$$

Donc

$$\lim_{s \rightarrow 0} \left\| \frac{f(x + s \cdot h) - f(x)}{s} - df_x(h) \right\| \cdot \frac{1}{\|h\|} = 0$$

ce qui permet de conclure comme $h \neq 0$. □

Lorsqu'elle existe, la quantité

$$\lim_{s \rightarrow 0} \frac{f(x + s \cdot h) - f(x)}{s}$$

s'appelle la dérivée de Gâteaux de f en x dans la direction h et sera notée $f'(x, h)$.

Attention! L'existence en toute direction h au point x d'une dérivée de Gâteaux de f n'implique pas la différentiabilité de f . On peut montrer que,

Proposition 1.2. Si $f'(x, h)$ est définie, pour tout $\lambda \in \mathbb{R}$, on a

$$f'(x, \lambda h) = \lambda f'(x, h).$$

Mais on peut avoir que $f'(x, h_1 + h_2) \neq f'(x, h_1) + f'(x, h_2)$. Par exemple, si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ désigne l'application définie pour $(x_1, x_2) \neq (0, 0)$ par

$$f(x_1, x_2) = \frac{x_1^3}{x_1^2 + x_2^2}$$

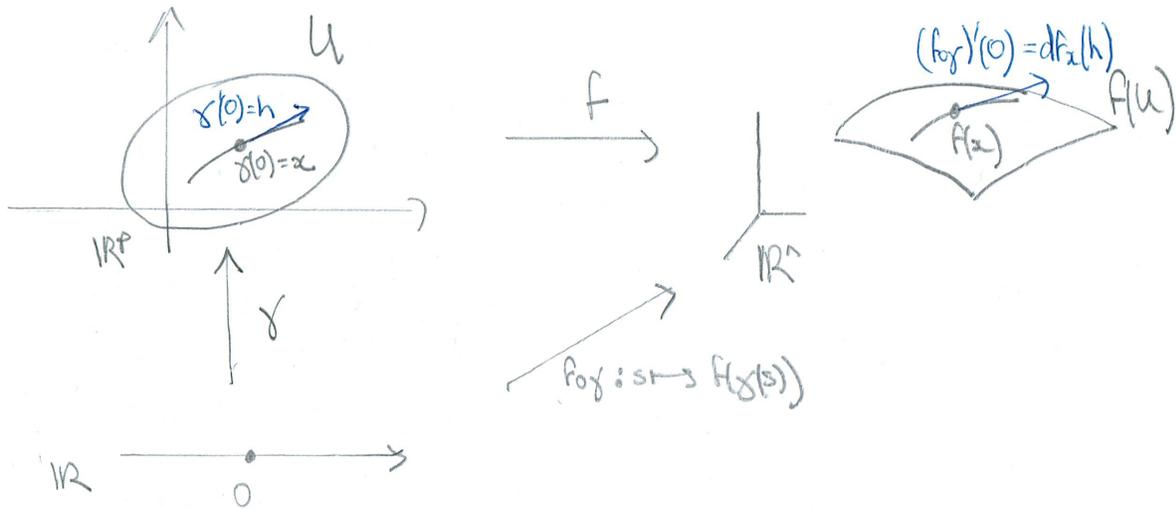
et prolongée par 0 en $(0, 0)$, alors l'application $h \mapsto f'(0, h)$ n'est pas linéaire (voir exercice).

Plus généralement, on peut montrer l'implication géométrique suivante :

Proposition 1.3. Si f est une application différentiable en x , pour toute courbe $\gamma :]-\varepsilon, \varepsilon[\rightarrow U \subset \mathbb{R}^p$ dérivable en 0 telle que $\gamma(0) = x$ et $\gamma'(0) = h$, la courbe $s \in \mathbb{R} \mapsto f \circ \gamma(s)$ de \mathbb{R}^n passant par $f(x)$ est dérivable en 0 de vecteur tangent $df_x(h)$:

$$\left. \frac{d}{ds} \right|_0 f(\gamma(s)) = df_x(h).$$

Voici un schéma de la situation.



Tout comme pour les fonctions d'une variable réelle, la notion de différentiabilité n'a d'intérêt que si le modèle infinitésimal se comporte convenablement lorsque l'on modifie le point x dans U .

Definition 1.4. $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ est dite (de classe) C^1 sur U si elle est différentiable sur U et l'application différentielle

$$df : x \in U \mapsto df_x \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n) \simeq \mathbb{R}^{pn}$$

est continue.

L'application différentielle, en tant qu'application linéaire, peut être représentée dans les bases canoniques de \mathbb{R}^p et \mathbb{R}^n par une matrice de taille $n \times p$ à coefficients réels appelée matrice jacobienne

$$Jac_x(f) = Mat_{(e_1, \dots, e_p), (e'_1, \dots, e'_n)}(df_x) = (df_x(e_j) \mid e'_i)_{ij},$$

(e_1, \dots, e_p) et (e'_1, \dots, e'_n) désignant les bases canoniques respectives de \mathbb{R}^p et \mathbb{R}^n , et $(\cdot \mid \cdot)$ le produit scalaire euclidien. L'application différentielle est alors continue si et seulement si les coefficients de la matrice jacobienne le sont.

Observer que si on note $f = (f_1, \dots, f_n)$ alors $(df_x(e_j)) \mid e'_i = d(f_i)_x(e_j)$ quantité que l'on note usuellement

$$\frac{\partial f_i}{\partial x_j}(x).$$

Autrement dit, si f est C^1 sur U , alors les dérivées partielles de f

$$x \mapsto df_x(e_j) = \frac{\partial f}{\partial x_j}(x) = \left(\frac{\partial f_1}{\partial x_j}(x), \dots, \frac{\partial f_n}{\partial x_j}(x) \right)$$

existent et sont continues. De manière étonnante, cette condition nécessaire sur les dérivées partielles est suffisante. Noter que l'on a alors

$$df_x(h) = \sum_{j=1}^p \frac{\partial f}{\partial x_j}(x) \cdot h_j.$$

Attention! Il est important ici de rappeler la signification géométrique des dérivées partielles de f : il s'agit juste des vecteurs tangents au temps 0 des courbes images par f des droites définissant le système de coordonnées local en x (voir figure 4). Les dérivées partielles peuvent donc exister indépendamment de la différentiabilité de la fonction en question puisqu'il s'agit en d'autres termes des dérivées de Gâteaux en x selon les vecteurs e_1, \dots, e_p .

Théorème 1.1. f est C^1 si et seulement si les dérivées partielles de f

$$x \mapsto \frac{\partial f}{\partial x_j}(x)$$

existent et sont continues.

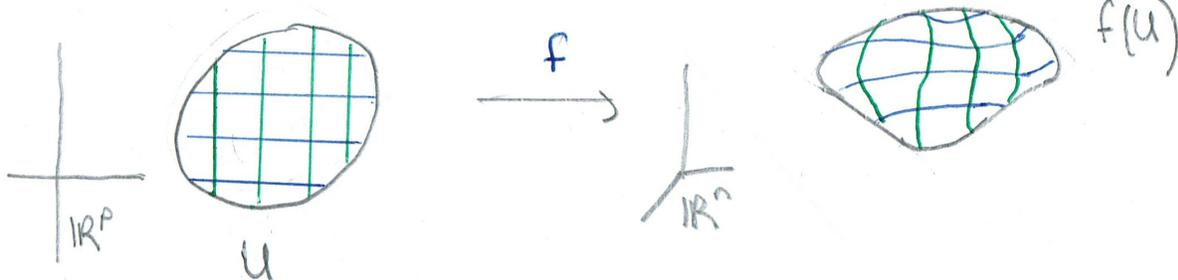


FIGURE 4. Application $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ de classe C^1 et de rang 2.

Démonstration. Supposons que les dérivées partielles de f existent et soient continues. Il suffit de montrer que f est alors différentiable sur U . La matrice jacobienne étant définie en x , le candidat naturel pour être la différentielle de f est l'application

$$h \in \mathbb{R}^p \mapsto Jac_x(f) \cdot h = \sum_{j=1}^p \frac{\partial f}{\partial x_j}(x) \cdot h_j.$$

On a pour tout $h \in \mathbb{R}^p$ tel que $x + h \in U$

$$f(x + h) - f(x) - \sum_{j=1}^p \frac{\partial f}{\partial x_j}(x) \cdot h_j =$$

$$\begin{aligned} & f(x_1 + h_1, \dots, x_{p-1} + h_{p-1}, x_p + h_p) - f(x_1 + h_1, \dots, x_{p-1} + h_{p-1}, x_p) - \frac{\partial f}{\partial x_p}(x_1 + h_1, \dots, x_{p-1} + h_{p-1}, x_p) \cdot h_p \\ & + \frac{\partial f}{\partial x_p}(x_1 + h_1, \dots, x_{p-1} + h_{p-1}, x_p) \cdot h_p - \frac{\partial f}{\partial x_p}(x_1, \dots, x_{p-1}, x_p) \cdot h_p \\ & + f(x_1 + h_1, \dots, x_{p-1} + h_{p-1}, x_p) - f(x_1, \dots, x_{p-1}, x_p) - \sum_{j=1}^{p-1} \frac{\partial f}{\partial x_j}(x) \cdot h_j \end{aligned}$$

La première partie et la dernière partie sont respectivement un $o(\|h_p\|)$ (cas de dimension 1) et un $o(\|(h_1, \dots, h_{p-1})\|)$ (par induction), et donc a fortiori leur somme est un $o(\|h\|)$. La seconde partie s'écrit

$$\left(\frac{\partial f}{\partial x_p}(x_1 + h_1, \dots, x_{p-1} + h_{p-1}, x_p) - \frac{\partial f}{\partial x_p}(x_1, \dots, x_{p-1}, x_p) \right) \cdot h_p = \varepsilon(h) \cdot h_p = o(\|h\|)$$

les dérivées partielles étant continues. □

Notions de différentiabilité d'ordre supérieur. Nous pouvons identifier $\mathcal{L}(\mathbb{R}^p, \mathbb{R}^n) \simeq \mathbb{R}^{pn}$, et interpréter ainsi la différentielle comme une application de \mathbb{R}^p dans \mathbb{R}^{pn} pour laquelle nous disposons de la notion de classe C^1 . Ainsi f est dite de classe C^2 si sa différentielle df est de classe C^1 . On introduit alors de manière inductive la notion de classe C^k pour tout entier k et on dit que f est de classe C^∞ si f est C^k pour tout entier k .

1.2. Comportements locaux déterminés par une propriété ponctuelle. Nous allons voir que certaines propriétés locales d'une fonction de classe C^1 sont complètement déterminées par sa différentielle en un point.

La philosophie est la suivante : étant donnée une application de classe C^1 , si sa différentielle en un point vérifie une propriété qui est stable par perturbation dans l'espace des applications linéaires, alors l'application elle-même vérifie cette propriété au voisinage de ce point.

1.2.1. *Théorème des accroissements finis.* Commençons par un premier exemple.

Théorème 1.2. (des accroissements finis) Soit $U \subset \mathbb{R}^p$ un ouvert convexe. Supposons que $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ soit une application C^1 telle qu'il existe $k > 0$ satisfaisant

$$\|df_x\| < k$$

pour tout $x \in U$. Alors

$$\|f(y) - f(x)\| < k\|y - x\|$$

pour tous $x, y \in U$.

Rappelons ici que la norme classique d'une application linéaire est définie par

$$\|L\| = \max\{\|L(h)\| \mid \|h\| = 1\}.$$

C'est donc la norme du plus grand vecteur dans l'image de la sphère unité. Par définition pour tout $h \in \mathbb{R}^p$ nous avons

$$\|L(h)\| \leq \|L\| \cdot \|h\|.$$

Ce résultat entre bien dans le cadre de notre philosophie, la propriété $\|L\| < k$ étant localement stable pour les applications linéaires : si f est de classe C^1 et vérifie $\|df_{x_0}\| < k$ en un point x_0 (c'est-à-dire $\|df_{x_0}(y) - df_{x_0}(x)\| < k\|y - x\|$ pour tout $x, y \in \mathbb{R}^p$), alors cela reste vrai dans un certain voisinage de x_0 et par le TAF $\|f(y) - f(x)\| < k\|y - x\|$ pour tout x, y dans ce voisinage de x_0 .

Démonstration. On se fixe $x, y \in U$ et on considère la courbe

$$\begin{aligned} \gamma : [0, 1] &\rightarrow \mathbb{R}^n \\ t &\mapsto f(x + t(y - x)) \end{aligned}$$

qui est bien définie comme $x + t(y - x)$ appartient à U pour tout $t \in [0, 1]$ par convexité. Cette courbe est de classe C^1 comme composition de deux applications C^1 , et on voit que

$$\|\gamma'(t)\| = \|df_{x+t(y-x)}(y - x)\| < k\|y - x\|.$$

Maintenant,

$$\|f(y) - f(x)\| = \|\gamma(1) - \gamma(0)\| = \left\| \int_0^1 \gamma'(t) dt \right\| \leq \int_0^1 \|\gamma'(t)\| dt < k\|y - x\|$$

en utilisant l'inégalité de Jensen appliquée à la fonction convexe $h \mapsto \|h\|$. □

Remarque 1.1. Nous avons ici invoqué l'inégalité

$$\left\| \int_0^1 u(t) dt \right\| \leq \int_0^1 \|u(t)\| dt,$$

pour toute fonction $u : [0, 1] \rightarrow \mathbb{R}^n$. Nous allons donc la démontrer. Tout d'abord, il faut remarquer que la norme peut être définie comme le maximum de certaines formes linéaires :

$$\|x\| = \max_{\|a\|=1} a_1 x_1 + \dots + a_n x_n.$$

En effet, pour tout $a \in \mathbb{R}^n$ tel que $\|a\| = 1$, nous avons

$$a_1 x_1 + \dots + a_n x_n = \langle a, x \rangle \leq \|a\| \|x\| = \|x\|$$

par l'inégalité de Cauchy-Schwarz avec égalité pour $a = x/\|x\|$.

Ensuite, on utilise la linéarité de l'intégrale

$$\begin{aligned} \left\| \int_0^1 u(t) dt \right\| &= \max_{\|a\|=1} \left(a_1 \int_0^1 u_1(t) dt + \dots + a_n \int_0^1 u_n(t) dt \right) \\ &= \max_{\|a\|=1} \int_0^1 (a_1 u_1(t) + \dots + a_n u_n(t)) dt \\ &\leq \int_0^1 \max_{\|a\|=1} (a_1 u_1(t) + \dots + a_n u_n(t)) dt \\ &\leq \int_0^1 \|u(t)\| dt. \end{aligned}$$

En fait, cette preuve marche pour toute fonction convexe φ car celles-ci peuvent toujours être décrites comme le maximum de certaines fonctions affines (penser au cas de dimension 1 pour se faire une idée). C'est ce qu'on appelle l'*inégalité de Jensen*.

1.2.2. *Théorème d'inversion locale.* Un second exemple de propriété des applications linéaires stable par perturbation est l'inversibilité.

Proposition 1.4. *Si $L_0 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ est inversible, alors il existe un voisinage ouvert de L_0 dans $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ formé exclusivement d'applications linéaires inversibles.*

Démonstration. La topologie sur $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ est induite par l'identification $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \simeq \mathbb{R}^{n^2}$ correspondant à la représentation matricielle dans la base canonique de \mathbb{R}^n . Ainsi, nous voyons que l'application

$$L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \mapsto \det_{can}(L)$$

est continue car le déterminant est une fonction polynômiale des coefficients de la matrice. Comme $\det(L_0) \neq 0$, on en déduit qu'il existe $\varepsilon > 0$ tel que l'intervalle $]\det_{can}(L_0) - \varepsilon, \det_{can}(L_0) + \varepsilon[$ ne contienne pas 0. L'image réciproque de cet intervalle par l'application \det_{can} est donc un ouvert de $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ formé exclusivement d'éléments inversibles. \square

En particulier, si $\{L_t\}_{t \in [0,1]}$ une déformation par des éléments de $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ d'une application L_0 inversible alors, pour t suffisamment petit, l'application linéaire L_t est aussi inversible.

Dire que $\{L_t\}$ est une déformation de L_0 signifie que l'application

$$t \in [0, 1] \mapsto L_t \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$$

est continue. Autrement dit, les coefficients des matrices L_t dans une base fixée dépendent continûment du paramètre t .

Nous en déduisons que si la différentielle d'une application f de classe C^1 est inversible en un point x , pour y suffisamment proche de x la différentielle df_y sera elle-aussi inversible. Ceci va impliquer que l'application f est localement inversible.

Avant d'énoncer plus soigneusement ce résultat, nous avons besoin d'introduire la notion de difféomorphisme.

Definition 1.5. *Une application $f : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^n$ est un difféomorphisme de U sur V si f est de classe C^1 , réalise une bijection de U sur V , et son inverse $f^{-1} : V \rightarrow U$ est aussi de classe C^1 .*

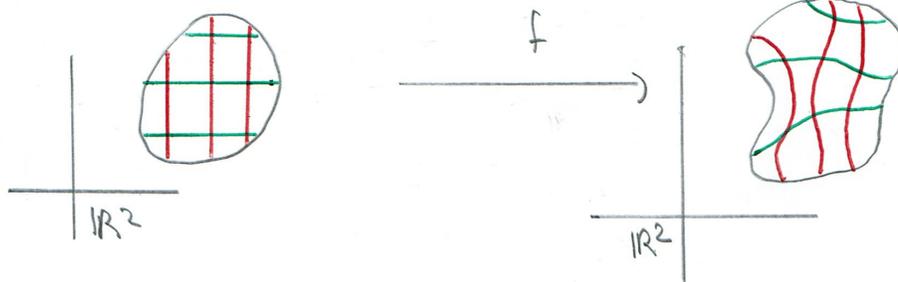


FIGURE 5. Représentation locale d'un C^1 -difféomorphisme pour $n = 2$.

Un difféomorphisme est donc tout simplement un élément inversible pour la loi de composition dans l'espace des applications de classe C^1 : f est C^1 , f^{-1} est C^1 et $f^{-1} \circ f = id_U$.

Il est facile d'observer que si f est un difféomorphisme, alors en tout point $x \in U$ sa différentielle df_x est inversible et l'on a

$$df_{f(x)}^{-1} = (df_x)^{-1}.$$

C'est une conséquence de la règle de différentiation $d(f \circ g)_x = df_{g(x)} \circ dg_x$ appliquée à la relation $f^{-1} \circ f = id_U$.

Réciproquement, d'après notre philosophie expliquée plus haut, si la différentielle d'une application de classe C^1 est inversible en un point, alors l'application elle-même est inversible au niveau local.

Théorème 1.3. (*d'inversion locale*) Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ une application de classe C^1 et $x_0 \in U$ telle que df_{x_0} soit inversible. Alors il existe un voisinage ouvert U_0 de x_0 et un voisinage ouvert V_0 de $f(x_0)$ tels que f soit un difféomorphisme de U_0 sur V_0 .

Démonstration. On procède par étapes.

Réduction du problème. Quitte à remplacer f par

$$x \mapsto df_{x_0}^{-1}(f(x_0 + x) - f(x_0)),$$

nous pouvons supposer que $x_0 = 0$, $f(0) = 0$ et $df_0 = id_{\mathbb{R}^n}$.

Construction de l'application réciproque f^{-1} . Nous allons utiliser pour cela le théorème du point fixe que voici.

Théorème 1.4. (*du point fixe*) Soit F un fermé non vide de \mathbb{R}^n et $T : F \rightarrow F$ une application contractante : il existe $0 \leq k < 1$ tel que pour tout $x, y \in F$ on ait

$$\|T(x) - T(y)\| \leq k \cdot \|x - y\|.$$

Alors il existe un unique point $x_0 \in F$ tel que $T(x_0) = x_0$.

Etant donné y proche de 0, résoudre l'équation $y = f(x)$ revient à trouver un point fixe de l'application

$$h(x) = y + x - f(x).$$

Etudions tout d'abord les propriétés de l'application

$$g : x \in U \mapsto x - f(x).$$

Cette application est C^1 et vérifie $g(0) = 0$ ainsi que $dg_0 = 0$. Par continuité de l'application différentielle, il existe $r > 0$ suffisamment petit tel que pour tout $x \in B(0, 2r) \subset U$ on ait

$$\|dg_x\| < \frac{1}{2}.$$

Le théorème des accroissements finis nous donne alors que pour tout $x, x' \in B(0, 2r)$ nous avons

$$\|g(x) - g(x')\| < \frac{1}{2}\|x - x'\|.$$

En particulier, $\|g(x)\| < \frac{1}{2}\|x\|$ en choisissant $x' = 0$ et donc $g(B(0, 2r)) \subset B(0, r)$ ce qui implique par continuité que $g(\overline{B(0, 2r)}) \subset \overline{B(0, r)}$.

Nous en déduisons facilement que pour tout $y \in B(0, r)$

$$h(\overline{B(0, 2r)}) \subset \overline{B(0, 2r)},$$

comme $\|h(x)\| \leq \|y\| + \|g(x)\|$. Or, par ailleurs, pour tous $x, x' \in B(0, 2r)$ nous avons

$$\|h(x) - h(x')\| = \|g(x) - g(x')\| < \frac{1}{2}\|x - x'\|,$$

formule qui reste vraie par continuité pour $x, x' \in \overline{B(0, 2r)}$ en remplaçant l'inégalité stricte par une inégalité large.

Ainsi l'application

$$h : \overline{B(0, 2r)} \rightarrow \overline{B(0, 2r)}$$

est bien contractante et nous pouvons lui appliquer le théorème du point fixe : il existe un unique point $x \in \overline{B}(0, 2r)$ tel que

$$h(x) = x \Leftrightarrow y = f(x).$$

Par construction, ce point x appartient à la boule ouverte $B(0, 2r)$ comme

$$\|x\| = \|h(x)\| \leq \|y\| + \|g(x)\| < 2r.$$

Par conséquent, la restriction

$$f : \underbrace{B(0, 2r) \cap f^{-1}(B(0, r))}_{:=U_0} \rightarrow \underbrace{B(0, r)}_{:=V_0}$$

est une bijection dont on notera f^{-1} l'application réciproque. Noter que U_0 est bien un voisinage ouvert de 0

L'application f^{-1} est lipschitzienne. Soient $x, x' \in U_0$ et $y, y' \in V_0$ tels que $y = f(x)$ et $y' = f(x')$. Alors

$$\|f^{-1}(y) - f^{-1}(y')\| = \|x - x'\| = \|y + g(x) - y' - g(x')\| \leq \|y - y'\| + \frac{1}{2}\|x - x'\|$$

d'où

$$\|f^{-1}(y) - f^{-1}(y')\| = \|x - x'\| \leq 2\|y - y'\|.$$

L'application f^{-1} est de classe C^1 . L'application df_0 est inversible, donc quitte à diminuer r , nous pouvons supposer que df_x l'est également pour tout $x \in U_0$. Nous commençons par prouver que $(df_x)^{-1}$ est bien la différentielle de f^{-1} en x .

Comme $\|(df_0)^{-1}\| = \|id_{\mathbb{R}^n}\| = 1$, on peut supposer que

$$\|(df_x)^{-1}\| \leq 2$$

pour $x \in B(0, 2r)$ quitte à diminuer r à nouveau.

Fixons donc $y \in B(0, r)$. Pour $k \in \mathbb{R}^n$ tel que $y + k \in B(0, r)$, on note $x = f^{-1}(y)$ puis on pose $h = f^{-1}(y + k) - f^{-1}(y)$. Ainsi $x + h = f^{-1}(y + k)$ et on calcule alors

$$\begin{aligned} \|f^{-1}(y + k) - f^{-1}(y) - (df_x)^{-1}(k)\| &= \|x + h - x - (df_x)^{-1}(f(x + h) - f(x))\| \\ &= \|(df_x)^{-1}(df_x(h) - f(x + h) + f(x))\| \\ &\leq \|(df_x)^{-1}\| \cdot \|df_x(h) - f(x + h) + f(x)\| \\ &\leq 2 \cdot \|df_x(h) - f(x + h) + f(x)\| = o(\|h\|). \end{aligned}$$

Or $\|h\| = \|f^{-1}(y + k) - f^{-1}(y)\| \leq 2\|k\|$ comme f^{-1} est 2-lipschitzienne, et donc

$$\|f^{-1}(y + k) - f^{-1}(y) - (df_x)^{-1}(k)\| = o(\|k\|).$$

Ainsi, l'application f^{-1} est différentiable en $y = f(x)$ et

$$d(f^{-1})_y = (df_x)^{-1} = (df_{f^{-1}(y)})^{-1}.$$

Maintenant, l'application

$$u \in Gl(\mathbb{R}^n) \mapsto u^{-1} \in Gl(\mathbb{R}^n)$$

est continue (penser à la formule de Cramer), et donc la composition

$$y \xrightarrow{f^{-1}} f^{-1}(y) \xrightarrow{df} df_{f^{-1}(y)} \xrightarrow{(\cdot)^{-1}} (df_{f^{-1}(y)})^{-1} = d(f^{-1})_y$$

est continue.

Nous déduisons que l'application f^{-1} est de classe C^1 et f est donc bien un difféomorphisme de U_0 sur V_0 . \square

On peut déduire du théorème d'inversion local certains corollaires dont les preuves seront traitées en exercice.

Corollaire 1.1. (Théorème d'inversion globale) f est un difféomorphisme de U sur $f(U)$ si et seulement si f est injective, de classe C^1 et df_x est inversible pour tout $x \in U$.

Corollaire 1.2. (Théorème des fonctions implicites) Soit $f : U \subset \mathbb{R}^p \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ une application de classe C^1 . Supposons qu'il existe un point $(x_0, y_0) \in U$ tel que la matrice

$$\left(\frac{\partial f}{\partial y_1}(x_0, y_0) \dots \frac{\partial f}{\partial y_n}(x_0, y_0) \right)$$

soit inversible. Alors il existe un voisinage U' ouvert de (x_0, y_0) dans U de la forme $U' = V'_1 \times V'_2 \subset \mathbb{R}^p \times \mathbb{R}^n$ et une application $\varphi : V'_1 \rightarrow V'_2$ de classe C^1 telle que

$$(x, y) \in U' \text{ avec } f(x, y) = f(x_0, y_0) \Leftrightarrow y = \varphi(x).$$

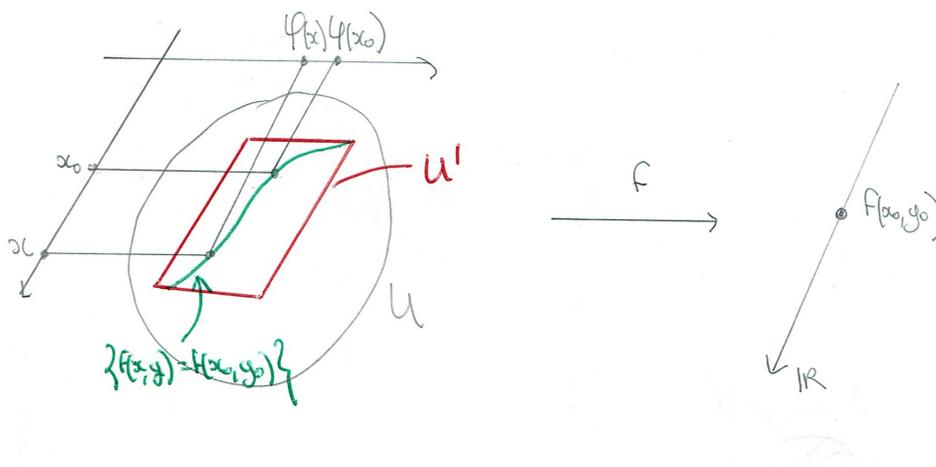


FIGURE 6. Théorème des fonctions implicites.

1.2.3. *Redressement des immersions et des submersions.* Voici enfin deux derniers exemples de propriété d'une application linéaire stable par perturbation : l'injectivité et la surjectivité.

Proposition 1.5. Soit $\{L_t\}_{t \in [0,1]}$ une déformation par des éléments de $\mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ d'une application L_0 injective (respectivement surjective). Alors, pour t suffisamment petit, l'application linéaire L_t est aussi injective (respectivement surjective).

Nous laissons cette proposition comme exercice. Remarquer que si L_0 est injective, alors nécessairement $p \leq n$, et si L_0 est surjective, nous avons $p \geq n$.

Nous introduisons les notions suivantes.

Definition 1.6. Soit $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ une application de classe C^1 . On dit que f est une immersion (resp. submersion) en x si sa différentielle df_x est injective (resp. surjective).

Les deux résultats suivants rentrent dans le cadre de notre philosophie, et décrivent aussi complètement les immersions et submersions du point de vue local à changement de coordonnées près.

Théorème 1.5. Soit $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ une application de classe C^1 , qui soit une immersion en $x \in U$. Alors il existe des voisinages ouverts U' de x et V' de $f(x)$ satisfaisant $f(U') \subset V'$, ainsi qu'un difféomorphisme $\varphi : V' \rightarrow \varphi(V')$ tel que

$$\varphi \circ f(x'_1, \dots, x'_p) = (x'_1, \dots, x'_p, 0, \dots, 0)$$

pour tout $x' = (x'_1, \dots, x'_p) \in U'$.

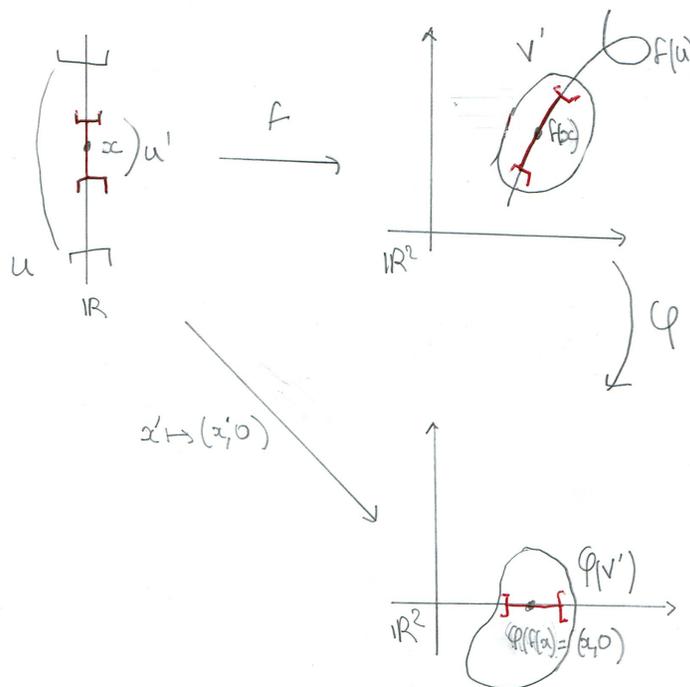


FIGURE 7. Redressement d'une immersion.

Ce résultat nous dit qu'une immersion peut être redressée (en modifiant l'espace d'arrivée par un difféomorphisme) en l'application linéaire injective canonique correspondante. Une courbe immergée avec un point double fait comprendre que l'injectivité de l'application peut n'être que locale.

Démonstration. Comme df_x est injective, nous avons que $p \leq n$. Notons pour $i = 1 \dots p$ par

$$v_i = df_x(e_i) \in \mathbb{R}^n$$

la i ème dérivée partielle de f . L'injectivité de df_x implique alors que cette famille est libre, et par le théorème de la base incomplète, nous pouvons choisir $n - p$ vecteurs v_{p+1}, \dots, v_n tels que (v_1, \dots, v_n) soit une base de \mathbb{R}^n . Nous définissons alors $\psi : U \times \mathbb{R}^{n-p} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ en posant

$$\psi(x'_1, \dots, x'_n) = f(x'_1, \dots, x'_p) + x'_{p+1}v_{p+1} + \dots + x'_n v_n.$$

Cette application vérifie $\psi(x, 0_{\mathbb{R}^{n-p}}) = f(x)$. Elle est aussi de classe C^1 . Pour cela il suffit d'observer que pour tout $x' = (x'_1, \dots, x'_p, x'_{p+1}, \dots, x'_n) \in U \times \mathbb{R}^{n-p}$

$$Jac_{x'}(\psi) = (Jac_{(x'_1, \dots, x'_p)}(f), v_{p+1}, \dots, v_n),$$

matrice qui dépend continûment du point x' . On remarque également que $Jac_{(x, 0_{\mathbb{R}^{n-p}})}(\psi) = (v_1, \dots, v_n)$ et donc que $d\psi_x$ est inversible. Par le théorème d'inversion local, ψ est un difféomorphisme d'un voisinage ouvert de $(x, 0_{\mathbb{R}^{n-p}})$ sur un voisinage ouvert V' de $\psi(x, 0_{\mathbb{R}^{n-p}}) = f(x)$.

Par construction,

$$\psi(x'_1, \dots, x'_p, 0, \dots, 0) = f(x'_1, \dots, x'_p)$$

et donc $\psi^{-1} \circ f(x'_1, \dots, x'_p) = (x'_1, \dots, x'_p, 0, \dots, 0)$. Il suffit alors de poser $\varphi = \psi^{-1}$. \square

Nous démontrerons en exercice le résultat correspondant pour les submersions qui s'énonce ainsi.

Théorème 1.6. Soit $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ une application de classe C^1 , qui soit une submersion en $x \in U$. Alors il existe un voisinage ouvert $U' \subset U$ de x , ainsi qu'un difféomorphisme $\varphi : U' \rightarrow \varphi(U')$ tel que

$$f \circ \varphi^{-1}(x'_1, \dots, x'_p) = (x'_1, \dots, x'_n)$$

pour tout $x' = (x'_1, \dots, x'_p) \in U'$.

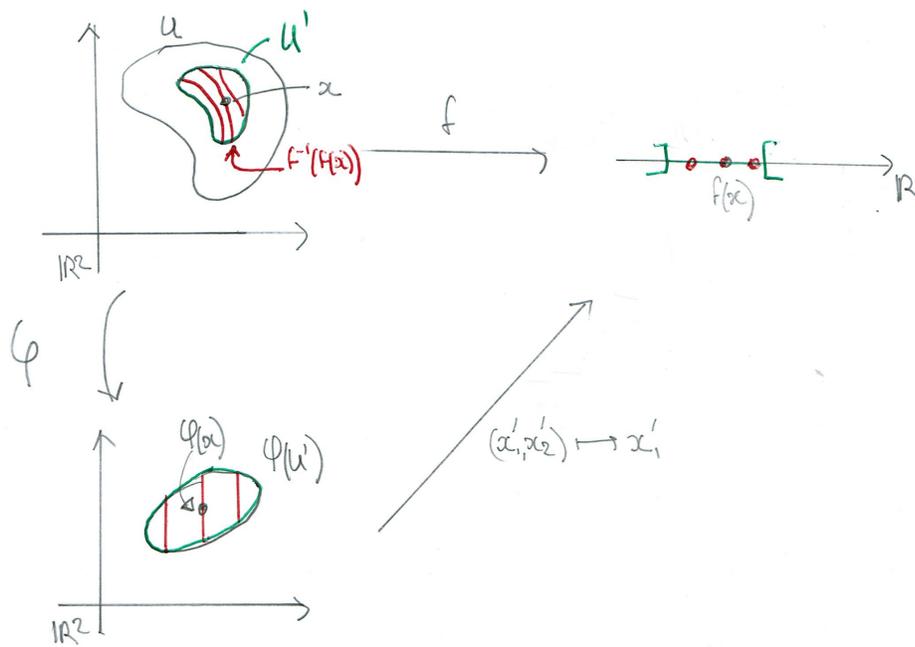


FIGURE 8. Redressement d'une submersion.

2. PREMIÈRE INCURSION DANS LE MONDE DES ÉQUATIONS DIFFÉRENTIELLES

2.1. Un peu de vocabulaire. Nous commençons par définir le terme d'équation différentielle ordinaire et montrer comment se ramener en toute généralité au cas des équations différentielles ordinaires du premier ordre.

Definition 2.1. Soit $k \in \mathbb{N}^*$. On appelle équation différentielle ordinaire d'ordre k sur \mathbb{R}^n une équation de la forme

$$\frac{d^k x}{dt^k} = f \left(t, x, \dots, \frac{d^{k-1} x}{dt^{k-1}} \right)$$

où $f : I \times U_0 \times U_1 \times \dots \times U_{k-1} \rightarrow \mathbb{R}^n$ désigne une fonction continue, I étant un intervalle de \mathbb{R} et les U_i des ouverts de \mathbb{R}^n .

On appelle alors solution de cette équation toute application $x : J \subset I \rightarrow \mathbb{R}^n$ satisfaisant l'égalité ci-dessus. Ceci implique que x est de classe C^k et que $\frac{d^i x}{dt^i}(t) \in U_i$ pour tout $t \in J$.

Lorsque f est indépendante de t , l'équation différentielle est dite autonome.

Pour un certain nombre de considérations (celles les plus générales, comme par exemple la question de l'existence de solutions), tout se ramène au cas d'ordre 1 et autonome en vertu du résultat suivant.

Proposition 2.1. Toute solution d'une équation différentielle ordinaire est équivalente à la solution d'une équation différentielle ordinaire d'ordre 1 et autonome qui lui est naturellement associée.

Démonstration. Si $x : J \subset I \rightarrow \mathbb{R}^n$ est une solution de l'équation

$$\frac{d^k x}{dt^k} = f \left(t, x, \dots, \frac{d^{k-1} x}{dt^{k-1}} \right),$$

alors l'application

$$\begin{aligned} y : J &\rightarrow \mathbb{R}^{1+nk} \\ t &\mapsto \left(t, x(t), \dots, \frac{d^{k-1} x}{dt^{k-1}} \right) \end{aligned}$$

est de classe C^1 et satisfait l'équation ordinaire d'ordre 1 et autonome sur \mathbb{R}^{1+nk} suivante

$$\frac{dy}{dt} = F(y(t))$$

avec

$$\begin{aligned} F : I \times U_0 \times \dots \times U_{k-1} &\rightarrow \mathbb{R} \times \mathbb{R}^n \times \dots \times \mathbb{R}^n \\ (s, u_0, \dots, u_{k-2}, u_{k-1}) &\mapsto (1, u_1, \dots, u_{k-1}, f(s, u_0, \dots, u_{k-1})) \end{aligned}$$

En effet,

$$\begin{aligned} \frac{dy}{dt} &= \frac{d}{dt} \left(t, x(t), \dots, \frac{d^{k-2} x}{dt^{k-2}}, \frac{d^{k-1} x}{dt^{k-1}} \right) = \left(1, x'(t), \dots, \frac{d^{k-1} x}{dt^{k-1}}, \frac{d^k x}{dt^k} \right) \\ &= \left(1, x'(t), \dots, \frac{d^{k-1} x}{dt^{k-1}}, f \left(t, x, \dots, \frac{d^{k-1} x}{dt^{k-1}} \right) \right) = F \left(t, x(t), \dots, \frac{d^{k-2} x}{dt^{k-2}}, \frac{d^{k-1} x}{dt^{k-1}} \right) \\ &= F(y(t)). \end{aligned}$$

Réciproquement, toute solution de cette équation ordinaire d'ordre 1 et autonome fournit une solution de l'équation différentielle ordinaire initialement considérée : si $y(t) = (s(t), u_0(t), \dots, u_{k-1}(t))$ est une solution de l'équation ordinaire d'ordre 1 et autonome ci-dessus définie sur un intervalle $J' \subset \mathbb{R}$, cela implique que

$$u'_i = u_{i+1}$$

pour tout $i = 0, \dots, k - 2$ et que

$$u'_{k-1} = f(s, u_0, \dots, u_{k-1})$$

donc on trouve

$$\frac{d^k u_0}{dt^k} = f\left(s(t), u_0, \dots, \frac{d^{k-1} u_0}{dt^{k-1}}\right).$$

Nous savons aussi que $s'(t) = 1$ et donc $s(t) = t + c$ où $c \in \mathbb{R}$. Comme y est solution, cela impose que $t + c \in I$ pour tout $t \in J'$. D'où la fonction $x(t) = u_0(t - c)$ est solution de l'équation différentielle ordinaire initiale sur l'intervalle $J = J' + c \subset I$. \square

2.2. Existence des solutions. Maintenant que nous avons vu comment se ramener à une équation différentielle ordinaire d'ordre 1 et autonome du type $\frac{dx}{dt} = f(x)$, nous allons interpréter celle-ci géométriquement.

Definition 2.2. *Un champ de vecteurs sur un ouvert $U \subset \mathbb{R}^n$ est la donnée d'une application $f : U \rightarrow \mathbb{R}^n$.*

Pour mieux comprendre la signification géométrique, on réalise le champ de vecteurs f en liant le vecteur $f(x)$ au point x (voir figure 9).

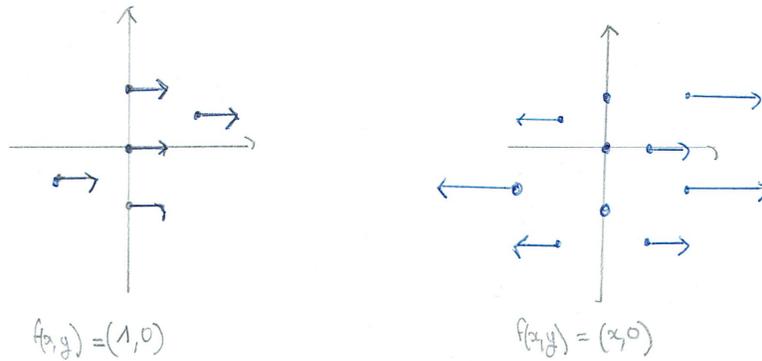


FIGURE 9. Exemples de champs de vecteurs de \mathbb{R}^2 .

Definition 2.3. *On appelle courbe intégrale du champ de vecteurs f une courbe $\gamma : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ de classe C^1 tangente en tout point à f , c'est-à-dire telle que*

$$\gamma'(t) = f(\gamma(t)).$$

Autrement dit, on interprète géométriquement les solutions de l'équation différentielle $\frac{dx}{dt} = f(x)$ comme les courbes intégrales du champ de vecteurs défini par f . Par exemple, si on considère le champ de vecteurs $f(x, y) = (-x, y)$, par tout point $(x_0, y_0) \in \mathbb{R}^2$ passe une unique courbe intégrale s'écrivant

$$\gamma_{(x_0, y_0)}(t) = (x_0 \cdot e^{-t}, y_0 \cdot e^t).$$

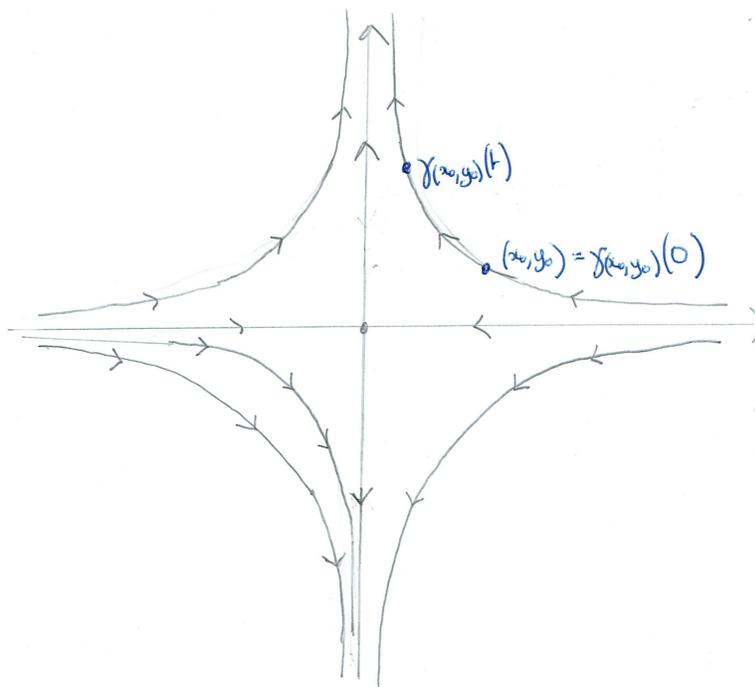
L'ensemble des courbes intégrales forme un dessin que l'on appelle *portrait de phase*, voir figure 10.

Nous allons voir maintenant que tout champ de vecteurs suffisamment régulier s'intègre toujours localement.

Théorème 2.1. (Cauchy-Lipschitz) *Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ une application de classe C^1 et $x_0 \in U$.*

a) *Alors il existe $\alpha > 0$ et $\gamma :] - \alpha, \alpha[\rightarrow U$ de classe C^2 tels que $\gamma(0) = x_0$ et $\gamma'(t) = f(\gamma(t))$ pour tout $t \in] - \alpha, \alpha[$.*

b) *De plus, si $\bar{\gamma} :] - \alpha', \alpha'[\rightarrow U$ désigne une autre courbe intégrale passant par x_0 en $t = 0$, alors γ et $\bar{\gamma}$ coïncident sur leur intervalle commun de définition.*

FIGURE 10. Portrait de phase du champ $f(x, y) = (-x, y)$.

Démonstration. Existence. Nous allons utiliser le théorème du point fixe de Picard dans la version suivante.

Théorème 2.2. (du point fixe de Picard) Soit (E, d) un espace métrique complet et $T : E \rightarrow E$ une application contractante : il existe $0 < k < 1$ tel que pour tout $x, y \in E$ on ait

$$d(T(x), T(y)) \leq k \cdot d(x, y).$$

Alors il existe un unique point $x_0 \in E$ tel que $T(x_0) = x_0$.

Tout d'abord, il nous faut remarquer que

$$\gamma'(t) = f(\gamma(t)) \text{ et } \gamma(0) = x_0 \Leftrightarrow \gamma(t) = x_0 + \int_0^t f(\gamma(s)) ds.$$

Etant donné $\alpha > 0$, nous définissons donc l'application

$$\begin{aligned} T : C^0([- \alpha, \alpha], U) &\rightarrow C^0([- \alpha, \alpha], \mathbb{R}^n) \\ \gamma &\mapsto (t \mapsto x_0 + \int_0^t f(\gamma(s)) ds) \end{aligned}$$

dont on cherche un point fixe.

Pour tout $\alpha > 0$ et $r > 0$, nous munissons l'espace $C^0([- \alpha, \alpha], \bar{B}(x_0, r))$ de la norme infini :

$$\|u\|_\infty = \sup_{t \in [- \alpha, \alpha]} \|u(t)\|.$$

Cet espace métrique est complet. Nous allons prouver qu'il existe $r > 0$ et $\alpha > 0$ tels que

- (1) $T(C^0([- \alpha, \alpha], \bar{B}(x_0, r))) \subset C^0([- \alpha, \alpha], \bar{B}(x_0, r))$,
- (2) T est contractante.

Nous pourrions ainsi conclure à l'existence d'un point fixe γ de T , c'est-à-dire une courbe intégrale de f passant par x_0 au temps $t = 0$.

Fixons tout d'abord un rayon $r > 0$ satisfaisant l'inclusion $\bar{B}(x_0, r) \subset U$. Comme $f : U \rightarrow \mathbb{R}^n$ est C^1 donc continue, nous pouvons poser

$$M := \max_{x \in \bar{B}(x_0, r)} \|f(x)\| < +\infty.$$

On calcule alors

$$\|T(\gamma(t)) - x_0\| = \left\| \int_0^t f(\gamma(s)) ds \right\| \leq M \cdot \alpha$$

en utilisant à nouveau l'inégalité de Jensen. Nous aurons donc satisfait le point (1) ci-dessus si l'on choisit $M \cdot \alpha \leq r$.

On calcule ensuite $\forall \gamma, \tilde{\gamma} \in C^0([- \alpha, \alpha], \bar{B}(x_0, r))$ que

$$\begin{aligned} \|T(\gamma) - T(\tilde{\gamma})\|_\infty &= \sup_{|t| < \alpha} \|T(\gamma)(t) - T(\tilde{\gamma})(t)\| \\ &= \sup_{|t| < \alpha} \left\| \int_0^t [f(\gamma(s)) - f(\tilde{\gamma}(s))] ds \right\| \\ &\leq \alpha \cdot \sup_{|s| < \alpha} \|f(\gamma(s)) - f(\tilde{\gamma}(s))\|. \end{aligned}$$

Comme f est de classe C^1 , sa différentielle est continue et donc

$$k := \sup_{\bar{B}(x_0, r)} \|df_x\| < \infty.$$

Par le TAF, on en déduit que f est k -lipschitzienne, et donc que

$$\|T(\gamma) - T(\tilde{\gamma})\|_\infty \leq \alpha \cdot k \cdot \|\gamma - \tilde{\gamma}\|_\infty.$$

L'application se trouve ainsi être contractante si on choisit α de sorte que $\alpha \cdot k < 1$. Ceci est compatible avec le choix précédent de α puisqu'au pire on diminue encore la valeur de α .

Les points (1) et (2) étant satisfaits pour un certain couple de valeur (r, α) , il existe par le théorème du point fixe une unique application $\gamma \in C^0([- \alpha, \alpha], \bar{B}(x_0, r))$ tel que $T(\gamma) = \gamma$. Cette courbe γ est alors une courbe intégrale de f passant par x_0 au temps $t = 0$. On constate que γ est de classe C^1 en vertu du théorème fondamental du calcul différentiel, comme

$$\gamma(t) = x_0 + \int_0^t f(\gamma(s)) ds.$$

On constate alors que γ est de classe C^2 puisque $\gamma' = f \circ \gamma$ avec f et γ de classe C^1 .

Unicité. L'unicité se montre de la manière suivante. On commence par montrer que les deux solutions coïncident sur un intervalle non vide contenant 0. On choisit $0 < \beta \leq \min\{\alpha, \alpha'\} \leq \alpha$ tel que $\bar{\gamma}(t) \in \bar{B}(x_0, r)$ pour tout $t \in J := [-\beta/2, \beta/2]$. On pose alors

$$Q := \max_{t \in J} \|\gamma(t) - \bar{\gamma}(t)\|.$$

Par continuité, ce maximum est atteint en un point $t_1 \in J$. Donc

$$\begin{aligned} Q = \|\gamma(t_1) - \bar{\gamma}(t_1)\| &= \left\| \int_0^{t_1} (\gamma'(s) - \bar{\gamma}'(s)) ds \right\| \\ &\leq \int_0^{t_1} \|f(\gamma(s)) - f(\bar{\gamma}(s))\| ds \\ &\leq \int_0^{t_1} k \|\gamma(s) - \bar{\gamma}(s)\| ds \\ &\leq \alpha \cdot k \cdot Q. \end{aligned}$$

Comme $\alpha \cdot k < 1$, on obtient $Q = 0$ et les deux solutions coïncident donc sur J .

Maintenant on note J^* le plus grand intervalle ouvert contenant 0 et tel que $\gamma = \bar{\gamma}$ sur J^* . Si J^* ne coïncide pas avec $] -\alpha, \alpha[\cap] -\alpha', \alpha'[,$ alors J^* admet un point t_2 à sa frontière contenu dans l'intervalle ouvert $] -\alpha, \alpha[\cap] -\alpha', \alpha'[/math>. Par continuité $\gamma(t_2) = \bar{\gamma}(t_2)$ et donc les courbes $s \mapsto \gamma(t_2 + s)$ et $s \mapsto \bar{\gamma}(t_2 + s)$ coïncident sur un voisinage ouvert de 0. L'intervalle J^* n'était donc pas maximal. $\square$$

En corollaire du théorème de Cauchy-Lipschitz, nous avons donc que toute équation différentielle ordinaire d'ordre $k \in \mathbb{N}^*$ admet des solutions locales, et que celles-ci sont uniques si on prescrit les valeurs en un point des dérivées de x jusqu'à l'ordre $k - 1$.

Du point de vue géométrique, nous avons donc que pour tout champ de vecteurs de classe C^1 , par tout point passe une courbe intégrale, et que deux courbes intégrales ne peuvent jamais se croiser (sauf à être confondues). De même, une courbe intégrale ne peut jamais s'autointersecter, sauf si elle est périodique.

2.3. Comportement des solutions. Maintenant que nous savons que toute équation différentielle ordinaire d'ordre 1 et autonome admet des solutions locales, ou de manière équivalente tout champ de vecteurs admet des courbes intégrales au voisinage de chaque point, nous pouvons définir la notion de solution maximale.

Definition 2.4. Soit $f : U \rightarrow \mathbb{R}^n$ un champ de vecteurs de classe C^1 .

Pour chaque point $x_0 \in U$, il existe un unique intervalle $J(x_0) \subset \mathbb{R}$ contenant 0 en son intérieur et une unique courbe $\gamma_{x_0} : J(x_0) \rightarrow U$ intégrale de f et telle que $\gamma_{x_0}(0) = x_0$ ayant la propriété de maximalité suivante : toute autre courbe intégrale $\bar{\gamma} : \bar{J} \rightarrow U$ passant par x_0 en $t = 0 \in \bar{J}$ vérifie l'inclusion $\bar{J} \subset J(x_0)$. Bien évidemment nous avons alors $\gamma_{x_0}(t) = \bar{\gamma}(t)$ pour tout $t \in \bar{J}$.

La courbe $\gamma_{x_0} : J(x_0) \rightarrow U$ est alors appelée solution (ou courbe intégrale) maximale de l'équation $\frac{dx}{dt} = f(x)$ passant par x_0 au temps $t = 0$.

Remarquer que $J(x_0)$ est un intervalle nécessairement ouvert que nous noterons $]a, b[$: sinon, on pourrait prolonger la solution en son extrémité en utilisant le théorème de Cauchy-Lipschitz. Mais il est possible que $a = -\infty$ ou $b = \infty$. Lorsque a ou b est fini, la question se pose de savoir ce qu'il se passe à la limite de l'intervalle maximal de définition.

Théorème 2.3. (des bouts) Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ un champ de vecteurs de classe C^1 , x_0 un point de U et

$$\gamma_{x_0} :]a, b[\rightarrow U$$

la courbe intégrale maximale issue de x_0 (c'est-à-dire vérifiant $\gamma_{x_0}(0) = x_0$).

Si $b < \infty$, alors pour tout compact $K \subset U$ il existe $\eta \in]a, b[$ tel que

$$\gamma_{x_0}(t) \notin K$$

pour tout $\eta < t < b$.

Respectivement, quitte à remplacer le champ f par $-f$, on voit que si $a > -\infty$, alors pour tout compact $K \subset U$ il existe $\eta \in]a, b[$ tel que $\forall a < t < \eta$ on ait $\gamma_{x_0}(t) \notin K$.

Ceci implique que lorsque l'intervalle maximal de définition d'une courbe intégrale est fini en une extrémité, en celle-ci la courbe tend soit vers le bord de U , soit à l'infini (voir figure 11).

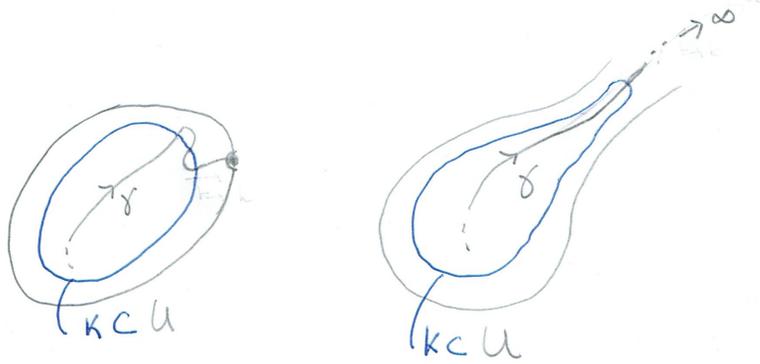


FIGURE 11. $b < \infty$ implique fuite au bord ou à l'infini.

Démonstration. Pour plus de commodité, notons par γ la courbe intégrale γ_{x_0} . Supposons le contraire : il existe un compact K et une suite (t_n) de points dans $]a, b[$ telle que

$$\lim_{n \rightarrow \infty} t_n = b$$

et $\gamma(t_n) \in K$ pour tout entier n . L'ensemble K étant compact, nous pouvons (quitte à extraire une sous-suite) également supposer qu'il existe $z \in K$ tel que

$$\lim_{n \rightarrow \infty} \gamma(t_n) = z.$$

Maintenant fixons $\alpha > 0$ et $r > 0$ tels que d'une part $B(z, 2r) \subset U$, et d'autre part

$$\alpha \cdot \sup_{\bar{B}(z,r)} \|f(x)\| \leq \frac{r}{2}$$

et

$$\alpha \cdot \sup_{\bar{B}(z,r)} \|df_x\| < \frac{1}{2}.$$

Parce que f est de classe C^1 , les applications

$$x_0 \in B(z, r) \mapsto \sup_{\bar{B}(x_0,r)} \|f(x)\|$$

et

$$x_0 \in B(z, r) \mapsto \sup_{\bar{B}(x_0,r)} \|df_x\|$$

sont continues. On en déduit qu'il existe $\varepsilon > 0$ tel que

$$\alpha \cdot \sup_{\bar{B}(x_0,r)} \|f(x)\| \leq r$$

et

$$\alpha \cdot \sup_{\bar{B}(x_0,r)} \|df_x\| < 1$$

pour tout $x_0 \in B(z, \varepsilon)$. Ainsi, par l'argumentation du théorème de Cauchy-Lipschitz, les domaines de définition $J(x_0)$ des courbes intégrales maximales issues de points x_0 appartenant à $B(z, \varepsilon)$ vérifient

$$] - \alpha, \alpha[\subset J(x_0).$$

On choisit alors $t_n \in]b - \alpha/2, b[$ tel que $\gamma(t_n) \in B(z, \varepsilon)$. La courbe intégrale issue de $\gamma(t_n)$ est définie sur l'intervalle $] - \alpha, \alpha[$ et permet donc de prolonger γ au-delà de b . Ceci contredit la maximalité de γ . \square

Nous considérons maintenant les courbes intégrales dans leur ensemble. Comme nous venons de le voir dans la démonstration du théorème des bouts, deux courbes intégrales démarrant en des points suffisamment proches seront définies sur un intervalle fermé commun. Nous pouvons quantifier à quel point elles resteront proches l'une de l'autre de la manière suivante.

Théorème 2.4. Soit f un champ de vecteurs de classe C^1 sur un ouvert U et $x_0 \in U$. Supposons que $[0, c] \subset J(x_0)$.

Alors il existe un voisinage U' de x_0 dans U et une constante $k \geq 0$ tel que, pour tout $x \in U'$, la courbe intégrale γ_x issue de x soit elle-aussi définie sur $[0, c]$ et vérifie l'estimée suivante :

$$\|\gamma_{x_0}(t) - \gamma_x(t)\| \leq \|x_0 - x\| \exp(k \cdot t)$$

pour tout $t \in [0, c]$.

Nous pouvons interpréter ce résultat de la manière suivante. D'une part, la courbe intégrale passant par x dépend de manière continue du point x (voir figure 12). Plus précisément,

$$\lim_{x \rightarrow x_0} \gamma_x(t) = \gamma_{x_0}(t)$$

pour tout $t \in [0, c]$. D'autre part, deux courbes démarrant relativement proche ne peuvent dévier l'une de l'autre qu'au plus d'une manière exponentielle en fonction du temps.

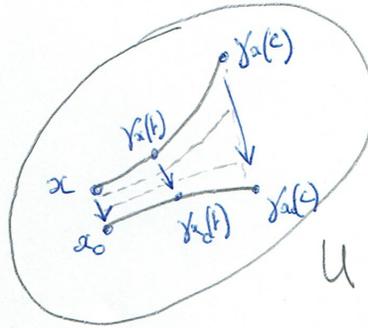


FIGURE 12. Convergence simple des courbes intégrales et divergence au plus exponentielle.

Démonstration. Par compacité de $[0, c]$ il existe $\varepsilon > 0$ tel que $x \in U$ dès que $\|x - \gamma_{x_0}(t)\| \leq \varepsilon$ pour un $t \in [0, c]$. On pose alors $K = \{x \in \mathbb{R}^n \mid d(x, \gamma_{x_0}([0, c])) \leq \varepsilon\} \subset U$. L'ensemble K est compact et donc l'application f est k -lipschitz sur K (en posant $k = \max_{x \in K} \|df_x\|$ et en appliquant le TAF).

Soit $\delta > 0$ suffisamment petit pour que $\delta \exp(k \cdot c) \leq \frac{\varepsilon}{2}$. En particulier, $\delta \leq \frac{\varepsilon}{2}$. Posons

$$U' = \{x \in \mathbb{R}^n \mid \|x - x_0\| < \delta\} \subset \overset{\circ}{K} \subset U.$$

Nous allons montrer que pour tout $x \in U'$ la courbe intégrale γ_x est définie sur $[0, c]$.

Soit $\tilde{\gamma}_x$ la courbe intégrale du champ de vecteurs $f|_K$ définie sur un intervalle maximal de la forme $]a, b[$.

Nous allons montrer que $b > c$.

Remarquons que $\tilde{\gamma}_x :]a, b[\rightarrow \overset{\circ}{K}$ est aussi une courbe intégrale de f issue de x . De la maximalité de γ_x , on déduit que $\gamma_x(t) = \tilde{\gamma}_x(t)$ pour tout $t \in]a, b[$. Posons pour $t \in [0, b[$

$$v(t) = \|\gamma_{x_0}(t) - \gamma_x(t)\|.$$

Alors

$$\begin{aligned} v(t) = \|\gamma_{x_0}(t) - \gamma_x(t)\| &= \|\gamma_{x_0}(0) - \gamma_x(0) + \int_0^t (f(\gamma_{x_0}(s)) - f(\gamma_x(s)))ds\| \\ &\leq v(0) + \int_0^t k \cdot v(s)ds \end{aligned}$$

puisque γ_{x_0} et γ_x ne sortent pas de $\overset{\circ}{K}$. Nous invoquons alors le lemme de Gronwall.

Lemme 2.1. (de Gronwall) Soit $u : [0, a] \rightarrow \mathbb{R}_+$ une application continue. Supposons qu'il existe $C > 0$ et $k \geq 0$ tel que

$$u(t) \leq C + \int_0^t k \cdot u(s)ds$$

pour tout $t \in [0, a]$. Alors

$$u(t) \leq C \exp(k \cdot t)$$

pour tout $t \in [0, a]$.

Démonstration. Posons $U(t) = C + \int_0^t k \cdot u(s)ds > 0$. Alors $u(t) \leq U(t)$ et

$$U'(t) = ku(t),$$

donc

$$\frac{U'(t)}{U(t)} = \frac{ku(t)}{U(t)} \leq k.$$

On en déduit que

$$\log U(t) \leq \log U(0) + kt$$

et comme $U(0) = C$ en passant à l'exponentielle il vient

$$U(t) \leq C \exp(k \cdot t)$$

d'où le résultat comme $u(t) \leq U(t)$. □

En appliquant le lemme de Gronwall à $u = v$ avec $C = v(0) = \|x_0 - x\|$, il vient pour tout $t \in [0, b[$

$$\|\gamma_{x_0}(t) - \gamma_x(t)\| \leq \|x_0 - x\| \exp(k \cdot t) \leq \delta \exp(k \cdot b)$$

Si $b \leq c$, cela impliquerait alors que

$$\|\gamma_{x_0}(t) - \gamma_x(t)\| \leq \delta \exp(k \cdot c) \leq \frac{\varepsilon}{2},$$

et donc en particulier $\gamma_x(t)$ ne quitte pas le compact $\{x \in \mathbb{R}^n \mid d(x, \gamma_{x_0}([0, c])) \leq \frac{\varepsilon}{2}\} \subset \overset{\circ}{K}$ pour tout $t \in [0, b[$. Ceci contredirait la maximalité en vertu du théorème des bouts.

Donc pour tout $x \in U'$ la courbe intégrale γ_x est définie sur $[0, c]$. Nous avons par ailleurs prouvé l'estimée suivante :

$$\|\gamma_{x_0}(t) - \gamma_x(t)\| \leq \|x_0 - x\| \exp(k \cdot t)$$

pour tout $t \in [0, c]$, ce qui conclut la preuve. □

Remarquons que, quitte à remplacer f par $-f$ et intervertir passé et futur des courbes, nous pouvons montrer que si un intervalle fermé $[c, d]$ est contenu dans $J(x_0)$, alors il existe un ouvert U' autour de x_0 tel que $[c, d] \subset J(x)$ pour tout $x \in U'$ et pour lequel pour tout $t \in [c, d]$

$$\|\gamma_{x_0}(t) - \gamma_x(t)\| \leq \|x_0 - x\| \exp(k \cdot t)$$

où $k \geq 0$ est une constante.

2.4. Flot d'un champ de vecteurs.

Definition 2.5. Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ un champ de vecteurs de classe C^1 . On définit le flot du champ de vecteur comme l'application

$$\begin{aligned} \varphi : \Omega = \{(t, x) \in \mathbb{R} \times U \mid t \in J(x)\} &\rightarrow U \\ (t, x) &\mapsto \gamma_x(t) \end{aligned}$$

où γ_x désigne la courbe intégrale maximale de f telle que $\gamma_x(0) = x$. Nous noterons

$$\varphi(t, x) = \varphi_t(x).$$

On peut écrire

$$\Omega = \bigcup_{x \in U} J(x) \times \{x\}.$$

Ainsi $\{0\} \times U \subset \Omega$ mais en général il n'existe pas d'intervalle ouvert J autour de 0 tel que $J \times U \subset \Omega$. Néanmoins, nous pouvons démontrer le résultat suivant.

Proposition 2.2. Ω est un ouvert de $\mathbb{R} \times U$, et l'application $\varphi : \Omega \rightarrow U$ est continue.

Démonstration. Montrons tout d'abord que Ω est ouvert. Fixons $(t_0, x_0) \in \Omega$ et choisissons $\alpha > 0$ tel que $[t_0 - \alpha, t_0 + \alpha] \subset J(x_0)$. On peut alors trouver un voisinage ouvert U' de x_0 tel que pour tout $x \in U'$ on ait $[t_0 - \alpha, t_0 + \alpha] \subset J(x)$ ainsi que la majoration de type exponentielle du théorème précédent. On a bien trouvé un voisinage ouvert

$$]t_0 - \alpha, t_0 + \alpha[\times U'$$

de (t_0, x_0) qui soit inclus dans Ω , et donc Ω est ouvert.

Maintenant, si $(t, x) \in]t_0 - \alpha, t_0 + \alpha[\times U'$, nous avons

$$\begin{aligned} \|\varphi_{t_0}(x_0) - \varphi_t(x)\| &= \|\gamma_{x_0}(t_0) - \gamma_x(t)\| \\ &\leq \|\gamma_{x_0}(t_0) - \gamma_{x_0}(t)\| + \|\gamma_{x_0}(t) - \gamma_x(t)\| \\ &\leq \|\gamma_{x_0}(t_0) - \gamma_{x_0}(t)\| + \|x_0 - x\| \exp(k \cdot t) \\ &\leq \|\gamma_{x_0}(t_0) - \gamma_{x_0}(t)\| + \|x_0 - x\| \exp(k \cdot (t_0 + \varepsilon)). \end{aligned}$$

Ainsi pour tout $\varepsilon > 0$, la continuité de $t \mapsto \gamma_{x_0}(t)$ implique l'existence de $\eta > 0$ tel que

$$\|\gamma_{x_0}(t_0) - \gamma_{x_0}(t)\| < \varepsilon$$

dès que $|t - t_0| < \eta$. On en déduit que pour $(t, x) \in]t_0 - \eta, t_0 + \eta[\times B\left(x_0, \frac{\varepsilon}{\exp(k \cdot (t_0 + \varepsilon))}\right)$

$$\|\varphi_{t_0}(x_0) - \varphi_t(x)\| \leq 2\varepsilon$$

montrant ainsi la continuité de φ . □

Nous observons ensuite que le flot vérifie la formule remarquable suivante.

Proposition 2.3. (Formule du flot)

Si $t_1 \in J(x)$ et $t_2 \in J(\varphi_{t_1}(x))$, alors $t_1 + t_2 \in J(x)$ et

$$\varphi_{t_1+t_2}(x) = \varphi_{t_2}(\varphi_{t_1}(x)).$$

Démonstration. Considérons la courbe $t \mapsto \varphi_{t_1+t}(x)$. C'est une courbe intégrale démarrant en $\varphi_{t_1}(x)$ et elle est maximale définie sur l'intervalle $J(x) - t_1$ (exercice). On en déduit que $J(\varphi_{t_1}(x)) = J(x) - t_1$ et que $\varphi_t(\varphi_{t_1}(x)) = \varphi_{t_1+t}(x)$ par unicité des courbes intégrales démarrant en un point. Ainsi $t_2 \in J(\varphi_{t_1}(x))$ ssi $t_1 + t_2 \in J(x)$ et on a alors la formule annoncée. □

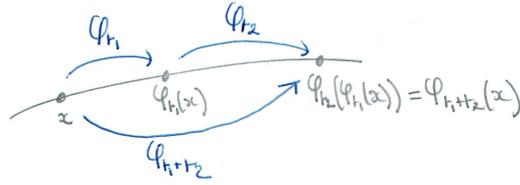


FIGURE 13. Formule du flot.

Observons en particulier que si $t \in J(x)$, alors $-t \in J(\varphi_t(x))$ et $\varphi_{-t}(\varphi_t(x)) = x$. Dans le cas où $\Omega = \mathbb{R} \times U$ (autrement dit, $J(x) = \mathbb{R}$ pour tout $x \in U$, on dit que le champ de vecteurs est complet), on dispose alors d'un morphisme de groupes

$$\begin{aligned} \varphi : \mathbb{R} &\rightarrow \text{Homéo}(U) \\ t &\mapsto \varphi_t \end{aligned}$$

de \mathbb{R} dans l'ensemble des homéomorphismes de U . Ceci justifie la notation $\varphi(t, \cdot) = \varphi_t$ comme la famille $\{\varphi_t\}_{t \in \mathbb{R}}$ est une famille à un paramètre d'homéomorphismes. En fait, on dispose même d'une famille à un paramètre de difféomorphismes de U en vertu du résultat suivant.

Théorème 2.5. *Si f est un champ de vecteurs de classe C^1 défini sur un ouvert U , alors le flot $\varphi : \Omega \rightarrow U$ est de classe C^1 .*

Autrement dit, lorsque le champ de vecteurs est complet (cad $\Omega = \mathbb{R} \times U$), chaque homéomorphisme φ_t est de classe C^1 et son inverse φ_{-t} aussi. On obtient ainsi une famille à un paramètre de difféomorphismes de U : la courbe de classe C^1 suivante

$$\begin{aligned} \varphi : \mathbb{R} &\rightarrow \text{Difféo}(U) \\ t &\mapsto \varphi_t \end{aligned}$$

est un morphisme de groupes de \mathbb{R} dans l'espace des difféomorphismes de U (en particulier $\varphi(0) = id_U$).

Démonstration. Devoir maison! □

2.5. Orbites. On se fixe un champ de vecteurs $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ de classe C^1 . Nous allons classer l'ensemble des courbes intégrales en fonction de leur image.

Definition 2.6. *L'orbite d'un point x de U est le sous-ensemble*

$$\mathcal{O}_x := \{\varphi_t(x) \mid t \in J(x)\} \subset U.$$

Cette définition permet de décomposer géométriquement le domaine U selon la dynamique de f .

Proposition 2.4. *Les orbites de f forment une partition de U , appelée portrait de phase de f . De manière équivalente, la relation définie par*

$$y \sim x \Leftrightarrow y \in \mathcal{O}_x$$

est une relation d'équivalence. Ainsi

$$U = \coprod_{[x] \in U/\sim} \mathcal{O}_x.$$

Démonstration. Il faut vérifier que la relation \sim est réflexive, symétrique et transitivité. Tout d'abord $x \in \mathcal{O}_x$ et donc $x \sim x$, doù la réflexivité. Ensuite,

$$y \sim x \Leftrightarrow y \in \mathcal{O}_x \Leftrightarrow \exists t \in J(x) \mid \varphi_t(x) = y \Leftrightarrow \exists -t \in J(y) \mid x = \varphi_{-t}(y) \Leftrightarrow x \in \mathcal{O}_y \Leftrightarrow x \sim y$$

d'où la symétrie. Enfin, si $y \in \mathcal{O}_x$ et $z \in \mathcal{O}_y$, alors il existe $t_1 \in J(x)$ tel que $\varphi_{t_1}(x) = y$ et $t_2 \in J(y) = J(\varphi_{t_1}(x))$ tel que $z = \varphi_{t_2}(y)$, et donc par la formule du flot $t_1 + t_2 \in J(x)$ et

$$\varphi_{t_1+t_2}(x) = \varphi_{t_2}(\varphi_{t_1}(x)) = z$$

cad $z \in \mathcal{O}_x$, d'où la transitivité. \square

Nous nous intéressons à un type particulier d'orbite, qui jouera un rôle central par la suite.

Definition 2.7. Une orbite est dite périodique s'il existe $T > 0$ dans $J(x)$ tel que

$$\varphi_T(x) = x.$$

On observe tout d'abord que si l'orbite de x est périodique, alors $J(x) = \mathbb{R}$. On définit alors l'ensemble des périodes comme

$$\mathcal{P} = \{T \in \mathbb{R} \mid \varphi_T(x) = x\} \subset \mathbb{R}$$

La formule du flot assure que \mathcal{P} est un sous-groupe additif de \mathbb{R} . Le nombre

$$T_0 = \inf\{T > 0 \mid \varphi_T(x) = x\}.$$

est fini par hypothèse de périodicité, et nous avons par continuité du flot que

$$\varphi_{T_0+t}(x) = \varphi_t(x)$$

pour tout $t \in \mathbb{R}$. Ainsi $T_0 \in \mathcal{P}$ et est appelé *période minimale* de \mathcal{O}_x .

Proposition 2.5. Si $T_0 > 0$, alors

$$\mathcal{P} = \mathbb{Z} \cdot T_0.$$

Démonstration. Soit $T \in \mathcal{P} \setminus \{0\}$ et posons $T' = T - [T/T_0] \cdot T_0$, la notation $[\cdot]$ désignant la partie entière. Si $T' \neq 0$ alors $0 < T' < T_0$ et pourtant $T' \in \mathcal{P}$: contradiction. Ainsi il existe un unique entier $k \in \mathbb{N}^*$ pour lequel $T = k \cdot T_0$ et on a $\mathcal{P} \subset \mathbb{Z} \cdot T_0$. L'inclusion réciproque est évidente. \square

Proposition 2.6. Si $T_0 = 0$, alors $\mathcal{P} = \mathbb{R}$ et $f(x) = 0$. On dit alors que x est un point fixe du flot.

Démonstration. Observons tout d'abord que \mathcal{P} est fermé puisque l'équation $\varphi_T(x) = \varphi(T, x) = x$ est continue en T . Si $T_0 = 0$, on montre que \mathcal{P} est dense dans \mathbb{R} . En effet, pour tout $a \in \mathbb{R}$ et pour tout $k \in \mathbb{N}^*$, il existe $T \in]0, 1/k[\cap \mathcal{P}$ et alors $[a/T] \cdot T$ est une période à distance $< 1/k$ de a , d'où la densité. Ainsi \mathcal{P} est fermé et dense dans \mathbb{R} , et donc $\mathcal{P} = \overline{\mathcal{P}} = \mathbb{R}$.

Maintenant, on sait que $\gamma_x(t) = \varphi_t(x) = x$ pour tout $t \in \mathbb{R}$. On en déduit que

$$f(x) = f(\gamma_x(0)) = (\gamma_x)'(0) = 0.$$

\square

Réciproquement, si $f(x) = 0$, alors la courbe $\gamma : t \in \mathbb{R} \mapsto x$ est une courbe intégrale de f issue de x et donc x est un point fixe du flot. Autrement dit,

Proposition 2.7. x est un point fixe du flot si et seulement si $f(x) = 0$.

2.6. Champs de vecteurs linéaires et portrait de phase en dimension 2. Soit $A \in \mathcal{L}(\mathbb{R}^n)$ un endomorphisme linéaire de \mathbb{R}^n , que l'on identifiera à sa matrice écrite dans la base canonique. A cette application linéaire est naturellement associé le champ de vecteurs f défini par $f(x) = A \cdot x$. Ce type de champ de vecteurs est dit linéaire. On considère l'équation différentielle associée

$$\frac{dx}{dt} = Ax$$

qui est alors dite linéaire à coefficients constants.

Proposition 2.8. *Pour tout $x_0 \in \mathbb{R}^n$, la solution maximale de condition initiale x_0 est définie sur \mathbb{R} et donnée par la formule*

$$x(t) = e^{tA}x_0.$$

Nous rappelons ici quelques propriétés de l'application exponentielle. Tout d'abord, elle est définie de la manière suivante : la série de terme générale

$$\sum_{k=0}^{\infty} \frac{A^k}{k!}$$

converge normalement sur les boules fermées $\overline{B}(0, R) = \{A \in M_n(\mathbb{R}) \mid \|A\| \leq R\}$ (ici, nous rappelons que $\|A\| = \max_{\|x\|=1} \|Ax\|$). En effet, pour tout $A \in \overline{B}(0, R)$ on a pour tout $k \in \mathbb{N}$

$$\left\| \frac{A^k}{k!} \right\| \leq \frac{\|A\|^k}{k!} \leq \sum_{k=0}^{\infty} \frac{R^k}{k!}$$

qui est le terme d'une série convergente. On voit facilement que $e^0 = Id$ et on calcule

$$e^A \cdot e^{-A} = \sum_{k=0}^{\infty} \frac{A^k}{k!} \cdot \sum_{k=0}^{\infty} \frac{(-A)^k}{k!} = \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{A^k}{k!} \cdot \frac{(-A)^{n-k}}{(n-k)!} = \sum_{n=0}^{\infty} \frac{A^n}{n!} (1-1)^n = Id.$$

On peut ainsi définir une application

$$\begin{aligned} \exp : M_n(\mathbb{R}) &\rightarrow GL_n(\mathbb{R}) \\ A &\mapsto e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!} \end{aligned}$$

appelée exponentielle de matrice. On calcule facilement

$$\frac{d}{dt} e^{tA} = \lim_{h \rightarrow 0} \frac{e^{(t+h)A} - e^{tA}}{h} = \sum_{k=1}^{\infty} \frac{A^k}{k!} \lim_{h \rightarrow 0} \frac{(t+h)^k - t^k}{h} = \sum_{k=1}^{\infty} \frac{t^{k-1} \cdot A^k}{(k-1)!} = A \cdot e^{tA}$$

et on obtient ainsi que les solutions de l'équation différentielle linéaire $\frac{dx}{dt} = Ax$ sont de la forme

$$x(t) = e^{tA}x_0.$$

En pratique, cette formule est utile lorsque l'on sait calculer e^A explicitement comme dans le cas suivant, ou alors dans certains cas de la dimension 2.

Cas diagonalisable. Supposons que A soit diagonalisable. Il existe alors une matrice $P \in GL_n(\mathbb{R})$ et n valeurs propres réelles $\lambda_1, \dots, \lambda_n$ (non nécessairement distinctes) telles que

$$A = P \cdot \text{Diag}(\lambda_1, \dots, \lambda_n) \cdot P^{-1}.$$

On observe alors que

$$e^{tA} = P \cdot \text{Diag}(e^{\lambda_1 \cdot t}, \dots, e^{\lambda_n \cdot t}) \cdot P^{-1}.$$

Ainsi, nous avons montré que :

Proposition 2.9. Soit A une matrice diagonalisable de valeurs propres $\lambda_1, \dots, \lambda_n$. Alors toute solution de l'équation différentielle

$$\frac{dx}{dt} = Ax$$

avec pour condition initiale $x(0) = x_0$ s'écrit sous la forme

$$x(t) = B \cdot \begin{pmatrix} e^{\lambda_1 t} \\ \dots \\ e^{\lambda_n t} \end{pmatrix}$$

pour une matrice $B \in M_n(\mathbb{R})$ satisfaisant $B \cdot \begin{pmatrix} 1 \\ \dots \\ 1 \end{pmatrix} = x_0$.

Remarquer que B est uniquement définie si les valeurs propres de A sont deux à deux distinctes.

Cas de la dimension 2. Nous allons maintenant déterminer l'allure des solutions dans le cas de la dimension 2 en traçant les différents types de portraits de phase des champs de vecteurs linéaires qui peuvent apparaître.

Etant donnée une matrice $A \in M_2(\mathbb{R})$, on rappelle que le polynôme caractéristique s'écrit

$$P_A(t) = \det(tI_2 - A) = t^2 - \text{tr}(A)t + \det(A).$$

On trace donc la parabole donnant le lieu d'annulation du discriminant $\Delta(A) = (\text{tr}(A))^2 - 4 \det(A)$ dans un repère $(\text{tr}(A), \det(A))$.

- (1) Considérons le cas du point N_1 . La matrice A admet deux valeurs propres distinctes $\lambda_1 > 0$ et $\lambda_2 > 0$ comme $\Delta(A) > 0$, $\det(A) > 0$ et $\text{tr}(A) > 0$. On se ramène donc à un changement de base près à étudier le système

$$\begin{cases} x' = \lambda_1 x \\ y' = \lambda_2 y \end{cases}$$

dans les coordonnées d'une base formée des deux vecteurs propres de A . On a donc pour solutions dans ce nouveau système de coordonnées

$$\gamma(t) = (x_0 e^{\lambda_1 t}, y_0 e^{\lambda_2 t}).$$

Ce cas-là s'appelle noeud répulsif.

- (2) Dans le cas du point N_2 , on aura pour écriture des solutions dans la base formée des vecteurs propres

$$\gamma(t) = (x_0 e^{\lambda_1 t}, y_0 e^{\lambda_2 t}).$$

avec cette fois λ_1 et λ_2 deux valeurs propres distinctes strictement négatives. On parle alors de noeud attractif.

- (3) Considérons le cas du point S . La matrice A admet deux valeurs propres distinctes λ_1 et λ_2 comme $\Delta(A) > 0$, mais de signe opposé puisqu'alors $\det(A) < 0$. On a donc pour solutions dans ce nouveau système de coordonnées

$$\gamma(t) = (x_0 e^{\lambda_1 t}, y_0 e^{\lambda_2 t})$$

avec $\lambda_1 < 0 < \lambda_2$.

On parle de point selle. Un exemple est donné par le champ linéaire $f(x, y) = (-x, y)$ déjà étudié (voir figure 10) et associé à la matrice

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

- (4) Considérons le cas du point NS . La matrice A admet deux valeurs propres distinctes dont l'une est nulle et l'autre est notée λ . Le signe de λ coïncide avec le signe de la trace de A . On a donc pour solutions dans ce nouveau système de coordonnées

$$\gamma(t) = (x_0, y_0 e^{\lambda t}).$$

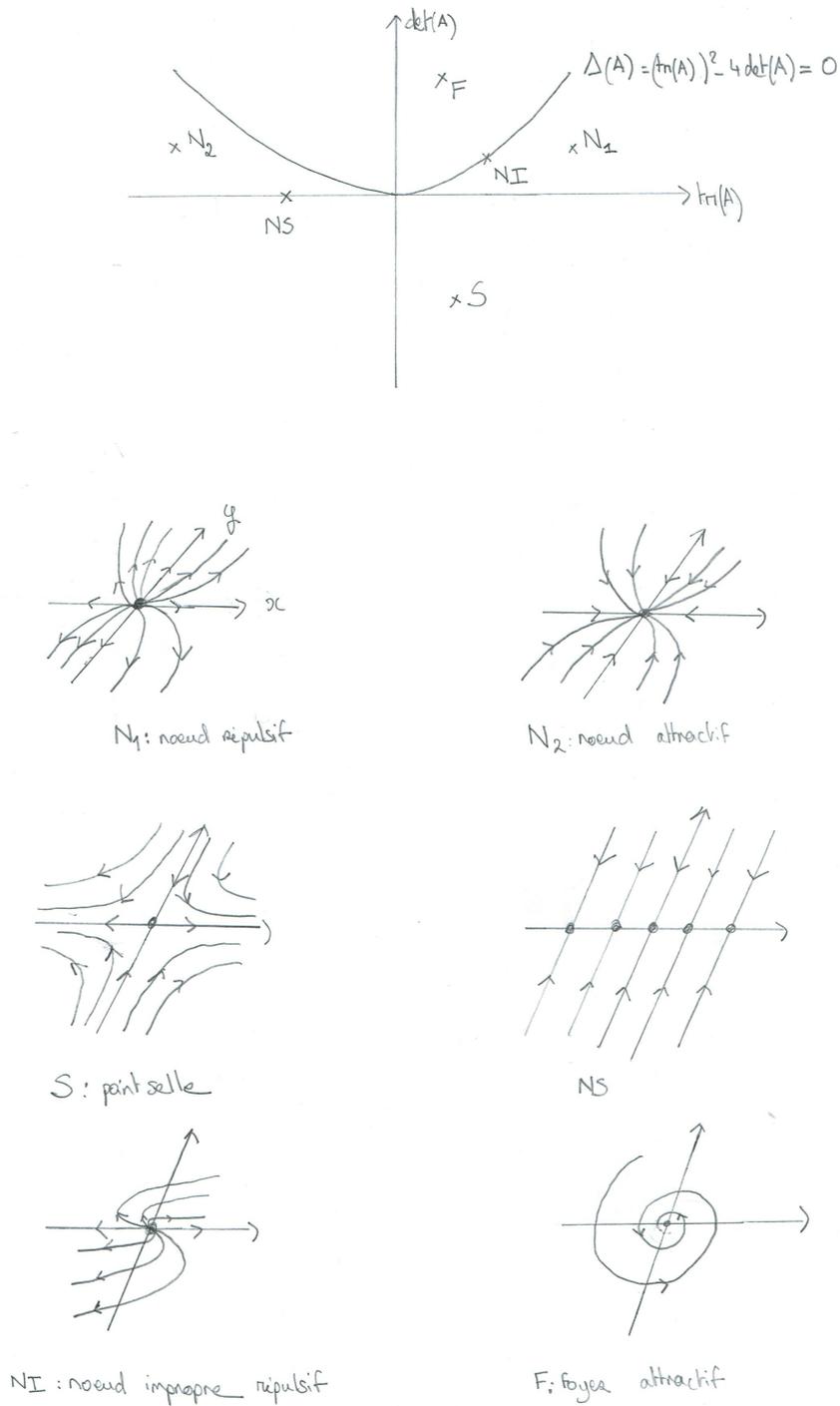


FIGURE 14. Portraits de phase d'un champ linéaire en dimension 2.

- (5) Considérons le cas du point NI . La matrice A admet une seule valeur propre λ , dont le signe concorde avec celui de la trace.
- (a) Si la matrice est néanmoins diagonalisable, alors $A = \lambda I_2$ et on retombe dans la situation du (1) ou (2).

- (b) Si la matrice n'est pas diagonalisable, alors elle admet néanmoins une réduction dite de Jordan, et donc il existe $P \in GL_2(\mathbb{R})$ telle que

$$A = P \cdot \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \cdot P^{-1}.$$

On en déduit par récurrence immédiate que

$$(tA)^n = P \cdot \begin{pmatrix} (\lambda \cdot t)^n & nt^n \lambda^{n-1} \\ 0 & (\lambda \cdot t)^n \end{pmatrix} \cdot P^{-1}.$$

et donc que

$$e^{tA} = P \cdot \begin{pmatrix} e^{\lambda \cdot t} & te^{\lambda \cdot t} \\ 0 & e^{\lambda \cdot t} \end{pmatrix} \cdot P^{-1}.$$

Dans la base de la réduction de Jordan, la solution s'écrit donc

$$\gamma(t) = (x_0 e^{\lambda \cdot t} + y_0 t e^{\lambda \cdot t}, y_0 e^{\lambda \cdot t}).$$

On parle de noeud impropre (répulsif si $\lambda > 0$ et attractif si $\lambda < 0$).

- (6) Dans le cas du point F , A admet deux valeurs propres complexes conjuguées que l'on notera $a \pm ib$. Alors il existe $P \in GL_2(\mathbb{R})$ telle que

$$A = P \cdot \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \cdot P^{-1}.$$

Si deux matrices S et T commutent, on a $e^{S+T} = e^S \cdot e^T$ et donc, les matrices $\begin{pmatrix} ta & 0 \\ 0 & ta \end{pmatrix}$ et $\begin{pmatrix} 0 & -bt \\ bt & 0 \end{pmatrix}$ commutant, on calcule

$$e^{tA} = P \cdot e^{ta} e^{\begin{pmatrix} 0 & -bt \\ bt & 0 \end{pmatrix}} \cdot P^{-1}$$

On voit alors facilement que

$$e^{\begin{pmatrix} 0 & -bt \\ bt & 0 \end{pmatrix}} = \begin{pmatrix} \cos(bt) & -\sin(bt) \\ \sin(bt) & \cos(bt) \end{pmatrix}$$

et donc les solutions s'écrivent sous la forme

$$\gamma(t) = e^{at} (x_0 \cos(bt) - y_0 \sin(bt), y_0 \cos(bt) + x_0 \sin(bt)).$$

On parle alors de foyer attractif si $a < 0$, foyer répulsif si $a > 0$ et de centre si $a = 0$.

3. DU CALCUL DIFFÉRENTIEL AUX SOUS-VARIÉTÉS

3.1. Préambule sur la différentiabilité d'ordre supérieure. Nous commençons par l'observation suivante.

Proposition 3.1. *Soit r un entier avec $2 \leq r \leq \infty$. Si $f : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^n$ est de classe C^r et f est un C^1 -difféomorphisme de U sur V , alors f est un C^r -difféomorphisme.*

Démonstration. Il faut montrer que f^{-1} est de classe C^r . Supposons que f^{-1} est de classe C^k avec $1 \leq k \leq r - 1$. Pour tout $y \in V$

$$df_y^{-1} = (df_{f^{-1}(y)})^{-1}$$

et donc l'application df_y^{-1} est obtenue comme la composition des trois applications suivantes

$$\begin{aligned} y \in V &\mapsto f^{-1}(y) \in U \\ x \in U &\mapsto df_x \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \\ u \in \mathcal{G}l(\mathbb{R}^n) &\mapsto u^{-1} \in \mathcal{G}l(\mathbb{R}^n) \end{aligned}$$

qui sont respectivement de classe C^k , C^{r-1} et C^∞ . Donc df^{-1} est de classe C^k et donc f^{-1} est de classe C^{k+1} . \square

Ainsi, le théorème d'inversion locale, et ses corollaires—à savoir les théorèmes de redressement des submersions et des immersions ainsi que le théorème des fonctions implicites—sont valables en remplaçant l'hypothèse f de classe C^1 par f de classe C^r pour r un entier allant de 1 à $+\infty$ et les difféomorphismes apparaissant dans la conclusion sont alors des C^r -difféomorphismes.

Dans tout ce qui suit sur les sous-variétés, nous supposons donc les applications toujours de classe C^∞ . En particulier, par difféomorphisme nous entendons des C^∞ -difféomorphismes à savoir des applications de classe C^∞ inversible dont l'application réciproque est également de classe C^∞ . Mais, d'après la proposition précédente, la théorie se développe de manière similaire dans la classe C^r avec un entier r quelconque ≥ 1 .

3.2. Définition, caractérisations et exemples.

Définition 3.1. *Soit $1 \leq m \leq n$ deux entiers. On dit qu'un sous-ensemble M de \mathbb{R}^n est une sous-variété de dimension m si pour chaque point $x \in M$ il existe un voisinage ouvert U de x dans \mathbb{R}^n et un difféomorphisme $h : U \rightarrow h(U)$ tel que*

$$h(U \cap M) = h(U) \cap (\mathbb{R}^m \times \{0\}).$$

h s'appelle un difféomorphisme linéarisant pour M au voisinage de x .

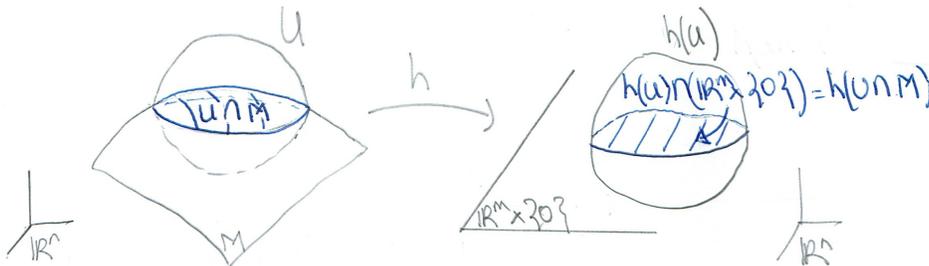


FIGURE 15. Difféomorphisme linéarisant d'une sous-variété M .

Autrement dit, une sous-variété est un sous-ensemble de \mathbb{R}^n qui est modelé localement à difféomorphisme près sur le sous-espace \mathbb{R}^m . En particulier, observer que les images de sous-variétés par des difféomorphismes

sont encore des sous-variétés. Parfois, il est utile d'utiliser la codimension qui est définie comme

$$\text{codim}(M) = n - \dim(M) = n - m.$$

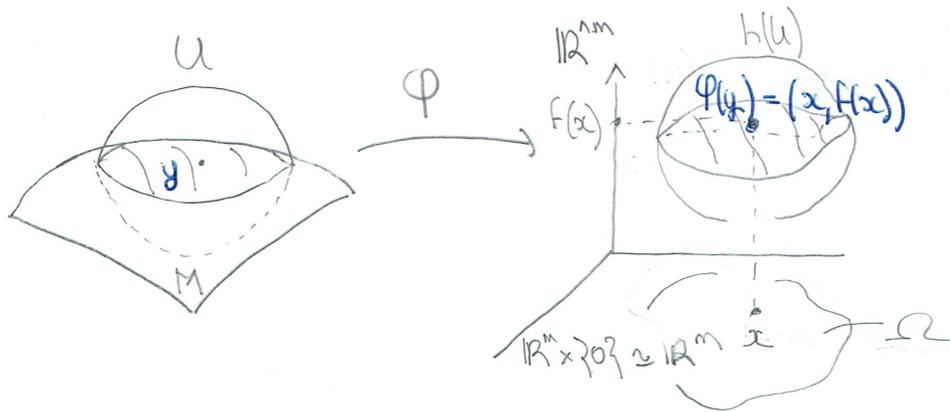
Le résultat suivant fournit les différentes caractérisations utiles d'un sous-ensemble M de \mathbb{R}^n comme sous-variété.

Théorème 3.1. *Un sous-ensemble M de \mathbb{R}^n est une sous-variété de dimension m si et seulement s'il vérifie une des propriétés équivalentes suivantes :*

- (1) **Grappe.** *En tout point de M , il existe un voisinage ouvert de ce point dans \mathbb{R}^n , un difféomorphisme $\varphi : U \rightarrow \varphi(U) \subset \mathbb{R}^n$ et $n - m$ fonctions f_{m+1}, \dots, f_n définies sur un ouvert Ω de \mathbb{R}^m tels que $y \in U \cap M$ si et seulement si il existe $x = (x_1, \dots, x_m) \in \Omega$ tel que*

$$\varphi(y) = (x_1, \dots, x_m, f_{m+1}(x_1, \dots, x_m), \dots, f_n(x_1, \dots, x_m)).$$

Autrement dit, M est donné localement, au difféomorphisme φ près, comme le graphe de la fonction $f = (f_{m+1}, \dots, f_n) : \Omega \subset \mathbb{R}^m \rightarrow \mathbb{R}^{n-m}$.

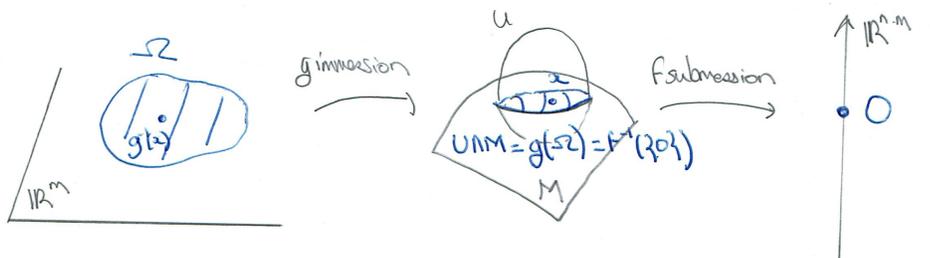


- (2) **Description implicite.** *Pour tout $x \in M$, il existe un voisinage U de x dans \mathbb{R}^n et une submersion $f : U \rightarrow \mathbb{R}^{n-m}$ telle que*

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

- (3) **Paramétrisation.** *Pour tout $x \in M$, il existe un voisinage U de x dans \mathbb{R}^n et une immersion $g : \Omega \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ telle que $g(\Omega) = U \cap M$ et g soit un homéomorphisme sur son image.*

Le couple (Ω, g) est appelé une paramétrisation de $M \cap U$.



Démonstration. (1) \Rightarrow [***M est une sous-variété***].

Pour cela il suffit de prouver que le graphe M d'une fonction $f = (f_{m+1}, \dots, f_n) : \Omega \subset \mathbb{R}^m \rightarrow \mathbb{R}^{n-m}$ est une sous-variété. Définissons pour cela $h : \Omega \times \mathbb{R}^{n-m} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ en posant

$$h(x) = (x_1, \dots, x_m, x_{m+1} - f_{m+1}(x_1, \dots, x_m), \dots, x_n - f_n(x_1, \dots, x_m))$$

pour tout point x de Ω . Alors la jacobienne de h est de forme triangulaire inférieure avec que des 1 sur la diagonale, et donc dh est inversible en tout point de son domaine de définition. Par le théorème d'inversion locale, au voisinage de chaque point de M il existe un voisinage U sur lequel h est un difféomorphisme local et on a pour tout $x \in U$

$$x \in M \Leftrightarrow \begin{cases} x_{m+1} = f_{m+1}(x_1, \dots, x_m) \\ \vdots \\ x_n = f_n(x_1, \dots, x_m) \end{cases} \Leftrightarrow h(x) \in \mathbb{R}^m \times \{0\}.$$

Autrement dit $x \in M \cap U \Leftrightarrow h(x) \in h(U) \cap (\mathbb{R}^m \times \{0\})$ et h est donc bien un difféomorphisme linéarisant.

[***M est une sous-variété***] \Rightarrow (1). Il existe un unique ouvert Ω de \mathbb{R}^m tel que $h(U \cap M) = \Omega \times \{0\}$. Posons $\varphi = h$ et

$$\begin{aligned} f_i : \Omega &\rightarrow \mathbb{R} \\ x &\mapsto f_i(x) = 0 \end{aligned}$$

pour $i = m+1, \dots, n$. Alors

$$\begin{aligned} y \in M \cap U &\Leftrightarrow h(y) \in h(U \cap M) \\ &\Leftrightarrow h(y) = (h_1(y), \dots, h_m(y), 0, \dots, 0) \text{ et } y \in U \\ &\Leftrightarrow h(y) \text{ appartient au graphe de } f = (f_{m+1}, \dots, f_n) : \Omega \rightarrow \mathbb{R}^{n-m}. \end{aligned}$$

(2) \Rightarrow [***M est une sous-variété***]. Par le théorème de redressement des submersions, il existe en tout point de M un voisinage ouvert U et un difféomorphisme $h : U \rightarrow h(U) \subset \mathbb{R}^n$ tel que

$$f \circ h^{-1}(x_1, \dots, x_n) = (x_{m+1}, \dots, x_n)$$

pour tout $x = (x_1, \dots, x_n) \in h(U)$. Ainsi

$$\begin{aligned} x \in M \cap U &\Leftrightarrow x \in U \text{ et } f(x) = 0 \in \mathbb{R}^{n-m} \\ &\Leftrightarrow x \in U \text{ et } f \circ h^{-1}(h(x)) = 0 \\ &\Leftrightarrow x \in U \text{ et } h(x) \in \mathbb{R}^m \times \{0\} \\ &\Leftrightarrow h(x) \in h(U) \cap (\mathbb{R}^m \times \{0\}) \end{aligned}$$

et donc h est un difféomorphisme linéarisant de M :

$$h(M \cap U) = h(U) \cap (\mathbb{R}^m \times \{0\}).$$

[***M est une sous-variété***] \Rightarrow (2). Si on note $h = (h_1, \dots, h_n)$ le difféomorphisme donné par la définition de sous-variété, alors l'application $f : U \rightarrow \mathbb{R}^{n-m}$ définie par $f(x) = (h_{m+1}(x), \dots, h_n(x))$ est une submersion. En effet, si on note $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$ la projection canonique sur les dernières coordonnées, alors $f = \pi \circ h$ et ainsi $df_x = \pi \circ dh_x$ qui est bien surjective.

On vérifie alors que

$$\begin{aligned} x \in M \cap U &\Leftrightarrow h(x) \in h(U \cap M) = h(U) \cap (\mathbb{R}^m \times \{0\}) \\ &\Leftrightarrow x \in U \text{ et } h(x) \in \mathbb{R}^m \times \{0\} \\ &\Leftrightarrow x \in U \text{ et } \pi \circ h(x) = 0 \\ &\Leftrightarrow x \in U \text{ et } f(x) = 0 \end{aligned}$$

et ainsi

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

(3) \Rightarrow [**M est une sous-variété**]. Par le théorème de redressement des immersions, il existe pour tout $x = g(x') \in M \cap U$ un voisinage ouvert $\Omega' \subset \Omega$ de x' , un voisinage ouvert $U' \subset U$ de x contenant $g(\Omega')$ et un difféomorphisme $h : U' \rightarrow h(U') \subset \mathbb{R}^n$ tel que

$$h \circ g(y_1, \dots, y_m) = (y_1, \dots, y_m, 0, \dots, 0)$$

pour tout $y_1, \dots, y_m \in \Omega'$. On peut supposer, quitte à le diminuer, que l'ouvert $h(U')$ est de la forme $\Omega' \times W$ où W est un ouvert de \mathbb{R}^{n-m} . On a alors

$$\begin{aligned} x \in M \cap U' &\Rightarrow \exists! x' \in \Omega \text{ tel que } x = g(x') \in U' \\ &\Rightarrow h(x) = h \circ g(x') \in h(U') \cap (\mathbb{R}^m \times \{0\}). \end{aligned}$$

Réciproquement, si $h(x) \in \mathbb{R}^m \times \{0\}$ pour $x \in U'$, alors $(h_1(x), \dots, h_m(x)) \in \Omega'$. Ainsi $h(x) = h(g(x'))$ ce qui implique que $x = g(x')$ comme h est un difféomorphisme local. Donc $x \in M$.

Nous concluons que h est un difféomorphisme linéarisant de M :

$$h(M \cap U') = h(U') \cap (\mathbb{R}^m \times \{0\}).$$

[**M est une sous-variété**] \Rightarrow (3). Nous pouvons écrire

$$h(U \cap M) = \Omega \times \{0\}$$

avec $\Omega \subset \mathbb{R}^m$ ouvert. On pose alors $g = h^{-1} \circ i : \Omega \rightarrow U$ où $i : \mathbb{R}^m \hookrightarrow \mathbb{R}^n$ désigne l'injection canonique. L'application est une immersion par construction. Quitte à diminuer Ω , on peut donc supposer que g est injective et donc bijective sur son image. Par construction,

$$g(\Omega) = M \cap U.$$

La continuité de $g^{-1} : M \cap U \rightarrow \Omega$ découle de ce que l'on peut choisir U borné ce qui implique que l'application g est fermée puisque continue, donc ouverte. \square

Un cas particulier de sous-variétés sont celles de codimension 1, ou de manière équivalente de dimension $n - 1$. Elles sont appelées *hypersurfaces*. On remarque aussi que l'image d'une courbe est une sous-variété de dimension 1 si et seulement si en chaque point elle ressemble à \mathbb{R} : en particulier elle n'admet pas de point double.

Voici quelques exemples classiques de sous-variétés.

Les sous-espaces affines. Les sous-espaces affines de dimension m sont des sous-variétés de dimension m car, par exemple, ils sont donnés comme le lieu des zéros de $n - m$ équations affines linéairement indépendantes. Les ouverts de sous-espaces affines sont des sous-variétés.

La sphère unité. La sphère unité

$$S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$$

est une hypersurface de \mathbb{R}^n , cad une sous-variété de dimension $n - 1$. En effet, on peut écrire

$$S^{n-1} = f^{-1}(0)$$

avec $f(x_1, \dots, x_n) = x_1^2 + \dots + x_n^2 - 1$ qui est classe C^∞ . On a $df_x = (2x_1, \dots, 2x_n)$ qui est non nulle dès que $x \in S^{n-1}$ et donc pour x dans un voisinage de S^{n-1} . Nous vérifions ainsi la caractérisation (2) du théorème précédent : f est une submersion au voisinage de chaque point de la sphère. Noter que S^{n-1} est compact.

Les tores. Dans \mathbb{R}^{2n} , l'ensemble suivant

$$\mathbb{T}^n = S^1 \times \dots \times S^1$$

est une sous-variété de dimension n . En effet, on peut écrire

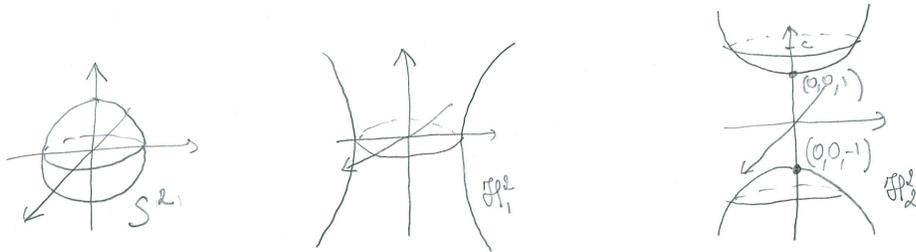
$$\mathbb{T}^n = \{x \in \mathbb{R}^{2n} \mid x_1^2 + x_2^2 = 1, \dots, x_{2n-1}^2 + x_{2n}^2 = 1\}$$

et on voit ainsi que $\mathbb{T}^n = f^{-1}(0)$ où $f(x) = (x_1^2 + x_2^2 - 1, \dots, x_{2n-1}^2 + x_{2n}^2 - 1)$ qui est une submersion au voisinage du tore. C'est une sous-variété compacte puisque $\mathbb{T}^n \subset S^{2n-1}$ est fermé.

L'hyperboloïde. L'ensemble suivant

$$\mathcal{H}_1^m = \{x \in \mathbb{R}^{m+1} \mid x_1^2 + \dots + x_m^2 - x_{m+1}^2 = 1\}$$

est une sous-variété de dimension m non compacte appelée hyperboloïde à une nappe. On peut la dessiner en dimension 3 en remarquant que les intersections avec les plans $\{x_3 = c\}$ sont des cercles de rayon $\sqrt{1 + c^2}$.

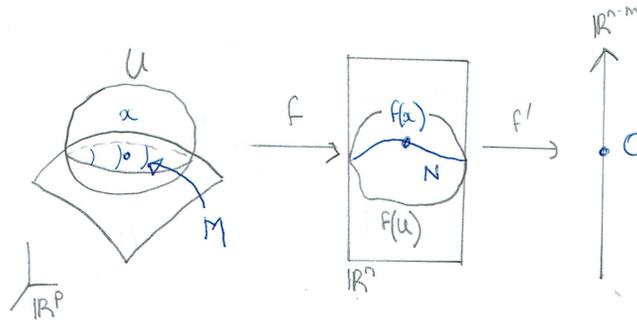


L'ensemble suivant

$$\mathcal{H}_2^m = \{x \in \mathbb{R}^{m+1} \mid x_1^2 + \dots + x_m^2 - x_{m+1}^2 = -1\}$$

est une sous-variété de dimension m non compacte appelée hyperboloïde à deux nappes. On peut la dessiner en dimension 3 en remarquant que les intersections avec les plans $\{x_3 = c\}$ sont des cercles de rayon $\sqrt{c^2 - 1}$ dès que $c^2 \geq 1$ et l'ensemble vide sinon.

Produit de deux sous-variétés. Nous pouvons construire des sous-variétés à partir d'autres sous-variétés de la manière suivante : si M_1, M_2 sont des sous-variétés de \mathbb{R}^{n_1} et de \mathbb{R}^{n_2} de dimension m_1 et m_2 respectivement, alors $M_1 \times M_2$ est une sous-variété de $\mathbb{R}^{n_1+n_2}$ de dimension $m_1 + m_2$. Cet exemple sera traité en exercice.



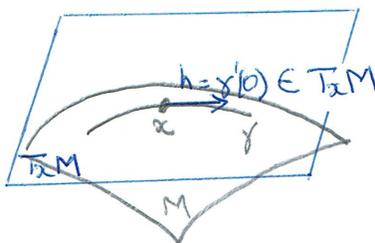
Préimage par une submersion. Soit $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ une submersion et N une sous-variété de \mathbb{R}^n de codimension $n - m$. Alors $M = f^{-1}(N)$ est une sous-variété de \mathbb{R}^p de codimension $n - m$. En effet, on se fixe $x \in M$, on choisit un voisinage ouvert V autour de $f(x) \in N$ et une submersion $f' : V \rightarrow \mathbb{R}^{n-m}$ telle que $f'^{-1}(\{0\}) = V \cap N$. Alors $f' \circ f$ est une submersion définie sur le voisinage ouvert $\tilde{U} = f^{-1}(V)$ autour de x tel que $(f' \circ f)^{-1}(\{0\}) = \tilde{U} \cap M$.

3.3. Espace tangent. On se donne une sous-variété M^m de \mathbb{R}^n et $x \in M$. Nous noterons $\mathcal{C}(x, M)$ l'ensemble des courbes γ sur M de classe C^1 et passant par x au temps $t = 0$:

$$\gamma :]-\varepsilon, \varepsilon[\rightarrow M \subset \mathbb{R}^n \text{ tel que } \gamma(0) = x.$$

Definition 3.2. L'espace tangent de M en x est l'ensemble défini comme

$$T_x M := \{\gamma'(0) \in \mathbb{R}^n \mid \gamma \in \mathcal{C}(x, M)\}.$$



Un exemple simple est lorsque $M = U \cap (\mathbb{R}^m \times \{0\}) \subset \mathbb{R}^n$ est déjà linéarisée. On observe qu'alors, pour tout $x \in M$ et $h \in \mathbb{R}^m \times \{0\}$ la courbe $\gamma(t) = x + th$ est contenue dans M pourvu que t soit suffisamment proche de 0. On obtient alors l'inclusion

$$\mathbb{R}^m \times \{0\} \subset T_x M$$

qui est en fait une égalité, l'inclusion réciproque étant triviale.

Cette remarque élémentaire nous amène alors à la proposition suivante :

Proposition 3.2. L'espace tangent $T_x M$ est un sous-espace vectoriel de dimension m de \mathbb{R}^n . Plus précisément,

(1) Si h est un difféomorphisme local linéarisant M au voisinage de x , on a

$$T_x M = dh_{h(x)}^{-1}(\mathbb{R}^m \times \{0\}).$$

(2) Si $f = (f_{m+1}, \dots, f_n) : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$ est une submersion définie au voisinage de x telle que $U \cap M = f^{-1}(0)$, alors

$$T_x M = \ker df_x = \bigcap_{i=m+1}^n \ker d(f_i)_x.$$

(3) Si M est donnée localement comme le graphe d'une fonction $f : \Omega \subset \mathbb{R}^m \rightarrow \mathbb{R}^{n-m}$, alors

$$T_{(x, f(x))} M = \{(h, k) \in \mathbb{R}^m \times \mathbb{R}^{n-m} \mid k = df_x(h)\}.$$

(4) Si (Ω, g) est une paramétrisation locale de M — g est une immersion et un homéomorphisme de Ω sur $M \cap U$ —, alors

$$T_x M = dg_{g^{-1}(x)}(\mathbb{R}^m).$$

Démonstration. Observons tout d'abord que si M est une sous-variété et $h : U \rightarrow V$ est un difféomorphisme, alors l'espace tangent en $h(x)$ de $N = h(U \cap M)$ est donné comme

$$T_{h(x)} N = dh_x(T_x M).$$

En effet, si $\gamma \in \mathcal{C}(x, M)$, alors $h \circ \gamma \in \mathcal{C}(h(x), N)$. Il reste alors à remarquer que

$$(h \circ \gamma)'(0) = dh_{\gamma(0)}(\gamma'(0)) = dh_x(\gamma'(0))$$

et ainsi $dh_x(T_x M) \subset T_{h(x)} N$. De même, nous obtenons $(dh_x)^{-1}(T_{h(x)} N) \subset T_x M$ ce qui implique l'inclusion réciproque.

(1) Nous pouvons donc écrire

$$T_x M = dh_{h(x)}^{-1}(\mathbb{R}^m \times \{0\})$$

où h est un difféomorphisme linéarisant. Ainsi, d'une part l'espace tangent est bien un sous-espace vectoriel de dimension m , et d'autre part la formule annoncée est vérifiée.

(2) Pour toute courbe $\gamma \in C(x, M)$, on a pour tout $t \in]-\varepsilon, \varepsilon[$ que $f_i(\gamma(t)) = 0$ pour $i = m+1, \dots, n$. Ceci implique $d(f_i)_x(\gamma'(0)) = 0$ en dérivant. Donc

$$T_x M \subset \bigcap_{i=m+1}^n \ker d(f_i)_x = \ker df_x$$

et on conclut grâce à la dimension et le théorème du rang, $df_x : \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$ étant une application surjective.

(3) On pose $F(x, y) = y - f(x)$ qui est une submersion définissant M comme $F^{-1}(0) = M \cap U$, puis on applique le calcul du (2) en remarquant que

$$dF_{(x, f(x))} = \pi_{\mathbb{R}^{n-m}} - df_x \circ \pi_{\mathbb{R}^m}$$

et donc que $\ker dF_{(x, f(x))} = \{(h, k) \in \mathbb{R}^m \times \mathbb{R}^{n-m} \mid k = df_x(h)\}$

(4) Si $\gamma :]-\varepsilon, \varepsilon[\rightarrow \Omega$ passe par $g^{-1}(x)$, alors $g \circ \gamma \in C(x, M)$ et donc

$$dg_{g^{-1}(x)}(\mathbb{R}^m) \subset T_x M.$$

On conclut alors en utilisant la dimension. □

L'espace tangent affine est le translaté de $T_x M$ suivant :

$$\mathcal{T}_x M = x + T_x M.$$

C'est le sous-espace affine approximant le mieux M au voisinage de x , et il est défini avec les notations du (2) dans la proposition précédente par le système d'équations

$$\begin{cases} \frac{\partial f_1}{\partial x_1}(x)(y_1 - x_1) + \dots + \frac{\partial f_1}{\partial x_n}(x)(y_n - x_n) = 0 \\ \dots \\ \frac{\partial f_{n-m}}{\partial x_1}(x)(y_1 - x_1) + \dots + \frac{\partial f_{n-m}}{\partial x_n}(x)(y_n - x_n) = 0 \end{cases}$$

où $y = (y_1, \dots, y_n)$.

Par exemple, pour la sphère unité en un point x

$$T_x S^{n-1} = \{y \in \mathbb{R}^n \mid \sum_{i=1}^n x_i y_i = 0\} \text{ et } \mathcal{T}_x S^{n-1} = \{y \in \mathbb{R}^n \mid \sum_{i=1}^n x_i y_i = 1\}.$$

Pour le tore $\mathbb{T}^n = S^1 \times \dots \times S^1$, on peut écrire

$$\begin{aligned} T_x \mathbb{T}^n &= \{y \in \mathbb{R}^{2n} \mid x_1 y_1 + x_2 y_2 = 0, \dots, x_{2n-1} y_{2n-1} + x_{2n} y_{2n} = 0\} \\ &= T_{(x_1, x_2)} S^1 \times \dots \times T_{(x_{2n-1}, x_{2n})} S^1 \end{aligned}$$

Nous verrons d'ailleurs qu'en général, si M_1 et M_2 sont deux sous-variétés, alors pour tout $(x_1, x_2) \in M_1 \times M_2$

$$T_{(x_1, x_2)} M_1 \times M_2 = T_{x_1} M_1 \times T_{x_2} M_2.$$

Definition 3.3. On définit l'espace tangent comme la réunion

$$TM = \bigcup_{x \in M} (\{x\} \times T_x M) \subset \mathbb{R}^n \times \mathbb{R}^n.$$

Proposition 3.3. TM est une sous-variété de \mathbb{R}^{2n} de dimension $2m$.

Démonstration. Soit $f : U \rightarrow \mathbb{R}^{n-m}$ une submersion telle que $M \cap U = f^{-1}(0)$. Alors

$$TM \cap (U \times \mathbb{R}^n) = \{(x, y) \in \mathbb{R}^{2n} \mid f(x) = 0 \text{ et } df_x(y) = 0\}.$$

Posons alors

$$\begin{aligned} F : U \times \mathbb{R}^n &\rightarrow \mathbb{R}^{2(n-m)} \\ (x, y) &\mapsto (f(x), df_x(y)) \end{aligned}$$

C'est une application de classe C^∞ et on voit que

$$Jac_{(x,y)}F = \begin{pmatrix} Jac_x f & 0 \\ Jac_x df \cdot (y) & Jac_x f \end{pmatrix}.$$

Les deux blocs diagonaux étant de rang $n - m$ (penser en terme de vecteurs colonnes indépendants), la matrice est de rang $2(n - m)$ et donc F est une submersion définissant localement TM . \square

On remarquera ici, que si nous avons travaillé avec la notion de sous-variété de classe C^r pour $r \in \mathbb{N}^*$, nous aurions obtenu que l'espace tangent était une sous-variété de classe C^{r-1} .

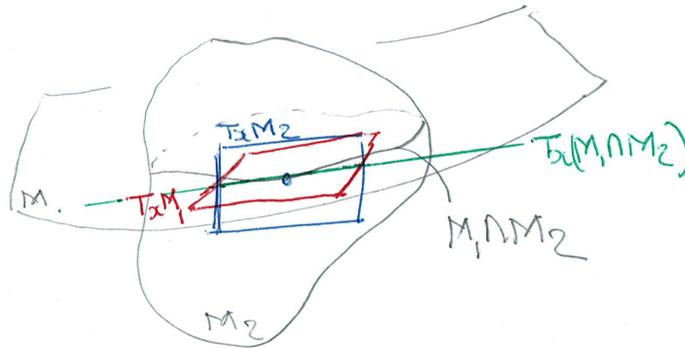
Par intersection de sous-variétés, nous pouvons produire de nouvelles sous-variétés à la condition suivante.

Proposition 3.4. Soient M_1 et M_2 deux sous-variétés de \mathbb{R}^n en position transverse, c'est-à-dire telles que pour tout $x \in M_1 \cap M_2$

$$\mathbb{R}^n = T_x M_1 + T_x M_2.$$

Alors $M_1 \cap M_2$ est une sous-variété de \mathbb{R}^n satisfaisant

- (1) $\text{codim}(M_1 \cap M_2) = \text{codim}(M_1) + \text{codim}(M_2)$,
- (2) $T_x(M_1 \cap M_2) = T_x M_1 \cap T_x M_2$.



Démonstration. Soit $x \in M_1 \cap M_2$. Pour $i = 1, 2$, soit $p_i = \text{codim}(M_i)$ et $f_i : U_i \rightarrow \mathbb{R}^{p_i}$ une submersion telle que $x \in U_i \cap M_i = f_i^{-1}(0)$. Nous posons $f = (f_1, f_2) : U_1 \cap U_2 \rightarrow \mathbb{R}^{p_1+p_2}$. On a alors

$$\ker df_x = \ker d(f_1)_x \cap \ker d(f_2)_x \text{ et } f^{-1}(0) = (U_1 \cap U_2) \cap (M_1 \cap M_2).$$

Observons alors que

$$\begin{aligned} \dim(\ker df_x) &= \dim(\ker d(f_1)_x \cap \ker d(f_2)_x) \\ &= \dim(\ker d(f_1)_x) + \dim(\ker d(f_2)_x) - \dim(\ker d(f_1)_x + \ker d(f_2)_x) \\ &= \dim M_1 + \dim M_2 - n \\ &= n - (p_1 + p_2). \end{aligned}$$

On en déduit que f est une submersion et donc $M_1 \cap M_2$ est une sous-variété de codimension $p_1 + p_2$. On a alors comme annoncé

$$T_x(M_1 \cap M_2) = \ker df_x = \ker d(f_1)_x \cap \ker d(f_2)_x = T_x M_1 \cap T_x M_2.$$

□

3.4. Application différentiable sur les sous-variétés et application tangente.

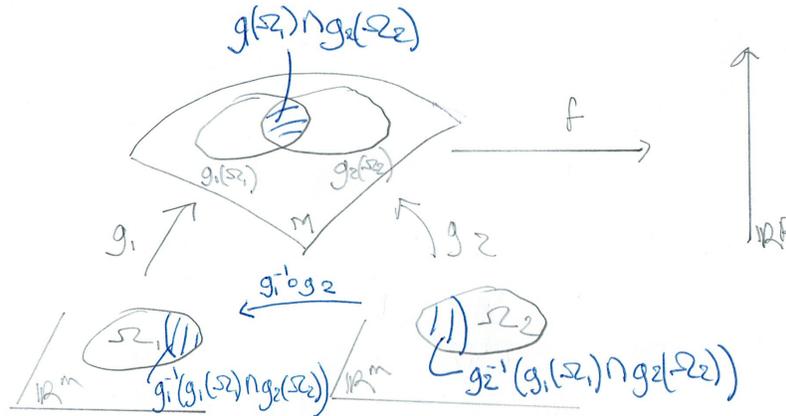
Definition 3.4. Soit M^m une sous-variété de \mathbb{R}^n et $r \in \mathbb{N} \cup \{+\infty\}$. Une application $f : M \rightarrow \mathbb{R}^p$ est dite de classe C^r si pour tout paramétrisation locale g de M la composition $f \circ g$ est de classe C^r .

Par exemple, si $f : U \rightarrow \mathbb{R}^p$ est une application de classe C^r définie sur un ouvert U contenant M , alors $f|_M : M \rightarrow \mathbb{R}^p$ est de classe C^r . Un exemple important est donné par les fonctions coordonnées de \mathbb{R}^n .

En pratique, nous avons besoin du critère suivant pour vérifier qu'une application est bien C^r .

Proposition 3.5. Soit M une sous-variété de \mathbb{R}^n et $\{g_i : \Omega_i \rightarrow \mathbb{R}^n\}_{i \in I}$ une collection finie de paramétrisations locales de M telle que $M = \cup_{i \in I} g_i(\Omega_i)$. Une application $f : M \rightarrow \mathbb{R}^p$ est de classe C^r ssi $f \circ g_i$ est de classe C^r pour tout $i \in I$.

Démonstration. Il suffit de remarquer que si $g_1 : \Omega_1 \rightarrow \mathbb{R}^n$ et $g_2 : \Omega_2 \rightarrow \mathbb{R}^n$ sont deux paramétrisations locales de M telles que $g_1(\Omega_1) \cap g_2(\Omega_2) \neq \emptyset$, et si $f \circ g_1$ est de classe C^r , alors automatiquement $f \circ g_2$ est de classe C^r sur $g_2^{-1}(g_1(\Omega_1) \cap g_2(\Omega_2))$.



Ceci découle de ce que l'application

$$g_1^{-1} \circ g_2 : g_2^{-1}(g_1(\Omega_1) \cap g_2(\Omega_2)) \rightarrow g_1^{-1}(g_1(\Omega_1) \cap g_2(\Omega_2))$$

est un difféomorphisme, comme $f \circ g_2 = (f \circ g_1) \circ (g_1^{-1} \circ g_2)$ sur $g_2^{-1}(g_1(\Omega_1) \cap g_2(\Omega_2))$. Pour cela, comme c 'est un homéomorphisme, il suffit de prouver qu'il est de classe C^∞ , et, son inverse ayant la même forme, il sera lui aussi de classe C^∞ .

Pour tout $x \in g_1(\Omega_1) \cap g_2(\Omega_2)$ il existe un difféomorphisme linéarisant h au voisinage de x et alors

$$g_1^{-1} \circ g_2 = g_1^{-1} \circ h^{-1} \circ h \circ g_2 = (h \circ g_1)^{-1} \circ (h \circ g_2).$$

Si nous notons $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ la projection sur le premier facteur de la décomposition $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}$, nous obtenons

$$g_1^{-1} \circ g_2 = g_1^{-1} \circ h^{-1} \circ h \circ g_2 = (\pi \circ h \circ g_1)^{-1} \circ (\pi \circ h \circ g_2).$$

Nous savons déjà que l'application $\pi \circ h \circ g_i : \Omega_i \subset \mathbb{R}^m \rightarrow \mathbb{R}^m$ est de classe C^∞ . Or l'application $\pi_{\mathbb{R}^m} \circ h \circ g_i : \Omega_i \rightarrow \mathbb{R}^m$ est de différentielle inversible au voisinage de tout point, comme

$$\begin{aligned} d(\pi_{\mathbb{R}^m} \circ h \circ g_i)_x(\mathbb{R}^m) &= \pi_{\mathbb{R}^m} \circ dh_{g_i(x)} \circ d(g_i)_x(\mathbb{R}^m) \\ &= \pi_{\mathbb{R}^m} \circ dh_{g_i(x)}(T_{g_i(x)}M) \\ &= \pi_{\mathbb{R}^m}(\mathbb{R}^m \times \{0\}) \\ &= \mathbb{R}^m. \end{aligned}$$

Par le théorème d'inversion locale, $\pi_{\mathbb{R}^m} \circ h \circ g_i : \Omega_i \rightarrow \mathbb{R}^m$ est donc un difféomorphisme local de classe C^∞ au voisinage de chaque point. Ceci implique que $(\pi_{\mathbb{R}^m} \circ h \circ g_1)^{-1}$ est de classe C^∞ sur $\pi \circ h(g_1(\Omega_1) \cap g_2(\Omega_2))$. On obtient ainsi que $g_1^{-1} \circ g_2 = (\pi \circ h \circ g_1)^{-1} \circ (\pi \circ h \circ g_2)$ est de classe C^∞ et donc un difféomorphisme. \square

Ainsi, nous disposons de la notion d'application de classe C^r pour toute application $f : M \subset \mathbb{R}^p \rightarrow N \subset \mathbb{R}^n$ entre deux sous-variétés.

Definition 3.5. Deux sous-variétés M et N sont dites C^r -difféomorphes si il existe une application bijective $f : M \rightarrow N$ telle que f et f^{-1} sont de classe C^r .

Maintenant, nous pouvons généraliser la notion de différentielle aux applications entre deux sous-variétés comme suit.

Proposition 3.6. Soit $f : M \rightarrow N$ une application de classe C^1 entre deux sous-variétés $M \subset \mathbb{R}^n$ et $N \subset \mathbb{R}^p$. Alors pour tout point $x \in M$ l'application définie par la formule

$$\begin{aligned} T_x f : T_x M &\rightarrow T_{f(x)} N \\ v = \dot{\gamma}(0) &\mapsto \frac{d(f \circ \gamma)}{dt}(0) \end{aligned}$$

est une application linéaire appelée application tangente de f en x .

En particulier, si $M = U$ and $N = V$ sont deux ouverts, alors $T_x f = df_x$.

Démonstration. Vérifions que la définition a bien un sens, et que l'application induite est bien linéaire. Pour cela, soit h est un difféomorphisme linéarisant de M au voisinage de x , cad le difféomorphisme $h : U \rightarrow h(U)$ vérifie

$$h(U \cap M) = h(U) \cap (\mathbb{R}^m \times \{0\}).$$

Quitte à diminuer U , on peut supposer que $h(U) = \Omega_1 \times \Omega_2 \subset \mathbb{R}^m \times \mathbb{R}^{n-m}$. On commence par étendre f localement en une fonction $\tilde{f} : U \rightarrow N$ en posant pour tout $z \in U$

$$\tilde{f}(z) = \tilde{f} \circ h^{-1}(x, y) = f \circ h^{-1}(x, 0)$$

où $h(z) = (x, y) \in h(U) \subset \mathbb{R}^m \times \mathbb{R}^{n-m}$. Nous pouvons également écrire $\tilde{f} = f \circ h^{-1} \circ p \circ h$ avec $p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ l'application définie par $p(x, y) = (x, 0)$. Ainsi, puisque l'application h^{-1} restreinte à $h(U) \cap (\mathbb{R}^m \times \{0\})$ est une paramétrisation locale de M , $f \circ h^{-1}$ restreinte à $p \circ h(U) = h(U) \cap (\mathbb{R}^m \times \{0\})$ est de classe C^1 . Nous en déduisons que \tilde{f} est bien C^1 sur U comme composée d'applications C^1 .

Ainsi, la formule proposée a bien un sens, puisque si $\gamma :]-\varepsilon, \varepsilon[\rightarrow U \cap M$ est une courbe de classe C^1 alors la composée $f \circ \gamma = \tilde{f} \circ \gamma :]-\varepsilon, \varepsilon[\rightarrow N$ est elle-aussi de classe C^1 . De plus, on a bien que l'image de $T_x f$ est contenue dans $T_{f(x)} N$. On a alors

$$T_x f(v) = d\tilde{f}_x(v)$$

pour tout $v \in T_x M$. Ceci nous donne la linéarité. \square

On voit facilement que si $f : M \rightarrow N$ et $g : N \rightarrow L$ sont deux applications C^1 entre sous-variétés, alors $g \circ f$ est aussi C^1 et

$$T(g \circ f)_x = T_{f(x)} g \circ T_x f.$$

3.5. Champs de vecteurs sur les sous-variétés. Soit U un ouvert de \mathbb{R}^n et M une sous-variété contenue dans U . On considère un champ de vecteurs $f : U \rightarrow \mathbb{R}^n$ de classe C^1 .

Definition 3.6. *Le champ de vecteurs f est dit tangent à M si pour tout $x \in M$ on a $f(x) \in T_x M$.*

Voici un exemple de champ de vecteurs sur la sphère unité de dimension impaire. On définit $f : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ en posant

$$f(x_1, x_2, \dots, x_{2n-1}, x_{2n}) = (-x_2, x_1, \dots, -x_{2n}, x_{2n-1}).$$

On voit facilement que $\langle x, f(x) \rangle = 0$ et donc pour tout $x \in S^{2n-1}$ on a bien $f(x) \in T_x S^{2n-1}$. On observe que ce champ ne s'annule nul part sur la sphère. On observe d'ailleurs que ce champ de vecteurs est également tangent au tore $\mathbb{T}^n \subset S^{2n-1}$ par construction.

Proposition 3.7. *Soit $f : U \rightarrow \mathbb{R}^n$ un champ de vecteurs de classe C^1 tangent à une sous-variété M fermée. Alors les courbes intégrales de f démarrant en un point de M sont entièrement contenues dans M .*

Démonstration. Soit $g : \Omega \subset \mathbb{R}^m \rightarrow U' \cap M \subset \mathbb{R}^n$ une paramétrisation locale de M au voisinage d'un point quelconque. Comme dg_x est injective, et que f est tangent à M , on peut tirer en arrière le champ de vecteurs f et ainsi obtenir un champ de vecteurs \tilde{f} sur Ω . Plus précisément, pour tout $x \in g(\Omega)$, l'application $dg_{g^{-1}(x)} : T_{g^{-1}(x)}\Omega = \mathbb{R}^m \rightarrow T_x M$ est un isomorphisme, et on définit le tiré en arrière de f sur Ω comme l'unique champ de vecteurs \tilde{f} tel que

$$dg_y(\tilde{f}(y)) = f(g(y))$$

pour tout $y \in \Omega$. Ce champ de vecteurs est de classe C^1 par le théorème des fonctions implicites, et admet donc une courbe intégrale maximale $\tilde{\gamma}$ passant par $y = g^{-1}(x)$. Posons $\gamma = g \circ \tilde{\gamma}$. Alors $\gamma(0) = x$ et on calcule que

$$\gamma'(t) = dg_{\tilde{\gamma}(t)}(\tilde{\gamma}'(t)) = dg_{\tilde{\gamma}(t)}(\tilde{f}(\tilde{\gamma}(t))) = f(\gamma(t))$$

et donc que γ est une courbe intégrale à f démarrant en x . Par construction $\gamma \subset M$. Soit maintenant γ_x la courbe intégrale maximale de f démarrant en x . Si la courbe sortait de M , il existerait un plus grand temps $t_0 < \infty$ tel que $\gamma_x(t) \in M$ pour $t \in [0, t_0[$. Mais par fermeture de M on aurait que $\gamma_x(t_0) \in M$ et donc en intégrant localement le champ de vecteurs autour de ce point on voit que $\gamma_x(t) \in M$ pour t au voisinage de t_0 , d'où une contradiction sur la maximalité de t_0 . \square

3.6. Point critique Vs. point régulier. Nous allons produire de nouvelles sous-variétés en prenant l'image inverse de certaines valeurs d'une fonction $f : M \rightarrow N$ lisse (c'est-à-dire de classe C^∞). Pour cela, nous introduisons la définition suivante :

Definition 3.7. *Soit $f : M \rightarrow N$ un application lisse entre deux sous-variétés. Un point $x \in M$ est dit critique si l'application tangente en x n'est pas surjective. Un point qui n'est pas critique est dit régulier. On note*

$$\mathcal{C} = \{x \in M \mid T_x f \text{ n'est pas surjective}\}$$

l'ensemble des points critiques de f . L'ensemble $f(\mathcal{C})$ s'appelle alors l'ensemble des valeurs critiques et son complémentaire $N \setminus f(\mathcal{C})$ l'ensemble des valeurs régulières.

Attention de remarquer que pour que $y \in N$ soit une valeur critique il suffit d'un point critique dans sa préimage. Un premier intérêt de distinguer valeur critique de valeur régulière est le résultat suivant.

Proposition 3.8. *Soit $f : M \rightarrow N$ une application de classe C^∞ entre deux sous-variétés $M \subset \mathbb{R}^n$ et $N \subset \mathbb{R}^p$. Si y est une valeur régulière de f , alors l'ensemble $f^{-1}(y) \subset M$ est une sous-variété de \mathbb{R}^n de dimension $\dim M - \dim N$.*

Démonstration. Soit $x \in f^{-1}(y)$. On considère une paramétrisation locale $g : \Omega \subset \mathbb{R}^{\dim M} \rightarrow \mathbb{R}^n$ de M au voisinage de x et un difféomorphisme $h : U \rightarrow V$ linéarisant de N au voisinage de y tel que $h(y) = 0$. D'une part, comme $h(U \cap N) = h(U) \cap (\mathbb{R}^{\dim N} \times \{0\})$, l'application $\pi_{\mathbb{R}^{\dim N}} \circ h : U \cap N \rightarrow \mathbb{R}^{\dim N}$ est bijective sur son image et son application tangente est bijective en tout point de U . D'autre part, la paramétrisation locale vue comme application de Ω dans M vérifie que $T_{g^{-1}(x)}g$ est un isomorphisme de $T_{g^{-1}(x)}\Omega = \mathbb{R}^{\dim M}$ sur T_xM . Ainsi la composée $(\pi_{\mathbb{R}^{\dim N}} \circ h) \circ f \circ g$ est une submersion sur $\mathbb{R}^{\dim N}$ sur son domaine de définition et donc la préimage de 0 à savoir $g^{-1}(f^{-1}(y))$ est une sous-variété de dimension $\dim M - \dim N$ au voisinage de $g^{-1}(x)$, il en est de même pour $f^{-1}(y)$ au voisinage de x par composition avec l'immersion g (l'image d'une sous-variété par une immersion étant encore une sous-variété). \square

On voit facilement que

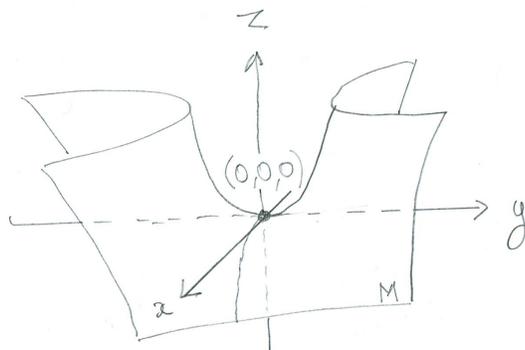
$$T_x(f^{-1}(y)) = \ker T_x f.$$

Il se trouve que les valeurs sont régulières génériquement dans le sens suivant : l'ensemble des valeurs régulières d'une application lisse $f : M \rightarrow N$ est un ensemble partout dense dans N . En particulier, on peut toujours assurer cette condition de régularité d'une valeur quitte à bouger un petit peu la valeur en question.

Le second intérêt que l'on peut présenter ici vient du calcul des variations et de la recherche d'extremum locaux.

Proposition 3.9. Soit $f : M \rightarrow \mathbb{R}$ une fonction de classe C^1 et x un point de M . Si x est un extremum local de f , x est un point critique.

La réciproque de ce résultat est bien évidemment fautive : prendre la fonctionnelle hauteur $f(x, y, z) = z$ définie sur la sous-variété $M = \{y^2 - x^2 = z\}$ appelée selle de cheval. L'origine est alors un point critique, mais n'est ni un minimum, ni un maximum local.



Démonstration. En effet, souvenons-nous qu'être extremum local signifie qu'il existe un voisinage U de x dans \mathbb{R}^n pour lequel pour tout $x' \in U \cap M$ nous avons $f(x) \leq f(x')$ dans le cas d'un minimum local, et $f(x) \geq f(x')$ dans le cas d'un maximum local. La question étant locale, on peut supposer que f est donnée localement comme la restriction d'une application $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ définie sur l'ouvert U voisinage de x .

Si $T_x f \neq 0$, on peut alors choisir un $v = \gamma'(0) \in T_x M$ tel que $T_x f(v) = df_x(v) = 1$. On a alors

$$\begin{aligned} f(\gamma(t)) &= f(\gamma(0) + vt + o(t)) \\ &= f(x) + df_x(tv + o(t)) + o(t) \\ &= f(x) + t + o(t) \end{aligned}$$

et donc x ne peut pas être un extremum local. On a donc nécessairement $T_x f = 0$. Le point x est un point critique pour f . \square

Le résultat suivant permet de trouver l'ensemble des points critiques d'une fonction définie sur une sous-variété.

Théorème 3.2 (des extrema liés). *Soit M une sous-variété définie au voisinage de l'un de ces points x par une submersion $g = (g_1, \dots, g_p) : U \rightarrow \mathbb{R}^p$. On se donne une fonction $f : U \rightarrow \mathbb{R}$ de classe C^1 .*

Alors x est un point critique de $f|_{U \cap M}$ si et seulement si il existe des réels $\lambda_1, \dots, \lambda_p$ tels que

$$df_x = \sum_{i=1}^p \lambda_i d(g_i)_x.$$

Les nombres $\lambda_1, \dots, \lambda_p$ sont appelés multiplicateurs de Lagrange.

Démonstration. Le point x est critique pour $f|_{U \cap M}$ si et seulement si l'application tangente n'est pas surjective, cad dans ce cas nulle. Ainsi x point critique ssi pour tout vecteur $v \in T_x M$ on $df_x(v) = 0$. Mais

$$T_x M = \bigcap_{i=1}^p \ker d(g_i)_x$$

et donc ceci est équivalent à avoir

$$\bigcap_{i=1}^p \ker d(g_i)_x \subset \ker df_x.$$

Ceci implique alors l'existence de réels $\lambda_1, \dots, \lambda_p$ tels que

$$df_x = \sum_{i=1}^p \lambda_i d(g_i)_x.$$

En effet, g est une submersion donc l'application linéaire $dg_x = (d(g_1)_x, \dots, d(g_p)_x)$ est de rang p et donc que la famille $(d(g_1)_x, \dots, d(g_p)_x)$ de formes linéaires de $(\mathbb{R}^n)^*$ est de rang p donc libre. On la complète en une base $(\varphi_1 = d(g_1)_x, \dots, \varphi_p = d(g_p)_x, \varphi_{p+1}, \dots, \varphi_n)$ et on écrit df_x dans cette base : il existe $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tels que

$$df_x = \sum_{i=1}^n \lambda_i \varphi_i = \sum_{i=1}^p \lambda_i d(g_i)_x + \sum_{i=p+1}^n \lambda_i \varphi_i$$

On note alors (y_1, \dots, y_n) la base antéduale, cad vérifiant $\varphi_i(y_j) = \delta_{ij}$. On a alors $y_j \in \bigcap_{i=1}^p \ker d(g_i)_x \subset \ker df_x$ pour tout $j = p+1, \dots, n$ dont on tire $\lambda_j = 0$ pour tout $j = p+1, \dots, n$. \square

Exemple d'application : l'isopérimétrie du triangle.

Ainsi, pour chercher un extremum, nous allons commencer par chercher les points critiques en cherchant les points admettant des multiplicateurs de Lagrange.

Pour se convaincre de la pertinence de la méthode, nous allons rechercher le triangle d'aire maximale à périmètre fixé $p > 0$. D'après la formule de Héron, si (x, y, z) est le triplet des longueurs des côtés du triangle, son aire est égale à

$$\frac{1}{4} \sqrt{(x^2 + y^2 + z^2)^2 - 2(x^4 + y^4 + z^4)}.$$

En effet, soit ABC un triangle de côté (x, y, z) . On note H le projeté orthogonal de B sur AC . D'après Pythagore,

$$\begin{aligned} x^2 &= (y - AH)^2 + BH^2 \\ &= (y - AH)^2 + z^2 - AH^2 \\ &= y^2 + z^2 - 2yAH \end{aligned}$$

d'où

$$AH = \frac{y^2 + z^2 - x^2}{2y}.$$

Comme son aire est donnée par la formule $S = \frac{1}{2}yBH$, on obtient

$$\begin{aligned} 16S^2 &= 4y^2BH^2 \\ &= 4y^2(z^2 - AH^2) \\ &= 4y^2\left(z^2 - \frac{1}{4y^2}(y^2 + z^2 - x^2)^2\right) \\ &= 4y^2z^2 - (y^2 + z^2 - x^2)^2 \\ &= 4y^2z^2 - y^4 - z^4 - x^4 - 2y^2z^2 + 2x^2y^2 + 2x^2z^2 \\ &= (x^2 + y^2 + z^2)^2 - 2(x^4 + y^4 + z^4). \end{aligned}$$

Nous allons donc maximiser la fonction

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

sous la contrainte

$$g(x, y, z) = x + y + z - p = 0.$$

Un triangle n'est défini, pour un triplet (x, y, z) , que si les trois coordonnées sont positives et si la somme de deux coordonnées est supérieure à la troisième. On pose donc

$$D = \{(x, y, z) \mid 0 \leq x \leq y + z, 0 \leq y \leq x + z \text{ et } 0 \leq z \leq x + y\}.$$

Sur la frontière de D , la fonction f est nulle (les triangles correspondants sont aplatis !). On cherche donc un point de l'intérieur de D tel que f soit maximale dans l'ensemble des points d'image par g nulle. Comme l'intersection de l'image réciproque de 0 par g et de D est un compact, il existe au moins un maximum.

On commence donc par chercher les points critiques de f sur $\overset{\circ}{D} \cap g^{-1}(0)$.

Comme g est une submersion sur $\overset{\circ}{D}$, l'ensemble $\overset{\circ}{D} \cap g^{-1}(0)$ est une sous-variété de \mathbb{R}^3 et nous pouvons appliquer le théorème des multiplicateurs. Ainsi un point $(x, y, z) \in \overset{\circ}{D} \cap g^{-1}(0)$ est critique pour $f|_{\overset{\circ}{D} \cap g^{-1}(0)}$ s'il existe λ tel que

$$df_{(x,y,z)} = \lambda dg_{(x,y,z)}.$$

Cette condition est équivalente au système d'équations

$$\begin{cases} 4x(-x^2 + y^2 + z^2) = \lambda, \\ 4y(x^2 - y^2 + z^2) = \lambda, \\ 4z(x^2 + y^2 - z^2) = \lambda, \\ x + y + z = p. \end{cases}$$

Les deux premières équations en utilisant la relation $x + y + z = p$ donnent $x = y$ ou $x + y = p/2$. On a donc les conditions suivantes

$$\begin{cases} x = y \text{ ou } x + y = p/2, \\ y = z \text{ ou } y + z = p/2, \\ x = z \text{ ou } x + z = p/2, \\ x + y + z = p, \end{cases}$$

On vérifie alors que la seule solution sur l'intérieur de D est pour $x = y = z = p/3$ correspondant au triangle équilatéral. Un seul point critique à l'intérieur de D donc c'est un maximum global.

4. SECONDE INCURSION DANS LE MONDE DES ÉQUATIONS DIFFÉRENTIELLES

4.1. **Théorème de redressement d'un champ de vecteurs.** Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ un champ de vecteurs de classe C^1 .

Définition 4.1. On appelle *intégrale première* de f une fonction $H \in C^1(U, \mathbb{R})$ telle que

$$dH_x(f(x)) = 0.$$

Par la formule de dérivation des fonctions composées, pour toute courbe intégrale $\gamma_x : J(x) \rightarrow \mathbb{R}^n$ de f on a

$$(H \circ \gamma_x)'(t) = dH_{\gamma_x(t)}(\dot{\gamma}_x(t)) = dH_{\gamma_x(t)}(f(\gamma_x(t))) = 0$$

et donc la fonction H est une intégrale première de f ssi elle est constante le long des courbes intégrales de f .

Définition 4.2. Une famille d'intégrales premières H_1, \dots, H_k de f est dite *indépendante* si en tout point $x \in U$ les formes différentielles $d(H_1)_x, \dots, d(H_k)_x$ sont linéairement indépendantes.

De manière équivalente, les intégrales premières H_1, \dots, H_k sont indépendantes si l'application $\mathcal{H} = (H_1, \dots, H_k) : U \rightarrow \mathbb{R}^k$ est une submersion. Alors pour tout $x \in U$ la courbe intégrale de f démarrant en x est contenue dans $\mathcal{H}^{-1}(\mathcal{H}(x))$ qui est une sous-variété de dimension $n - k$. En particulier, on obtient ainsi que $k \leq n - 1$, autrement dit :

Proposition 4.1. Un champ de vecteurs admet au plus $n - 1$ intégrales premières indépendantes.

On comprend donc que la détermination d'intégrales premières permet de cerner la position géométrique des courbes intégrales, et donc de résoudre l'équation différentielle associée. Optimalement, on trouve $n - 1$ intégrales premières linéairement indépendantes, et donc la courbe intégrale est déterminée géométriquement—cad à paramétrisation près— comme la composante connexe de $\mathcal{H}^{-1}(\mathcal{H}(x)) = \bigcap_{i=1}^{n-1} H_i^{-1}(H_i(x))$ contenant x . Dans le cas où le champ de vecteurs ne s'annule pas, celui-ci admet toujours localement $n - 1$ intégrales premières linéairement indépendantes.

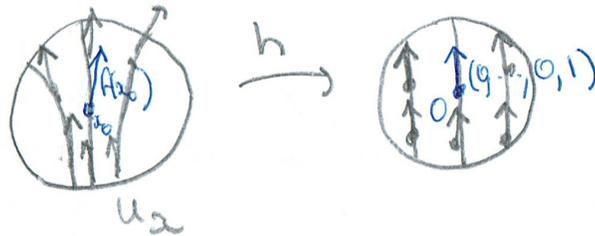
Théorème 4.1 (De redressement d'un champ de vecteurs non nul en un point). Soit $x \in U$ tel que $f(x) \neq 0$. Alors il existe un C^1 -difféomorphisme h d'un voisinage ouvert U_x de x dans un voisinage ouvert V de 0 tel que pour tout $x' \in U_x$

$$dh_{x'}(f(x')) = (0, \dots, 0, 1).$$

En particulier, pour tout $x' \in U_x$ et t suffisamment petit on a

$$\gamma_{x'}(t) = h^{-1}(h_1(x'), \dots, h_{n-1}(x'), h_n(x') + t)$$

et donc les fonctions h_1, \dots, h_{n-1} forment $n - 1$ intégrales premières indépendantes de f .



Démonstration. Soit H un hyperplan vectoriel tel que

$$\mathbb{R}^n = H \oplus \mathbb{R}f(x).$$

La formule suivante

$$\psi(u, t) = \varphi_t(x + u)$$

définit une application ψ sur un voisinage de $(0, 0)$ dans $H \times \mathbb{R}$, à valeurs dans un voisinage de x dans \mathbb{R}^n . Le flot étant de classe C^1 , la fonction ψ l'est aussi. On a par construction

$$d\psi_{(0,0)}(0_H, 1) = \left. \frac{d}{dt} \right|_0 \psi(0, t) = \left. \frac{d}{dt} \right|_0 \varphi_t(x) = f(x).$$

Par ailleurs, pour tout $u \in H$, on a

$$\psi(u, 0) = \varphi(x + u, 0) = x + u = \psi(0, 0) + u$$

et donc

$$d\psi_{(0,0)}(u, 0) = u.$$

Comme $\mathbb{R}^n = H \oplus \mathbb{R}f(x)$, il en résulte que $d\psi_{(0,0)} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est surjective, et donc bijective. On applique alors le théorème d'inversion locale, et ainsi ψ est un difféomorphisme d'un voisinage ouvert de $(0, 0)$ sur un voisinage ouvert de x . Puis, on vérifie que pour tout $(u, t) = \psi^{-1}(x')$ dans ce voisinage

$$d\psi_{(u,t)}(0_H, 1) = \left. \frac{d}{ds} \right|_{s=0} \psi(u, t + s) = \left. \frac{d}{ds} \right|_{s=0} \varphi(x + u, t + s) = f(\varphi(x + u, t)).$$

Enfin, en posant $h = \psi^{-1}$, on obtient que

$$\begin{aligned} (0, \dots, 0, 1) &= (d\psi_{(u,t)})^{-1} f(\varphi(x + u, t)) \\ &= dh_{h^{-1}(u,t)} f(h^{-1}(u, t)) \\ &= dh_{x'} f(x'). \end{aligned}$$

□

4.2. Linéarisation d'un champ au voisinage d'un zéro. Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ un champ de vecteurs de classe C^1 .

Proposition 4.2. *Soit x_0 un zéro du champ de vecteurs f . Pour tout intervalle compact $[a, b] \subset J(x_0)$ contenant 0, le flot admet le développement suivant*

$$\varphi_t(x_0 + h) = x_0 + e^{tdf_{x_0}} \cdot h + o(\|h\|)$$

pour tout $t \in [a, b]$ et h suffisamment petit.

Démonstration. Rappelons que le flot φ_t de f est de classe C^1 (cf DM1). Ainsi

$$\varphi_t(x + h) = \varphi_t(x) + d(\varphi_t)_x(h) + o_t(\|h\|)$$

en prenant bien garde que le petit o dépend a priori de t . Si t est restreint à un intervalle compact, le petit o peut être uniformément contrôlé. Pour cela, observons tout d'abord que si $F : \Omega \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$ est une application de classe C^1 , alors pour tout $x \in \Omega$ et $h \in \mathbb{R}^p$ tel que $x + h \in \Omega$ on a

$$F(x + h) = F(x) + dF_x(h) + \int_0^1 [dF_{x+sh} - dF_x](h) ds.$$

Pour cela, il suffit de poser $g(t) = F(x + th)$ et observer que $g(1) = g(0) + \int_0^1 g'(s) ds$. Si maintenant $t \in [a, b]$, alors on obtient que

$$\varphi_t(x + h) - \varphi_t(x) - d(\varphi_t)_x(h) = \int_0^1 (d(\varphi_t)_{x+sh} - d(\varphi_t)_x)(h) ds.$$

Ensuite on a

$$\frac{\| \int_0^1 (d(\varphi_t)_{x+sh} - d(\varphi_t)_x)(h) ds \|}{\|h\|} \leq \int_0^1 \|d(\varphi_t)_{x+sh} - d(\varphi_t)_x\| ds$$

et donc comme le flot est de classe C^1 , l'application $(u, t) \mapsto d(\varphi_t)_{x+u}$ est C^0 et donc uniformément continue sur un voisinage compact de la forme $B(0, r) \times [a, b]$. En particulier, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour tout $t \in [a, b]$ et $s \in [0, 1]$

$$\|d(\varphi_t)_{x+sh} - d(\varphi_t)_x\| < \varepsilon$$

dès que $\|h\| < \eta$. Ainsi

$$\int_0^1 \|d(\varphi_t)_{x+sh} - d(\varphi_t)_x\| ds \rightarrow 0$$

uniformément pour tout $t \in [a, b]$ lorsque $h \rightarrow 0$ et on peut écrire

$$\int_0^1 \|d(\varphi_t)_{x+sh} - d(\varphi_t)_x\| ds := \varepsilon(h).$$

Ainsi

$$\| \int_0^1 (d(\varphi_t)_{x+sh} - d(\varphi_t)_x)(h) ds \| \leq \|h\| \cdot \varepsilon(h),$$

et pour tout $t \in [a, b] \subset J(x)$, nous avons bien

$$\varphi_t(x+h) = \varphi_t(x) + d(\varphi_t)_x(h) + o(\|h\|).$$

Soit maintenant x_0 un zéro du champ de vecteurs f . Alors x_0 est un point fixe du flot : pour tout $t \in \mathbb{R}$ on a $\varphi_t(x_0) = x_0$ et donc au voisinage de ce zéro on a

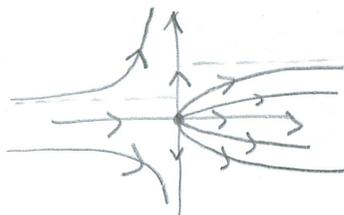
$$\varphi_t(x_0+h) = x_0 + d(\varphi_t)_{x_0}(h) + o(\|h\|).$$

De plus, sa différentielle vérifie l'équation différentielle linéarisée de f suivante (cf DM1)

$$(d(\varphi_t)_{x_0}(h))'(t) = df_{\varphi_t(x_0)}(d(\varphi_t)_{x_0}(h)) = df_{x_0}(d(\varphi_t)_{x_0}(h)).$$

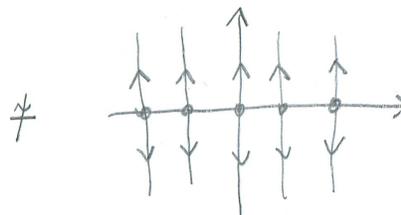
En particulier, $d(\varphi_t)_{x_0}(h) = e^{t df_{x_0}} \cdot h$ d'où le résultat. \square

Ainsi le portrait de phase de f sera une perturbation au voisinage du zéro x_0 du portrait de phase du champ de vecteurs linéaire F associé à f et donné par $F(h) = df_{x_0}(h)$. Il faut néanmoins faire attention car les perturbations de petit ordre peuvent néanmoins grandement modifier l'allure générale du portrait de phase. Par exemple, le champ de vecteurs $f(x, y) = (x^2, y)$ n'est aucunement topologiquement conjugué au champ de vecteurs linéarisé en l'origine $F(x, y) = (0, y)$.



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

$$\gamma_{(x_0, y_0)}^f(t) = \left(\frac{x_0}{1-x_0 t}, y_0 e^t \right)$$



$$F(x, y) = (0, y) = df_{(x_0, y_0)}$$

$$\gamma_{(x_0, y_0)}^F(t) = (x_0, y_0 e^t)$$

Dans certains cas, ce principe est néanmoins vrai. Citons, sans le prouver, le célèbre résultat suivant.

Théorème 4.2 (Hartman-Grobman). *Soit x_0 un zéro du champs de vecteurs f tel que les valeurs propres complexes de df_{x_0} soient toutes de partie réelle non nulle. Alors il existe deux ouverts U de x_0 et V de 0 dans \mathbb{R}^n ainsi qu'un homéomorphisme local $h : U \rightarrow V$ tel que $h(x_0) = 0$ et vérifiant*

$$h(\varphi_t(x)) = e^{tdf_{x_0}} \cdot h(x).$$

On dit que les flots de f et de df_{x_0} sont topologiquement conjugués.

4.3. Divergence d'un champ de vecteurs : théorème de Liouville et de récurrence de Poincaré. Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ un champ de vecteurs de classe C^1 .

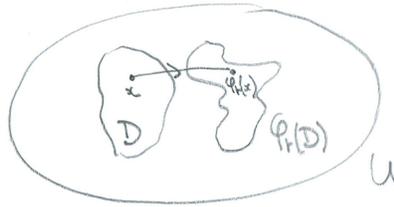
Definition 4.3. *La divergence d'un champ de vecteurs f est définie comme la fonction*

$$\operatorname{div} f(x) = \operatorname{trace}(df_x).$$

Nous allons montrer le résultat suivant :

Théorème 4.3 (de Liouville). *La condition $\operatorname{div} f = 0$ sur U équivaut à ce que le flot φ_t préserve le volume : pour tout domaine mesurable $D \subset U$ tel que $t \in J(x)$ pour tout $x \in D$, nous avons*

$$\operatorname{vol}(\varphi_t(D)) = \operatorname{vol}(D).$$



Démonstration. Si on note λ la mesure de Lebesgue sur \mathbb{R}^n , nous avons

$$\operatorname{vol}(\varphi_t(D)) = \int_{\varphi_t(D)} d\lambda = \int_U |\det d(\varphi_t)| d\lambda = \int_D \det d(\varphi_t) d\lambda$$

puisque $\det d(\varphi_t)_x > 0$ pour tout $(t, x) \in \Omega$ comme $d(\varphi_0)_x = id$ (le flot préserve l'orientation). Ainsi pour tout domaine mesurable $D \subset U$ et $t \in \cap_{x \in D} J(x)$, nous avons

$$\operatorname{vol}(\varphi_t(D)) = \operatorname{vol}(D)$$

ssi $\det d(\varphi_t)_x = 1$. Comme rappelé précédemment, $d(\varphi_t)_x(h)$ est solution de l'équation linéarisée

$$u' = A(t)(u)$$

de condition initiale h avec $A(t) = df_{\varphi_t(x)}$. Le résultat découle alors du lemme suivant :

Lemma 4.1. *Soit $A : J \rightarrow \mathcal{L}(\mathbb{R}^n)$ une application continue d'un ouvert J de \mathbb{R} dans l'ensemble des endomorphismes de \mathbb{R}^n . On note $u(h, t)$ la solution de l'équation*

$$u' = A(t)u$$

et de condition initiale $h \in \mathbb{R}^n$. Celle-ci est définie pour tout $t \in J$ et on note $u(t) : h \mapsto u(h, t)$ l'élément de $\mathcal{L}(\mathbb{R}^n)$ ainsi défini. Alors $t \in J \mapsto \det u(t)$ est solution de l'équation

$$v' = \operatorname{trace}(A(t))v$$

de condition initiale $v(0) = 1$. En particulier,

$$\det u(t) = \exp \left(\int_0^t \text{trace } A(s) ds \right)$$

et donc l'endomorphisme $u(t)$ est inversible.

Preuve du Lemme. On calcule

$$\begin{aligned} (\det u(t))' &= (\det(u_1(t), \dots, u_n(t)))' \\ &= \sum_{i=1}^n \det(u_1(t), \dots, A(t)u_i(t), \dots, u_n(t)) \\ &= \text{trace } A(t) \cdot \det(u_1(t), \dots, u_n(t)) \\ &= \text{trace } A(t) \cdot \det u(t) \end{aligned}$$

d'où la formule annoncée :

$$\det u(t) = \exp \left(\int_0^t \text{trace } A(s) ds \right)$$

La formule entre trace et déterminant se démontre facilement : la forme n -linéaire alternée

$$(v_1, \dots, v_n) \mapsto \sum_i \det_B(v_1, \dots, Av_i, \dots, v_n)$$

est proportionnel à la forme

$$(v_1, \dots, v_n) \mapsto \det_B(v_1, \dots, v_n)$$

et ce coefficient de proportionnalité se trouve en prenant $(v_1, \dots, v_n) = (e_1, \dots, e_n)$ les vecteurs composant la base B . \square

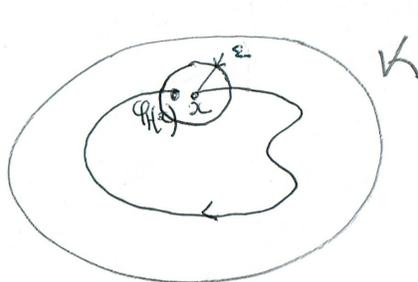
Ainsi, en appliquant le lemme ci-dessus,

$$\det(d(\varphi_t)_x) = \exp \left(\int_0^t \text{trace}(df_{\varphi_s}(x)) ds \right) = \exp \left(\int_0^t \text{div } f(\varphi_s(x)) ds \right),$$

formule dont nous déduisons l'équivalence annoncée. \square

Nous allons maintenant énoncer le théorème de récurrence de Poincaré. Pour cela, nous avons besoin de la définition suivante.

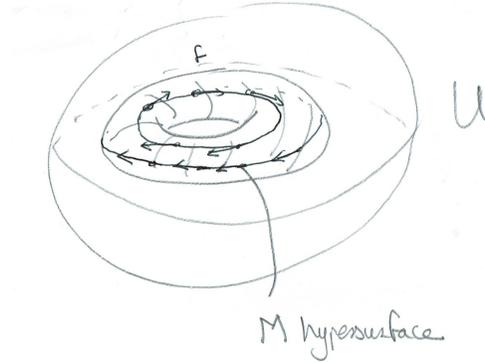
Definition 4.4. Un point x est dit récurrent, si pour tout $\varepsilon > 0$ et $t_0 \in \mathbb{R}$, il existe $t > t_0$ tel que $\varphi_t(x) \in B(x, \varepsilon)$.



La seconde notion dont nous avons besoin est la suivante.

Definition 4.5. Un compact $K \subset U$ est invariant par le flot de f si pour tout $x \in K$, $\varphi_t(x) \in K$ pour tout $t \in J(x)$.

Si un compact K est invariant par le flot, le théorème des bouts implique que $J(x) = \mathbb{R}$ pour tout $x \in K$. Un exemple de tel compact est donné par la fermeture de l'intérieur d'une hypersurface fermée sans bord pour laquelle le champ de vecteurs serait tangent le long de cette même hypersurface.



Bien qu'il soit peu probable d'avoir un champ de vecteurs dont toutes les orbites sont périodiques, si le flot préserve les volumes à l'intérieur d'un compact invariant, alors presque toute orbite sera récurrente :

Théorème 4.4 (de récurrence de Poincaré). *Soit f un champ de vecteurs à divergence nulle sur un compact K invariant par son flot. Alors presque (respectivement à la mesure de Lebesgue) tout point de K est récurrent.*

Démonstration. L'application

$$T := \varphi_{t_0} : K \rightarrow K$$

est un difféomorphisme, et donc mesurable puisque continue. Par le théorème de Liouville, nous savons aussi que T préserve la mesure de Lebesgue de \mathbb{R}^n : pour tout domaine borélien (mesurable pour la mesure de Lebesgue) $D \subset K$ on a $\text{vol}(T(D)) = \text{vol}(\varphi_{t_0}(D)) = \text{vol}(D)$.

On choisit ensuite une base dénombrable $\{U_n\}_{n \in \mathbb{N}}$ d'ouverts de K (considérer les traces sur K des boules ouvertes centrées en des points à coordonnées rationnelles et de rayon de la forme $1/k$ où $k \in \mathbb{N}^*$). Tout ouvert de K pour la topologie induite s'écrit comme une réunion d'ouverts de la collection $\{U_n\}$. On note $N(k)$ l'ensemble des n pour lesquels l'ouvert U_n est la trace d'une boule dont le rayon est $< 1/k$. On a alors

$$K = \bigcup_{n \in N(k)} U_n$$

quelque soit $k \in \mathbb{N}$.

Notons \tilde{U}_n l'ensemble des points x de U_n repassant une infinité de fois par U_n : $x \in \tilde{U}_n \Leftrightarrow T^k(x) \in U_n$ pour une infinité de $k \in \mathbb{N}$.

Lemma 4.2. *Pour tout $n \in \mathbb{N}$, l'ensemble \tilde{U}_n est mesurable et*

$$\text{vol}(U_n \setminus \tilde{U}_n) = 0.$$

Preuve du Lemme. Soit

$$C_m = U_n \setminus (\bigcup_{j \geq m} T^{-j}(U_n))$$

l'ensemble des points de U_n dont les images à partir de la m -ème itérée n'appartiennent plus à U_n . Alors C_m est mesurable et donc

$$\tilde{U}_n = U_n \setminus (\bigcup_m C_m)$$

est également mesurable.

On observe alors que $T^{-m}(C_m) \cap C_m = \emptyset$.

En effet,

$$\begin{aligned} T^{-m}(C_m) \cap C_m &= T^{-m}(C_m \cap T^m(C_m)) \\ &\subset T^{-m}(U_n \cap T^m(C_m)) = T^{-m}(\emptyset) = \emptyset. \end{aligned}$$

De même, $T^{-2m}(C_m)$ est disjoint de C_m , mais aussi de $T^{-m}(C_m)$ car préimage par T^m de deux parties disjointes. Ainsi $\{T^{-km}(C_m)\}_{k \in \mathbb{N}}$ est une famille de sous-ensembles mesurables contenus dans K , de même mesure et deux à deux disjoints. Comme la réunion a une mesure au plus égale à celle de K , on en déduit que $\text{vol}(C_m) = 0$ et donc $\text{vol}(U_n \setminus \tilde{U}_n) = 0$. \square

On s'intéresse à l'ensemble

$$\Omega = \bigcap_{k \in \mathbb{N}} \bigcup_{n \in N(k)} \tilde{U}_n$$

qui est exactement l'ensemble des points récurrents. On a alors

$$\begin{aligned} \text{vol}(K \setminus \Omega) &= \text{vol}(K \setminus \bigcap_{k \in \mathbb{N}} \bigcup_{n \in N(k)} \tilde{U}_n) \\ &= \lim_{k \rightarrow \infty} \text{vol}(K \setminus \bigcup_{n \in N(k)} \tilde{U}_n) \\ &= \lim_{k \rightarrow \infty} \text{vol}(\bigcup_{n \in N(k)} U_n \setminus \bigcup_{n \in N(k)} \tilde{U}_n) \\ &\leq \lim_{k \rightarrow \infty} \text{vol}(\bigcup_{n \in N(k)} (U_n \setminus \tilde{U}_n)) \\ &= 0. \end{aligned}$$

\square

4.4. Ensemble ω -limite et théorème de Poincaré-Bendixon. On se donne un champ de vecteurs $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ complet et de classe C^1 et on note φ_t son flot.

Definition 4.6. Soit $x \in U$. On appelle ensemble ω -limite de x l'ensemble

$$\omega(x) := \bigcap_{t > 0} \overline{\{\varphi_s(x) \mid s \geq t\}}$$

des valeurs d'adhérence de la trajectoire $\varphi_t(x)$ lorsque $t \rightarrow \infty$. De même, on appelle ensemble α -limite de x l'ensemble

$$\alpha(x) := \bigcap_{t < 0} \overline{\{\varphi_s(x) \mid s \leq t\}}.$$

Autrement dit,

$$\omega(x) = \{y \in \mathbb{R}^n \mid \exists \{t_n\} \rightarrow \infty \text{ avec } y = \lim_{n \rightarrow \infty} \varphi_{t_n}(x)\}.$$

Par exemple, si l'orbite de x est périodique, alors

$$\omega(x) = \alpha(x) = \mathcal{O}_x.$$

Autre exemple, dans le cas d'un noeud attractif en dimension 2, on a $\alpha(x) = \emptyset$ et $\omega(x) = \{(0, 0)\}$ (et inversément pour un noeud répulsif).

Un autre exemple important est celui d'une courbe fermée le long de laquelle la courbe intégrale issue de x vient s'enrouler. Dans ce cas, l'ensemble ω -limite est la courbe en question, voir Figure 16.

Par ailleurs, si le champ est à divergence nulle et s'il existe un compact K invariant par le flot, d'après le théorème de récurrence de Poincaré, on a $x \in \omega(x)$ et $x \in \alpha(x)$ pour presque tout $x \in K$.

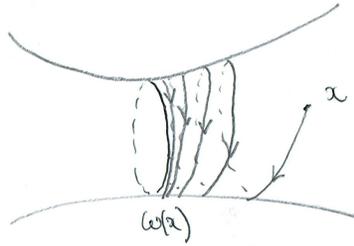


FIGURE 16. L'ensemble limite est une courbe fermée.

Proposition 4.3 (Propriétés dynamiques). *Pour tout x in U ,*

(i) $\omega(x) = \omega(\varphi_t(x))$ pour tout $t \in \mathbb{R}$;

(ii) $\omega(x)$ est invariant par le flot : pour tout $y \in \omega(x)$, $\varphi_t(y) \in \omega(x)$ pour tout $t \in \mathbb{R}$. En particulier, $\gamma_y \subset \omega(x)$.

(iii) Si H est une intégrale première de f , alors $\omega(x) \subset H^{-1}(H(x))$.

Démonstration. La propriété (i) se démontre ainsi : si $y \in \omega(x)$, alors il existe une suite $t_n \rightarrow \infty$ telle que $\varphi_{t_n}(x) \rightarrow y$ lorsque $n \rightarrow \infty$. Alors $y = \lim_{n \rightarrow \infty} \varphi_{t_n - t}(\varphi_t(x)) \in \omega(\varphi_t(x))$. Ainsi $\omega(x) \subset \omega(\varphi_t(x))$ pour tout $x \in U$ et pour tout $t \in \mathbb{R}$, et on déduit réciproquement que

$$\omega(\varphi_t(x)) \subset \omega(\varphi_{-t}(\varphi_t(x))) = \omega(x).$$

De même, si $y = \lim_{n \rightarrow \infty} \varphi_{t_n}(x) \in \omega(x)$, alors pour tout $t \in \mathbb{R}$ on a $\varphi_t(y) = \lim_{n \rightarrow \infty} \varphi_{t+t_n}(x) \in \omega(x)$. D'où l'invariance par le flot (ii).

Enfin, pour tout $y \in \omega(x)$, on a $H(y) = H(\lim_{n \rightarrow \infty} \varphi_{t_n}(x)) = \lim_{n \rightarrow \infty} H(\varphi_{t_n}(x)) = H(x)$ par continuité et invariance par le flot de H d'où (iii). \square

Proposition 4.4 (Propriétés topologiques). *Soit $x \in U$. Alors*

(iv) $\omega(x)$ est un fermé de U ;

Si de plus, on suppose qu'il existe un compact K de U tel que $\gamma_x(t) \in K$ pour $t \geq 0$, alors

(v) $\omega(x)$ est un compact connexe non vide.

Démonstration. La fermeture se démontre ainsi : si $y \notin \omega(x)$, il existe $\varepsilon > 0$ et $t_0 > 0$ tels que

$$\|\varphi_t(x) - y\| \geq \varepsilon$$

pour tout $t \geq t_0$. Ainsi pour tout $z \in B(y, \varepsilon/2)$ on a

$$\|\varphi_t(x) - z\| \geq \varepsilon/2$$

pour tout $t \geq t_0$ et donc $z \notin \omega(x)$. Donc $B(y, \varepsilon/2) \cap \omega(x) = \emptyset$ et donc le complémentaire de $\omega(x)$ est ouvert ce qui vérifie (iv).

Démontrons (v). Comme $\omega(x)$ est fermé contenu dans un compact, c'est un compact. Le fait que $\omega(x)$ est non vide découle de ce que par exemple la suite $\varphi_n(x)$ admet au moins une valeur d'adhérence.

Montrons que $\omega(x)$ est connexe. Supposons pour cela qu'il existe U_1, U_2 deux ouverts disjoints tels que $\omega(x) \subset U_1 \cup U_2$ et $\omega(x) \cap U_i \neq \emptyset$. Donc il existe une suite strictement croissante $\{t_n\}$ tendant vers l'infini telle que

$$\varphi_{t_{2n+1}}(x) \in U_1 \text{ et } \varphi_{t_{2n}}(x) \in U_2.$$

Donc pour tout n il existe $\tau_n \in]t_{2n}, t_{2n+1}[$ tel que $\varphi_{\tau_n}(x) \notin U_1 \cup U_2$. Mais cette suite $\{\varphi_{\tau_n}(x)\} \subset K$ admet une valeur d'adhérence qui est dans $\omega(x) \setminus U_1 \cup U_2$ par construction : contradiction. \square

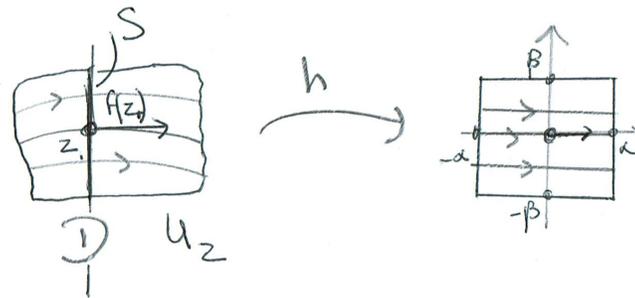
Théorème 4.5 (Poincaré-Bendixon). *On considère un champ de vecteurs sur un ouvert $U \subset \mathbb{R}^2$ de classe C^1 . On suppose qu'il existe un point $z \in U$ et un compact $K \subset U$ tels que :*

- $\gamma_z(t) \in K$ pour $t \geq 0$;
- K ne contient pas de point fixe du flot.

Alors l'ensemble limite $\omega(z)$ consiste en une orbite périodique appelée cycle limite.

Démonstration. Tout d'abord, remarquons que l'ensemble limite $\omega(z)$ un compact connexe non-vide qui ne contient pas de point fixe.

Choisissons $z_0 \in \omega(z)$. Nous allons montrer tout d'abord que l'orbite de z_0 , qui est contenue dans $\omega(z)$ est périodique. En effet, on sait que $\varphi_t(z_0) \in \omega(z)$ pour tout $t \geq 0$ et que $\omega(z)$ est compact. Ceci implique que $\omega(z_0) \neq \emptyset$. On choisit alors un point $z_1 \in \omega(z_0)$. Comme $\omega(z_0) \subset \omega(z)$, on sait que $f(z_1) \neq 0$, et donc on peut redresser le flot : on choisit une droite D du plan transverse à $f(z_1)$ passant par z_1 et un difféomorphisme $h : U_{z_1} \rightarrow]-\alpha, \alpha[\times]-\beta, \beta[$ tel que les courbes intégrales de f s'envoient sur les courbes de la forme $(x+t, y)$. On peut également supposer que pour tout $z' \in S := D \cap U_{z_1}$ nous avons $h(z') \in \{0\} \times \mathbb{R}$.



Nous avons besoin du lemme suivant :

Lemme 4.3. *Sur la section transverse S , il ne peut y avoir deux éléments z'_1 et z'_2 distincts d'un même ensemble limite ω : si z'_1 et z'_2 sont des points de $S \cap \omega(z')$ pour un certain $z' \in U$, alors $z'_1 = z'_2$.*

Preuve du lemme. Si pour $i = 1, \dots, n$ on a $z''_i = \varphi_{t_i}(z) \in S$ avec $t_1 \leq \dots \leq t_n$, alors la suite (z''_i) est monotone le long de S dans le sens suivant : les points sont définis comme $h(z''_i) = (0, y_i)$ vérifie $y_1 \leq \dots \leq y_n$ ou $y_n \leq \dots \leq y_1$. On visualise cela à l'aide du théorème de Jordan sur la figure 17.

Maintenant, si on avait deux éléments z'_1 et z'_2 distincts sur S appartenant à un même ensemble limite, cela contredirait nécessairement la monotonie. \square

Ainsi, comme $z_1 = \lim_{n \rightarrow \infty} \varphi_{t_n}(z_0)$, nécessairement la courbe γ_{z_0} coupe transversalement une infinité de fois la section S , et les points d'intersection étant tous des éléments de $\omega(z)$, ils sont confondus par le lemme précédent. En particulier, l'orbite de z_0 est périodique.

Montrons maintenant que $\omega(z)$ se réduit à \mathcal{O}_{z_0} . En effet, dans le cas contraire, l'ensemble $\omega(z) \setminus \mathcal{O}_{z_0}$ est alors un ouvert non vide de $\omega(z)$. La connexité de $\omega(z)$ implique alors que \mathcal{O}_{z_0} n'est pas ouvert dans $\omega(z)$ et

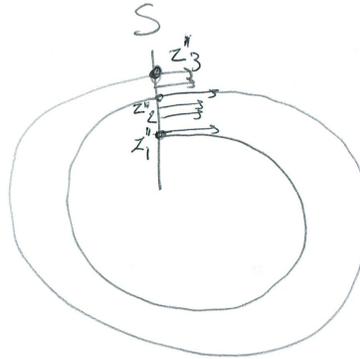


FIGURE 17. Monotonie le long de la section transverse.

donc il existe $z_2 \in \mathcal{O}_{z_0}$ tel que $z_2 = \lim_{n \rightarrow \infty} z'_n$ avec $z'_n \in \omega(z) \setminus \mathcal{O}_{z_0}$. Comme $f(z_2) \neq 0$, on peut redresser le champ en z_2 et obtenir une section locale S transverse au flot. On voit alors que S doit contenir un point de $\omega(z) \setminus \mathcal{O}_{z_0}$ comme cet ensemble est invariant par le flot (voir figure 18). Contradiction. \square

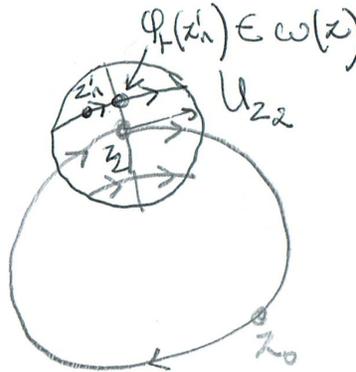


FIGURE 18

4.5. Stabilité et fonctions de Liapounov. Considérons un champ de vecteurs $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ de classe C^1 , et $x_0 \in U$ un zéro de f . Rappelons que x_0 est alors un point fixe du flot : $\varphi_t(x_0) = x_0$ pour tout $t \in \mathbb{R}$.

Definition 4.7. On dit que x_0 est un point fixe stable si, pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que

$$\|x - x_0\| \leq \delta \Rightarrow \forall t \in J(x) \cap \mathbb{R}_+, \|\varphi_t(x) - x_0\| \leq \varepsilon.$$

Par le théorème des bouts, nous déduisons alors que $J(x) \cap \mathbb{R}_+ = \mathbb{R}_+$ pour tout $x \in B(x_0, \delta)$. Par exemple, pour un noeud attractif, l'origine est un point fixe stable ; pour un foyer attractif aussi, mais pas pour un noeud ou un foyer répulsif. Pour un foyer de centre l'origine, celle-ci est également stable.

Definition 4.8. On dit que x_0 est un point d'équilibre asymptotiquement stable s'il est stable et s'il existe $\delta > 0$ tel que

$$\|x - x_0\| \leq \delta \Rightarrow \lim_{t \rightarrow \infty} \varphi_t(x) = x_0.$$

Par exemple, pour un noeud ou foyer attractif, l'origine est un point fixe asymptotiquement stable ; en revanche, pour un foyer de centre l'origine, celle-ci n'est pas asymptotiquement stable.

Définition 4.9. Soit $L : W \rightarrow \mathbb{R}$ une fonction C^0 définie sur un voisinage W de x_0 . On dit que L est une fonction de Liapounov pour le champ f au voisinage de x_0 si

(i) x_0 est un minimum strict pour L sur W ;

(ii) Si pour tout $x \in W \setminus \{x_0\}$ la fonction $t \mapsto L(\varphi_t(x))$ est strictement décroissante sur son intervalle de définition.

Observer que la précaution du (ii) vient de ce que a priori il n'y a pas de raison pour laquelle $\varphi_t(W) \subset W$ pour $t \geq 0$. Un critère de stabilité asymptotique est donné par le résultat suivant.

Théorème 4.6 (Théorème de Liapounov). Si un point d'équilibre x_0 du champ de vecteurs f admet une fonction de Liapounov, alors x_0 est asymptotiquement stable.

Démonstration. Quitte à restreindre L , on peut choisir W compact. On peut également supposer que $L(x_0) = 0$. Soit alors $\varepsilon > 0$ tel que $B(x_0, \varepsilon) \subset W$. Comme $W \setminus B(x_0, \varepsilon)$ est fermé, donc compact, $\min_{x \in W \setminus B(x_0, \varepsilon)} L(x) > \frac{1}{n_0}$ pour n_0 assez grand. Ainsi, l'ensemble $V_{n_0} = L^{-1}([0, \frac{1}{n_0}])$ est contenu dans $B(x_0, \varepsilon)$ et forme un voisinage compact de x_0 : pour cela, il faut remarquer que d'après (i), x_0 est un minimum strict et donc $L^{-1}(\frac{1}{2n_0}, \frac{1}{n_0})$ est un voisinage ouvert de x_0 contenu dans V_{n_0} .

Il est aisé de vérifier que $\varphi_t(V_{n_0}) \subset V_{n_0}$ puisque $t \mapsto L(\varphi_t(x))$ est décroissante pour $t \geq 0$. Ainsi, x_0 est un point d'équilibre stable : on peut trouver $\delta > 0$ tel que $B(x_0, \delta) \subset V_{n_0}$ (comme V_{n_0} est un voisinage ouvert de x_0), et ainsi pour tout $x \in B(x_0, \delta)$ on a bien $\varphi_t(x) \in V_{n_0} \subset B(x_0, \varepsilon)$.

Maintenant soit $x \in B(x_0, \delta) \subset V_{n_0}$ et notons $a = \lim_{t \rightarrow \infty} L(\varphi_t(x))$ qui existe bien puisque l'on a une fonction strictement décroissante et minorée par 0. Il faut montrer que $a = 0$, ce qui nous permettra de conclure comme x_0 est l'unique minimum de L sur V_{n_0} avec $L(x_0) = 0$. Si $a > 0$, on considère une suite $\{\varphi_{t_n}(x)\}_n$ avec la condition que $(t_n)_n$ soit une suite strictement croissante tendant vers l'infini. Quitte à en extraire une sous-suite convergente (cette suite étant contenue dans V_{n_0} compact), on peut supposer que cette suite converge vers un point de V_{n_0} que l'on notera x_1 . Par continuité, on a $L(x_1) = \lim_{n \rightarrow \infty} L(\varphi_{t_n}(x)) = a > 0$, et donc en particulier $x_1 \neq x_0$. Mais alors, pour tout $t > 0$, on a

$$a = L(x_1) > L(\varphi_t(x_1)) = \lim_{n \rightarrow \infty} L(\varphi_{t+t_n}(x)) = a$$

d'où une contradiction. □

F. BALACHEFF, LABORATOIRE PAUL PAINLEVÉ, BAT. M2, UNIVERSITÉ DES SCIENCES ET TECHNOLOGIES, 59 655 VILLENEUVE D'ASCQ, FRANCE.

E-mail address: florent.balacheff@math.univ-lille1.fr