

Introduction aux systèmes dynamiques

François LABOURIE ¹

26 avril 2007

¹Univ. Paris-Sud, Laboratoire de Mathématiques, Orsay F-91405 Cedex ; CNRS, Orsay cedex, F-91405

Table des matières

I	Équations différentielles ordinaires	7
1	Premières notions	9
1.1	Le problème à résoudre	9
1.1.1	Équations d'ordre supérieur	10
1.1.2	Équations autonomes	11
1.2	Durée de vie des solutions	11
1.3	Portrait de phase	12
1.4	Variation des fonctions	13
2	Équations linéaires	17
2.1	Rappels d'algèbre linéaire	17
2.2	Équations à coefficients constants	19
2.3	Équations linéaires dépendant du temps	22
2.4	Équations affines	27
2.5	Théorème de Liouville	29
3	Cauchy-Lipschitz	31
3.1	Hypothèses et énoncé	31
3.1.1	Fonctions k -lipschitziennes	31
3.1.2	Un premier énoncé.	32
3.1.3	Démonstration du théorème	33
3.2	Solutions maximales	36
3.2.1	Unicité, solutions maximales et prolongement.	36
3.2.2	Prolongement, sortie des compacts et explosion en temps fini	37
3.3	Retour sur Cauchy-Lipschitz	38
4	Le lemme de Gronwall	39
4.1	Énoncés	39
4.1.1	Gronwall avec des inégalités strictes	39
4.1.2	Gronwall classique	41
4.1.3	Sur et sous-solutions	41
4.2	Contrôle de la durée de vie	42

4.2.1	Utilisation de Gronwall pour montrer qu'une durée de vie est infinie	42
4.2.2	Utilisation de Gronwall pour montrer qu'une durée de vie est finie	43
4.3	Divergence des solutions	44
5	Le flot	49
5.1	Dépendance continue	49
5.2	Dépendance par rapport à un paramètre	51
5.3	Gronwall large	52
5.4	Flot comme solution	53
5.5	Dépendance C^k	54
6	Etude géométrique	57
6.1	Equations autonomes en dimension 2	57
6.1.1	Portrait de phase	57
6.1.2	Les trajectoires ne se coupent pas	58
6.1.3	Points fixes	58
6.2	Barrières et entonnoirs	59
6.2.1	Barrières	59
6.2.2	Entonnoirs et entonnoirs étranglés	60
6.2.3	Anti-entonnoirs et anti-entonnoirs étranglés	61
6.3	Equations autonomes	64
6.3.1	Rappels sur les courbes planes	64
6.3.2	Barrières et régionnement	65
6.3.3	Courbes de niveaux et orbites	67
7	Stabilité et équilibre	71
7.1	Champ de vecteur	71
7.2	Points d'équilibre	72
7.3	Fonction de Liapounov	73
7.4	Stabilité et valeurs propres	76
7.5	Champ de gradient	77
8	Linéarisation	81
8.1	La différentielle du flot	81
8.1.1	Équation linéarisée	81
8.1.2	La différentielle	81
9	Difféomorphisme	83
9.1	Changement de coordonnées	83
9.2	Image par un difféomorphisme	84
9.3	Hors de l'équilibre	86

10 Orbites périodiques	89
10.1 L'application premier retour de Poincaré	89
10.1.1 Dérivée de l'application premier retour	90
10.2 Orbite périodique stable	91
11 Cols	95
11.1 Instabilité et spectre	95
11.2 Séparatrices pour les cols	97
11.2.1 Col ou point selle	98
11.2.2 Séparatrice	98
11.3 Existence de quatre séparatrices	98
11.3.1 Une équation différentielle associée	99
11.3.2 Retour aux orbites du champ de vecteur et preuve du lemme 11.3.0.2	100
11.3.3 Unicité des séparatrices	101
11.3.4 Position des tangentes	101
II Introduction à la Mécanique Classique	103
12 Fonctions invariantes	105
12.1 Définitions	105
13 Mécanique Lagrangienne	107
13.1 Lagrangien	107
13.1.1 Equations d'Euler-Lagrange	107
13.1.2 Points critiques de l'action	107
13.1.3 Equation de Newton	109
13.1.4 Sous-variétés	109
13.1.5 Exemple et exercice : surface de révolution	109
13.2 Symétries	110
13.2.1 Exemple et exercice, mouvement dans un champ central .	111
13.2.2 Exemple et exercice, mouvement dans un champ central (bis)	112
13.2.3 Problème à deux corps	113
14 Mécanique Hamiltonienne	115
14.1 Equations d'Hamilton	115
14.1.1 Transformation de Legendre	115
14.2 Crochet de Poisson	117
14.2.1 Algèbre des intégrales premières	117
14.2.2 Evolution des observables	118
14.3 Coordonnées symplectiques	119
14.3.1 Variables en involution	119
14.3.2 Variables conjuguées	120
14.3.3 Coordonnées symplectiques	121

14.4	Systèmes complètement intégrables	122
14.4.1	Systèmes complètement intégrables	122
14.4.2	Méthode d'Hamilton-Jacobi	123
14.4.3	Au delà des systèmes complètement intégrables	128
14.5	Toupie de Lagrange	
	Mouvement du solide	129
III	Introduction à la théorie ergodique	133
15	Mesures Invariantes	135
15.1	Existence de mesures invariantes	137
15.2	Théorème de récurrence de Poincaré	138
15.3	Théorèmes ergodiques	141
15.3.1	Théorèmes ergodiques	141
15.4	Ergodicité du décalage de Bernouilli	144
15.4.1	Du décalage aux systèmes hyperboliques	146
15.4.2	Feuilletages absolument continus	148
16	Codage et entropie	149
16.1	Codage	149
16.1.1	Codage, partition génératrice	149
16.2	Entropie	150
16.2.1	Entropie d'une mesure sur un ensemble fini	150
16.2.2	Entropie d'une mesure sur $F^{\mathbb{N}}$	151
16.2.3	Entropie d'une transformation	151
16.3	Calcul de l'entropie	152
16.4	L'entropie comme invariant	154

Première partie

Équations différentielles
ordinaires

Chapitre 1

Premières notions

1.1 Le problème à résoudre

Soit E un espace vectoriel. Si $x : t \rightarrow x(t)$ est une courbe à valeurs dans E , on note $\dot{x}$ la dérivée de x par rapport à la variable t .

Soit maintenant Ω un ouvert de E , I un intervalle ouvert de $\mathbb{R}$, et f une application de $I \times \Omega$ à valeurs dans E . Une *équation différentielle* dans E est une expression de la forme

$$\dot{x} = f(t, x).$$

Résoudre cette équation différentielle avec les *conditions initiales* (t_0, x_0) , c'est trouver une *solution*, c'est-à-dire une courbe x à valeurs dans E définie sur un intervalle J contenant t_0 , non réduit à un point et inclus dans I , vérifiant

$$\begin{cases} x(t_0) &= x_0 \\ \forall t, \dot{x}(t) &= f(t, x(t)). \end{cases}$$

Lorsque $\dim(E) = 1$, on parle souvent d'*équations différentielles scalaires*. Lorsque $E = \mathbb{R}^n$, on parle traditionnellement de systèmes d'équations différentielles. Dans ce cas, en remplaçant la courbe x par la famille de fonction $(x_1, \dots, x_n)$ on peut écrire

$$\begin{cases} \dot{x}_1 &= f_1(t, x_1, \dots, x_n) \\ &\vdots \\ \dot{x}_n &= f_n(t, x_1, \dots, x_n) \end{cases}$$

La première origine des équations différentielle est physique : la mécanique du point où l'on étudie l'évolution ou trajectoire d'une particule, dans l'espace E , au cours du temps, sous l'effet de contraintes indiquées par la fonction f . On parlera donc aussi de *trajectoires* ou *orbites* autant que de *solutions* et la variable t sera appelée *temps* et notée traditionnellement t .

EXEMPLE :

1. L'équation *proie-prédateur*, ou de *Lotka-Volterra*. On image une situation écologique donnée par deux populations. L'une mesurée par y est celle des prédateurs, l'autre mesurée par x est celle des proies. Les prédateurs mangent les proies qui, elles, possèdent une source de nourriture inépuisable. Par ailleurs, tout ce petit monde se reproduit par les voies naturelles. Ainsi, l'évolution de la population des prédateurs est proportionnelle d'une part à quantité de nourriture supérieure à un certain seuil c'est-à-dire $x - 1$, d'autre part à la quantité de reproducteurs y

$$\dot{y} = y(x - 1).$$

De même, l'accroissement population de x est proportionnelle à x la quantité de reproducteurs, et corrigée par le nombre nombre de proies mangées par les prédateurs, quantité proportionnelle à xy .

$$\dot{x} = x - xy = x(1 - y).$$

Les constantes biologiques "d'appétit/libid" ou de "réussite à la chasse" ont toutes été ajustées pour donner la forme simple ci-dessus. Nous y reviendrons en 1.3.

1.1.1 Équations d'ordre supérieur

Nous venons de définir les équations d'ordre 1 (en dérivation). En notant $x^{(k)}$ la dérivée d'ordre k de la courbe x , une équation différentielle d'ordre k s'écrit

$$x^{(k)} = f(t, x, x^{(1)}, \dots, x^{(k-1)}). \quad (1.1)$$

Cette équation peut se ramener à l'étude de l'équation (d'ordre 1) suivante, définie elle sur $E^k = E \oplus \dots \oplus E$:

$$\left\{ \begin{array}{lcl} \dot{x}_1 & = & x_0 \\ \dot{x}_2 & = & x_1 \\ & \dots & \\ \dot{x}_{k-1} & = & x_{k-2} \\ \dot{x}_k & = & f(t, x_0, x_1, \dots, x_{k-1}) \end{array} \right. \quad (1.2)$$

En effet, nous avons immédiatement

Proposition 1.1.1.1 *la courbe x est solution de (1.1) si et seulement si la courbe $(x, x^{(1)}, \dots, x^{(k-1)})$ est solution de (1.2).*

Ceci est particulièrement utile pour les équations d'ordre 2 provenant de la physique Newtonienne.

EXERCICE :

Ramenez à l'étude d'équations d'ordre 1, les équations suivantes

1. l'*oscillateur harmonique* : $\ddot{x} = ax$.
2. le *pendule sans frottements* : $\ddot{x} = a \sin(x)$.
3. le *pendule avec frottements* $\ddot{x} = a \sin x - \dot{x}$.

1.1.2 Équations autonomes

Une équation différentielle est *autonome* si la fonction qui la définit ne dépend pas du temps. Autrement dit, une équation autonome s'écrit

$$\dot{x} = f(x).$$

Une propriété immédiate des solutions est la suivante

Proposition 1.1.2.1 *Une courbe x est solution d'une équation autonome avec condition initiale (t_0, x_0) , si et seulement si la courbe $t \rightarrow x(t - t_1 + t_0)$ est solution de la même équation avec conditions initiales (x_0, t_1) .*

Il n'y pas de perte de généralités à étudier les équations autonomes (ce que nous ferons très souvent par la suite) car toute équation différentielle sur E

$$\dot{x} = f(t, x) \tag{1.3}$$

est équivalente à l'équation autonome suivante sur $\mathbb{R} \times E$

$$\begin{cases} \dot{x} &= f(y, x) \\ \dot{y} &= 1. \end{cases} \tag{1.4}$$

En effet

Proposition 1.1.2.2 *la courbe x est solution de (1.3) si et seulement si la courbe (y, x) avec $y(t) = t$ est solution de (1.4)*

EN CONCLUSION, ON NE RESTREINT PAS LA GÉNÉRALITÉ DU PROBLÈME EN ÉTUDIANT LES ÉQUATIONS AUTONOMES : LES ÉQUATIONS D'ORDRE SUPÉRIEURE, OU DÉPENDANT DU TEMPS, D'APPARENCE PLUS GÉNÉRALES SE RAMÈNENT À L'ÉTUDE DE CE CAS PARTICULIER.

1.2 Durée de vie des solutions

L'un des premiers problèmes dans l'étude des équations différentielles est l'étude de la durée de vie des solutions : on cherche à trouver un intervalle maximal sur lequel est définie la solution. Même dans le cas d'une équation autonome

$$\dot{x} = f(x).$$

avec f définie sur E tout entier, la solution n'est pas nécessairement définie sur $\mathbb{R}$ tout entier comme le montre l'exemple suivant défini sur $\mathbb{R}$

$$\dot{x} = -x^2.$$

Elle admet comme solution

$$x(t) = -\frac{1}{t},$$

qui ne peut s'étendre à tout $\mathbb{R}$.

1.3 Portrait de phase pour les équations autonomes

Dans le cas autonome,

$$\dot{x} = f(x).$$

La fonction f est un *champ de vecteurs* : à chaque point de l'espace, on associe un vecteur. La courbe solution est alors tangente en tout point à ce champ de vecteurs (pensez à de la limaille de fer dans un champ magnétique plan). On tracera souvent en dimension 2 le *portrait de phase* qui comprendra un certain nombre d'orbites significatives permettant de rendre compte de l'évolution du système. En particulier on cherchera

- les *points fixes* ou *points d'équilibres* : les points où f s'annule. Ils correspondent aux trajectoires constantes.
- quelques courbes aboutissant aux points fixes
- les *orbites périodiques* : c'est à dire les solutions x , pour lesquelles il existe une constante T non nulle telle que $x(t + T) = x(t)$.

Quelques portraits de phase

Tout d'abord en dimension 1, on regarde l'équation

$$\dot{x} = ax,$$

suivant le signe de a . En dimension 2, on étudiera le système

$$\begin{cases} \dot{x} &= ax + b \\ \dot{y} &= cy + d \end{cases} \quad (1.5)$$

associé à la matrice

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

dans les cas suivants. On détaillera l'étude en fonction du signe des constantes

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

Puis, en utilisant les nombres complexes.

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

Enfin

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Voici enfin le portrait de phase de l'équation de Lotka-Volterra. Celui-ci est obtenu par ordinateur car nous n'avons pas les moyens théoriques pour le moment de l'étudier, nous y reviendrons plus tard.

Notions qualitatives

On ne sait que rarement résoudre les équations différentielles ordinaires. Par contre, on sait obtenir des informations qualitatives sur l'évolution d'un système. Il est ainsi important de comprendre ce qui se passe au voisinage au voisinage des points d'équilibre. Prenons l'exemple du pendule avec frottements : il a clairement deux points d'équilibre. Celui où le pendule est à la verticale en bas du point d'oscillation, celui au contraire où le pendule est à la verticale en haut de ce point d'oscillation. Intuitivement, l'un est stable, l'autre est instable. Nous justifierons plus tard ce point de vue sans résoudre explicitement les équations. Pour le moment introduisons quelques notions qualitatives.

On voit sur les portraits de phase que les points fixes - ou plutôt les orbites dans leur voisinage - peuvent avoir des comportements très différents. On peut définir quelques comportements importants de la manière *qualitative* suivante.

- un point fixe x_1 est *attracteur*, s'il possède un voisinage U , tel que toute orbite x de condition initiale (t_0, x_0) avec $x_0 \in U$ vérifie

$$\exists t_1, \forall t > t_1, x(t) \in U, \text{ et } \lim_{t \rightarrow \infty} x(t) = x_1.$$

- un point fixe x_1 est *répulsor*, s'il possède un voisinage U , tel que toute orbite x de condition initiale (x_0, t) avec $x_0 \in U$ vérifie

$$\exists t_1, \forall t < t_1, x(t) \in U, \text{ et } \lim_{t \rightarrow -\infty} x(t) = x_1.$$

- On dira qu'un ensemble U est *stable vers le futur* (ou ω -stable) si toute solution de condition initiale (t_0, x_0) , avec x_0 dans U reste dans U pour $t \geq t_0$.
- On dira qu'un ensemble U est *stable vers le passé* (ou α -stable) si toute solution de condition initiale (t_0, x_0) , avec x_0 dans U reste dans U pour $t \leq t_0$.

Comme vous le savez α et ω sont les première et dernière lettres de l'alphabet grec et par analogie la création du Monde et la fin des Temps.

En général, on se sait pas résoudre les équations différentielles. Il est important cependant d'obtenir des informations qualitatives sur les comportements des solutions. Nous verrons en particulier une méthode nous permettant d'étudier les points fixes et de montrer dans certains cas qu'ils sont attracteurs.

1.4 Variation des fonctions

On dira qu'une fonction g est *invariante* ou que c'est une *intégrale première* de l'équation différentielle

$$\dot{x} = f(x),$$

si pour toute solution x la fonction $t \rightarrow g(x(t))$ est constante. Lorsque l'on sait que g est une intégrale première, les trajectoires seront tracées sur les ensembles de niveaux de la fonction g .

Proposition 1.4.0.3 *Soit*

$$f = (f_1, \dots, f_n),$$

une application C^1 d'un ouvert Ω de $\mathbb{R}^n$ dans $\mathbb{R}^n$. La fonction g de classe C^1 est une intégrale première de l'équation différentielle

$$\dot{x} = f(x),$$

si et seulement si

$$\sum_{i=1}^{i=n} \partial_i g \cdot f_i = 0. \quad (1.6)$$

DÉMONSTRATION : Remarquons tout d'abord que pour toute fonction g de classe C^1 Ω dans $\mathbb{R}$ et toute solution $x =$, on a

$$\frac{d}{dt} g \circ x(s) = \sum_{i=1}^{i=n} \partial_i g(x(s)) \cdot \dot{x}_i(s) \quad (1.7)$$

$$= \sum_{i=1}^{i=n} \partial_i g(x(s)) f_i(x(s)). \quad (1.8)$$

Supposons donc tout d'abord que g est une fonction satisfaisant la condition (1.6). Alors pour toute solution x , d'après l'équation (1.8), on a bien $\frac{d}{dt} g \circ x = 0$. Ainsi, g est constante sur chaque orbite : c'est une intégrale première du mouvement.

Réciproquement supposons que g soit une intégrale première du mouvement. Nous admettons (par anticipation) que par tout point x_0 de Ω passe une solution x . Plus précisément, x est définie sur un intervalle ouvert I contenant 0 et vérifie $x(0) = x_0$. Alors en utilisant l'équation (1.8) en $s = 0$, on obtient

$$\sum_{i=1}^{i=n} \partial_i g(x_0) f_i(x_0) = 0.. \quad (1.9)$$

Ceci étant vrai pour tout x_0 , la condition (1.6) est vérifiée. $\square$

EXERCICE :

1. Donnez une interprétation géométrique de la condition (1.6) en termes du gradient de g et du champ de vecteur associé à l'équation.
2. On considère l'équation

$$m\ddot{x} = -\nabla f(x),$$

où l'on a noté ∇f le gradient de f . Montrez que

$$\frac{1}{2} m \dot{x}^2 + f,$$

est une intégrale première du mouvement.

3. En déduire le portrait de phase du pendule sans frottements.
4. On considère une fonction H de deux variables et le *système Hamiltonien*

$$\begin{cases} \dot{x} &= \partial_y H \\ \dot{y} &= -\partial_x H. \end{cases}$$

Montrez que H est une intégrale première du mouvement.

Physiquement, il est plus important de prévoir l'étude de l'évolution des fonctions, ou *observables*, que de celle des particules elles-mêmes.

Chapitre 2

Équations linéaires

2.1 Rappels d'algèbre linéaire

Nous renvoyons à [1].

Norme d'endomorphisme

Nous considérons E un espace vectoriel euclidien (ou hermitien) de dimension finie. Nous munirons $\text{End}(E)$ de la norme définie par

$$\|A\| = \sup\{\|A(x)\|; \|x\| \leq 1\}.$$

Nous avons alors

$$\|A(x)\| \leq \|A\| \cdot \|x\|, \quad \|A \cdot B\| \leq \|A\| \cdot \|B\|.$$

On remarque que si A est un endomorphisme de $\mathbb{R}^n$, sa norme est égale à celle en tant qu'endomorphisme de $\mathbb{C}^n$.

Exponentielle de matrice

Soit A une matrice – ou endomorphisme linéaire – son *exponentielle* est définie par la suite donnée par la série.

$$\exp(A) = 1 + A + \frac{1}{2}A^2 + \dots + \frac{1}{n!}A^n + \dots \quad (2.1)$$

Cette série est uniformément convergente pour A variant dans un ensemble borné. Nous rappelons quelques propriétés :

Proposition 2.1.0.4 [DIFFÉRENTIELLE DE L'EXPONENTIELLE] *L'application*

$$\begin{cases} \mathbb{R} & \rightarrow & \text{End}(E) \\ t & \mapsto & \exp(tA) \end{cases},$$

est différentiable et

$$\frac{d}{dt} \exp(tA) = A \cdot \exp(tA) = \exp(tA) \cdot A. \quad (2.2)$$

Proposition 2.1.0.5 [PROPRIÉTÉS ALGÈBRIQUES] *On a*

$$\exp(C^{-1}AC) = C^{-1} \exp(A)C, \quad (2.3)$$

$$\exp(-A) = (\exp(A))^{-1}. \quad (2.4)$$

Si A et B commute – c'est-à-dire $AB = BA$ – alors

$$\exp(A+B) = \exp(A) \exp(B) = \exp(B) \cdot \exp(A). \quad (2.5)$$

EXERCICE :

1. Calculez l'exponentielle des matrices suivantes :

$$D = \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix}, \quad C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad E = \begin{pmatrix} 1 & b \\ 0 & a \end{pmatrix}.$$

Matrices nilpotentes

Soit A une matrice sur un espace vectoriel complexe de dimension n . Si toutes les valeurs propres de A sont nulles, on dit que A est *nilpotente*. Alors $A^n = 0$. On rappelle enfin qu'une matrice nilpotente est conjuguée à une matrice triangulaire supérieure dont les éléments diagonaux sont nuls. Nous avons

Proposition 2.1.0.6 *Soit U une matrice nilpotente sur l'espace euclidien E de dimension n . Il existe alors un réel positif K tel que*

$$\forall t, \quad |t| > 1, \quad \text{alors } \|\exp(tU)\| \leq K|t|^n. \quad (2.6)$$

DÉMONSTRATION : En effet, nous avons

$$\exp(tU) = \sum_{p=0}^{n-1} \frac{t^p U^p}{p!}.$$

Et donc, si $|t| > 1$:

$$\begin{aligned} \|\exp(tU)\| &\leq \sum_{p=0}^{n-1} \left\| \frac{t^p U^p}{p!} \right\| \\ &\leq \sum_{p=0}^{n-1} \frac{\|U\|^p}{p!} |t|^n \\ &\leq \sum_{p=0}^{p=\infty} \frac{\|U\|^p}{p!} |t|^n \\ &\leq \exp(\|U\|) \cdot |t|^n \end{aligned}$$

Nous pouvons donc prendre $K = \exp(\|U\|)$.

Décomposition de Jordan-Dunford

Nous rappelons la proposition

Proposition 2.1.0.7 [JORDAN-DUNFORD] *Soit A une matrice à coefficient complexes. Alors, il existe une matrice diagonalisable D , une matrice nilpotente U telles que*

$$A = D + U \quad (2.7)$$

$$DU = UD. \quad (2.8)$$

2.2 Les équations linéaires à coefficients constants

Par définition une *équation linéaire à coefficients constants* s'écrit sous la forme

$$\dot{x} = A(x), \quad (2.9)$$

où A est une application linéaire de E dans lui-même. Dans la pratique, on identifie E à $\mathbb{R}^n$ par le choix d'une base et A à sa matrice dans cette base. Une première propriété est immédiate

Proposition 2.2.0.8 *Si x et y sont solutions de (2.9), alors $x+y$ est également solution. Plus généralement l'espace des solutions définies sur un intervalle I est un espace vectoriel.*

Nous savons résoudre explicitement ces équations.

Nous avons alors

Théorème 2.2.0.9 *Les solutions de (2.9), sont les courbes*

$$x(t) = \exp(tA)x_0.$$

Elles sont en particulier définies sur $\mathbb{R}$ tout entier.

DÉMONSTRATION : En effet, d'après la formule 2.2, nous avons bien

$$\frac{d}{dt}(\exp(tA) \cdot x_0) = A \cdot \exp(tA) \cdot x_0.$$

Nous en déduisons que les courbes $x(t) = \exp(tA)x_0$ sont solutions. Réciproquement, si $x(t)$ est solution, on vérifie que $\exp(-At) \cdot x(t)$ est constante. et donc

$$\exp(-A(t)) \cdot x(t) = x(0).$$

Ainsi,

$$x(t) = \exp(A(t)) \cdot x(0).$$

□

EXERCICE :

1. Montrez que si x est une solution dans E de $\dot{x} = A(x)$, alors $B(x)$ est solution dans F de $\dot{y} = BAB^{-1}(y)$ pour tout isomorphisme B de E dans F .
2. Montrez que l'espace des solutions $\mathcal{E}$ est un espace vectoriel et que pour tout réel t_0 ,

$$\begin{cases} \mathcal{E} & \rightarrow E \\ x & \mapsto x(t_0), \end{cases}$$

est un isomorphisme.

Attracteur pour une équation linéaire

Théorème 2.2.0.10 Soit A une endomorphisme linéaire d'un espace vectoriel complexe. On suppose que pour toute valeur propre λ de A , on a $\Re(\lambda) < 0$. Alors 0 est un point fixe attracteur de (2.9).

Plus précisément, il existe $\varepsilon > 0$ et $T > 0$, tels que si $x(t)$ est solution alors

$$\forall t > t_0 + T, \quad \|x(t)\| \leq \|x(t_0)\|e^{-\varepsilon t}. \quad (2.10)$$

De même, pour toute valeur propre λ de A , on a $\Re(\lambda) > 0$, alors 0 est un point fixe répulseur de (2.9)

DÉMONSTRATION : Il suffit de vérifier ce résultat pour une matrice A à coefficients complexes.

Première étape : On suppose dans un premier temps que A est diagonale avec toutes ses valeurs propres $(\lambda_1, \dots, \lambda_n)$ de partie réelle strictement négative. Alors $\exp(tA)$ est aussi diagonale et ses valeurs propres sont $(\exp(\lambda_1), \dots, \exp(\lambda_n))$. Nous avons alors

$$\exists \varepsilon > 0 \quad \forall t, \quad \|\exp(tA)\| \leq e^{-t\varepsilon}. \quad (2.11)$$

En effet, on a

$$\begin{aligned} \|\exp(tA)\| &\leq \sup_{i \in \{1, \dots, n\}} (|e^{t\lambda_i}|) \\ &\leq \sup_{i \in \{1, \dots, n\}} (e^{t\Re(\lambda_i)}) \end{aligned}$$

On peut donc prendre $\varepsilon = -\inf_{i \in \{1, \dots, n\}} (\Re(\lambda_i))$.

Deuxième étape : On suppose maintenant que A est diagonalisable. C'est à dire qu'il existe une matrice D diagonale dont toutes les valeurs propres ont une partie réelle strictement négative, et une matrice B telle que

$$A = BDB^{-1}.$$

En particulier,

$$\exp(tA) = B \exp(tD) B^{-1}.$$

Et donc

$$\|\exp(tA)\| \leq \|B\| \cdot \|B^{-1}\| \cdot \|\exp(tD)\|.$$

En combinant cette inégalité et l'assertion (2.11) appliquée à l'endomorphisme D , nous obtenons qu'il existe $\varepsilon > 0$ tel pour tout $t > 0$

$$\|\exp(tA)\| \leq \|B\| \cdot \|B^{-1}\| e^{-t\varepsilon}.$$

Nous en déduisons

$$\exists K_1 > 0, \exists \varepsilon > 0 \forall t, \|\exp(tA)\| \leq K_1 \cdot e^{-t\varepsilon}. \quad (2.12)$$

Troisième étape : On se place enfin dans le cas général. On écrit d'après la décomposition de Jordan-Dunford (2.1.0.7), $A = D + U$ avec D diagonalisable et U unipotente. Comme D et U commute, on a d'après l'assertion (2.5)

$$\exp(tA) = \exp(tD) \cdot \exp(tU).$$

Combinant les inégalités (2.12) et (2.6), nous obtenons enfin

$$\exists K_2, \exists \varepsilon > 0 \forall t > 1, \|\exp(tA) \cdot x\| \leq K_2 |t|^n \cdot e^{-t\varepsilon}.$$

Il suffit de prendre $K_2 = K_1 \cdot K$.

Démonstration du théorème : Nous pouvons maintenant finir de démontrer le théorème. Nous savons que si $x(t)$ est solution de (2.9). Alors

$$x(t) = \exp((t - t_0)A)x(t_0).$$

Par ailleurs, nous savons que

$$\lim_{t \rightarrow \infty} \frac{1}{K_2} t^{-n} e^{\frac{\varepsilon}{2}t} = +\infty.$$

Il existe donc $T > 0$ tel que

$$\forall s \geq T, s^n \leq \frac{1}{K_2} e^{\frac{\varepsilon}{2}s}.$$

Nous en déduisons donc

$$\forall t > t_0 + T, \|x(t)\| \leq \|x(t_0)\| e^{-\frac{\varepsilon}{2}t}.$$

□

2.3 Équations linéaires dépendant du temps

Soit I un intervalle ouvert, et f une application continue définie de $I \times E$ dans E . L'équation

$$\dot{x} = f(t, x),$$

est dite *linéaire*, si f est linéaire en x , c'est-à-dire pour tout $\lambda \in \mathbb{R}$, pour tout $t \in I$, pour tout $x, y \in E$, nous avons

$$\begin{aligned} f(t, \lambda x) &= \lambda f(t, x), \\ f(t, x + y) &= f(t, x) + f(t, y). \end{aligned} \quad (2.13)$$

Une manière équivalente de décrire une équation linéaire est de l'écrire sous la forme

$$\dot{x} = A(t)x, \quad (2.14)$$

où A est une application continue de I à valeurs dans les endomorphismes linéaires de E :

$$A : \begin{cases} I & \rightarrow \text{End}(E) \\ t & \mapsto A(t). \end{cases}$$

Contrairement au cas autonome, il n'est pas possible en général de résoudre ces équations "explicitement" à savoir à partir d'une procédure n'utilisant qu'une combinaison d'opérations algébriques, d'intégrales de différenciation etc. Nous montrerons pour ces équations notre premier grand théorème.

Théorème 2.3.0.11 *Pour tout x_0 de E et temps t_0 de I , il existe une solution x définie sur I tout entier et solution de*

$$\dot{x} = A(t)x.$$

De plus, si y est une solution de même condition initiale définie au voisinage de t_0 , y coïncide avec x sur son intervalle de définition.

EXERCICE :

Nous connaissons des solutions explicites par intégration dans quelques cas très particuliers :

1. On suppose que $\dim(E) = 1$, montrez que les solutions avec conditions initiales (t_0, x_0) de

$$\dot{x} = a(t)x,$$

sont données

$$x(t) = x(t_0)e^{\int_{t_0}^t a(s)ds}.$$

2. On suppose que pour tout t et s on a $A(t)A(s) = A(s)A(t)$. Montrez alors

$$x(t) = x(t_0) \exp \left(\int_{t_0}^t A(s)ds \right).$$

est solution de

$$\dot{x} = A(t)x.$$

avec conditions initiales (t_0, x_0) .

Avant de démontrer ce théorème, nous allons l'utiliser afin d'introduire une notion – celle de *résolvante* – qui nous donne une information sur le comportement collectif des solutions.

Résolvante d'une équation linéaire.

Nous admettrons dans ce paragraphe le théorème 2.3.0.11

Soit à nouveau une application continue

$$A : \begin{cases} I & \rightarrow \text{End}(E) \\ t & \mapsto A(t). \end{cases}$$

Par définition, pour tout t_0 dans I , la *résolvante* est la courbe

$$R_{t_0} : I \rightarrow \text{End}(E),$$

et solution de l'équation

$$\dot{R}_{t_0}(t) = A(t) \circ R_{t_0}^t, \quad R_{t_0}^{t_0} = \text{Id}. \quad (2.15)$$

EXERCICE :

1. Qu'est-ce que la résolvante d'une équation linéaire à coefficients constants $\dot{x} = Ax$?
2. En utilisant le théorème 2.3.0.11, montrez l'existence et l'unicité de la résolvante.

La proposition élémentaire suivante donne le lien entre la résolvante et les solutions de $\dot{x} = A(t)x$ dans E .

Proposition 2.3.0.12 *Soit R_{t_0} la résolvante de l'équation (2.14). Alors la solution de l'équation linéaire (2.14) dans E avec condition initiale (t_0, x_0) est*

$$x(t) = R_{t_0}^t.x_0$$

Autrement dit, la résolvante (une solution particulière) d'une équation, permet de construire toutes les solutions. Bien sûr, le gain n'est que conceptuel car la nouvelle équation a bien plus de paramètres.

Voici maintenant, un point de vue un peu plus abstrait sur la résolvante à éviter lors d'une première lecture :

EXERCICE :

1. L'espace des solutions $\mathcal{E}$ définies sur I est un espace vectoriel.
2. Pour tout t , l'application d'évaluation

$$\text{ev}_t \begin{cases} \mathcal{E} & \rightarrow E \\ x & \mapsto x(t), \end{cases}$$

est un isomorphisme.

3. on a

$$R_s^t = \text{ev}_t \circ \text{ev}_s^{-1}.$$

4. on a

$$R_t^s \circ R_u^t = R_u^s, \quad (R_t^s)^{-1} = R_s^t.$$

Preuve du théorème 2.3.0.11

Ce théorème découle du résultat suivant.

Théorème 2.3.0.13 *Soit J un intervalle compact inclus dans I . Alors pour tout x_0 de E et temps t_0 de J , il existe une unique courbe x définie sur J de conditions initiales (t_0, x_0) et solution de*

$$\dot{x} = A(t)x.$$

• *quelques rappels.* Nous rappelons le théorème du point fixe de Picard qui utilise la notion d'espace métrique complet pour laquelle nous renvoyons à [2].

Théorème 2.3.0.14 [POINT FIXE DE PICARD] *Soit M un espace métrique complet et T une application strictement contractante de M dans lui-même, c'est-à-dire telle qu'il existe $K < 1$ vérifiant*

$$\forall x, y \in M, \quad d(T(x), T(y)) \leq Kd(x, y).$$

Alors T possède un unique point fixe.

Notons T^p l'itérée p -fois de T :

$$T^p = \underbrace{T \circ T \circ \dots \circ T}_{p \text{ fois}}.$$

Rappelons la preuve de ce théorème

DÉMONSTRATION : Une récurrence simple nous montre que

$$d(T^p(x), T^p(y)) \leq K^p d(x, y). \quad (2.16)$$

Montrons tout d'abord l'unicité du point fixe. Si x_0 et y_0 sont deux points fixes, nous obtenons que

$$d(x_0, y_0) = d(T^p(x_0), T^p(y_0)) \leq K^p d(x_0, y_0).$$

Puisque $K < 1$, nous en déduisons en faisant tendre p vers l'infini que $d(x_0, y_0) = 0$ et donc que $x_0 = y_0$.

Montrons maintenant l'existence du point fixe. Soit x un point quelconque. Posons $y_p = T^p(x)$. Nous déduisons de l'inégalité (2.16) que

$$d(y_p, y_{p+1}) \leq K^p d(x, T(x)).$$

Nous en déduisons par l'inégalité triangulaire si $q > p$

$$\begin{aligned} d(y_p, y_q) &\leq \sum_{i=p}^{i=q-1} d(y_i, y_{i+1}) \\ &\leq K^p d(x, T(x)) \sum_{i=0}^{i=q-p-1} K^i \\ &\leq K^p \frac{d(x, T(x))}{1-K}. \end{aligned}$$

Ainsi, la suite $\{y_p\}_{p \in \mathbb{N}}$ est une suite de Cauchy. Par hypothèse M est complet, cette suite converge donc. Soit x_0 sa limite. L'application T étant continue nous avons :

$$T(x_0) = \lim_{p \rightarrow \infty} T(y_p) = \lim_{p \rightarrow \infty} y_{p+1} = x_0.$$

Nous venons de démontrer que x_0 est un point fixe de T . $\square$

Corollaire 2.3.0.15 *On suppose qu'il existe $p > 0$ tel que T^p est strictement contractante. Alors T a un unique point fixe.*

DÉMONSTRATION : Tout point fixe de T est point fixe de T^p . Donc T a au plus un point fixe d'après le théorème de Picard appliqué à T^p . Par ailleurs, le théorème de Picard nous donne l'existence d'un point fixe x de T^p . Mais alors

$$T^p(T(x)) = T(T^p(x)) = T(x).$$

Donc $T(x)$ est aussi point fixe de T^p . Par unicité du point fixe on a $T(x) = x$. Ainsi, x est point fixe de T . $\square$

Nous munissons E d'une norme euclidienne. Nous considérons l'espace $M = C^0(J, E)$ des courbes continues de J à valeurs dans E muni de la norme

$$\|x\|_\infty = \sup_{t \in J} \|x(t)\|.$$

Nous rappelons que M est alors un espace métrique complet pour la métrique

$$d(x, y) = \|x - y\|_\infty.$$

• *Une construction.* Nous considérons l'application

$$T : \begin{cases} M & \rightarrow M \\ y & \mapsto T(y) \text{ où } T(y)(t) = x_0 + \int_{t_0}^t A(u)y(u)du. \end{cases}$$

Nous remarquons que x est solution de $\dot{x} = A(t)x$, avec conditions initiales (t_0, x_0) si et seulement si $T(x) = x$. D'après le corollaire du théorème de Picard, il nous suffit de démontrer le lemme suivant.

Lemme 2.3.0.16 *Il existe $p > 0$ tel que T^p est strictement contractante : il existe $K < 1$ vérifiant*

$$\forall x \in M, \quad \|T^p(x) - T^p(y)\| \leq K\|x - y\|.$$

DÉMONSTRATION : On a

$$(T(x) - T(y))(t) = \int_{t_0}^t A(s)(x(s) - y(s))ds.$$

Comme J est borné et que A est continue. Il existe une constante K_0 telle que

$$\forall s \in J, \quad \|A(s)\| \leq K_0.$$

On obtient ainsi

$$\|(T(x) - T(y))(t)\| \leq K_0 \int_{t_0}^t \|x(s) - y(s)\| ds$$

Nous allons montrer par récurrence

$$\|(T^p(x) - T^p(y))(t)\| \leq K_0^p \|x - y\|_\infty \frac{|t - t_0|^p}{p!} \quad (2.17)$$

En effet, supposons cette inégalité pour un certain p . Alors

$$\begin{aligned} \|(T^{p+1}(x) - T^{p+1}(y))(t)\| &\leq K_0 \int_{t_0}^t \|T^p x(s) - T^p y(s)\| ds \\ &\leq K_0^{p+1} \|x - y\|_\infty \int_{t_0}^t \frac{|s - t_0|^p}{p!} ds \\ &\leq K_0^{p+1} \|x - y\|_\infty \frac{|s - t_0|^{p+1}}{(p+1)!} \end{aligned}$$

Nous venons de montrer l'inégalité 2.17. Nous en déduisons en posant $J = [a, b]$ que pour tout p

$$\|T^p(x) - T^p(y)\| \leq \frac{(K_0(b-a))^p}{p!} \|x - y\|_\infty.$$

Comme

$$\lim_{p \rightarrow \infty} \frac{(K_0(b-a))^p}{p!} = 0,$$

Il existe $p > 0$ tel que

$$K = K_0 \frac{(b-a)^p}{p!} < 1$$

C'est ce que nous voulions démontrer. $\square$

2.4 Équations différentielles affines et variation des constantes

Soit

$$A : \begin{cases} I & \rightarrow & \text{End}(E) \\ t & \mapsto & A(t). \end{cases}, \quad b : \begin{cases} I & \rightarrow & E \\ t & \mapsto & b(t). \end{cases}$$

Une *équation différentielle affine* dans un espace vectoriel E se présente sous la forme

$$\dot{x} = A(t).x + b(t). \quad (2.18)$$

L'*équation différentielle linéaire associée* est

$$\dot{y} = A(t).y \quad (2.19)$$

Les solutions de ses deux équations ont une relation naturelle entre elles :

Proposition 2.4.0.17 *Si x est une solution d'une équation différentielle affine et si y est une solution de l'équation différentielle linéaire associée alors $x + y$ est une solution de l'équation différentielle affine. De même si x et z sont deux solutions de l'équation différentielle affine, alors $x - z$ est une solution de l'équation linéaire associée.*

Une conséquence de cette proposition est que si l'on connaît une solution de l'équation affine, alors toutes les autres solutions s'obtiennent en lui additionnant une solution de l'équation linéaire associée.

Ceci est bien utile lorsque l'on connaît une solution particulière. Sinon, on utilise la méthode de la *variation des constantes*. Supposant connue la résolvante de l'équation linéaire associée, cette méthode permet de trouver toutes les solutions de l'équation affine par une simple intégration.

- Soit $t \rightarrow R_{t_0}^t$ la résolvante de l'équation linéaire. Par définition de la résolvante, les solutions de l'équation linéaire (2.19) sont de la forme $R_{t_0}^t c$ où c est une constante.
- On cherche ensuite les solutions de l'équation affine (2.18) sous la forme

$$x(t) = R_{t_0}^t c(t)$$

où

$$c : \begin{cases} I & \rightarrow & \text{End}(E) \\ t & \mapsto & A(t). \end{cases}, \quad c(t_0) = x_0.$$

Comme $R_{t_0}^t$ est inversible d'inverse $R_{t_0}^t$, on a $c(t) = R_{t_0}^t(x(t))$. Autrement dit, on *fait varier la constante c* . Comme,

$$\dot{x} = \dot{R}_{t_0}^t c + R_{t_0}^t \dot{c} \quad (2.20)$$

$$= A R_{t_0}^t . c + R_{t_0}^t \dot{c}, \quad (2.21)$$

$$= A x + R_{t_0}^t \dot{c} \quad (2.22)$$

On voit que résoudre l'équation (2.18) en x , est équivalent à résoudre en c l'équation.

$$\dot{c} = R_{t_0}^{-1}b, \quad c(t_0) = x_0.$$

Cette dernière se résout par intégration :

$$c(t) = x_0 + \int_{t_0}^t (R_{t_0}^s)^{-1}b(s)ds.$$

Par ailleurs, rappelons que

$$(R_{t_0}^s)^{-1} = R_s^{t_0}, \text{ et } R_{t_0}^t \circ R_s^{t_0} = R_s^t.$$

Nous pouvons donc résumer cette discussion sous le théorème suivant (Réfléchissez à l'unicité).

Théorème 2.4.0.18 [ÉQUATIONS DIFFÉRENTIELLES AFFINES] *On considère l'équation différentielle affine*

$$\dot{x} = A(t)x + b(t).$$

Soit $t \rightarrow R_{t_0}^t$ la résolvante de l'équation linéaire associée

$$\dot{y} = A(t)y.$$

La solution de condition initiale (t_0, x_0) est définie sur I tout entier et est donnée par

$$x(t) = R_{t_0}^t x_0 + \int_{t_0}^t R_s^t b(s)ds \quad (2.23)$$

On remarque que le premier terme de droite de l'égalité (2.23) est la solution de l'équation linéaire associée de condition initiale (t_0, x_0) et que le deuxième terme est la solution de l'équation différentielle associée de conditions initiales $(0, t_0)$. Cette décomposition est tout à fait naturelle d'après la proposition 2.4.0.17

Une dernière remarque : il est inutile de se souvenir de la forme précise de la solution donnée par l'équation (2.23). Par contre, *il faut se souvenir de la méthode qui nous y a menés.*

EXERCICE :

1. Donnez explicitement les solutions de l'équation affine à coefficients constants

$$\dot{x} = Ax + b,$$

où A et b ne dépendent pas du temps.

2. En utilisant l'exercice précédent, résoudre en discutant suivant les valeurs des constantes a, b, c l'équation différentielle du deuxième ordre

$$\ddot{x} = ax + bx + c.$$

3. Montrez que l'espace $\mathcal{F}$ des solutions de l'équation différentielle affine est un espace affine dont l'espace vectoriel associé est l'espace $\mathcal{E}$ des solutions de l'équation différentielle linéaire associée.
4. Montrez que les solutions de l'équation différentielle affine sur E avec conditions initiales (t_0, x_0) se déduisent des solutions avec conditions initiales $((x_0, 1), t_0)$ de l'équation différentielle linéaire sur $\mathbb{R} \times E$

$$\begin{cases} \dot{x} &= A(t)x + b(t)s \\ \dot{s} &= 0. \end{cases}$$

2.5 Wronskien et théorème de Liouville : l'évolution de déterminant de la résolvante

Soit

$$\dot{x} = A(t)x,$$

une équation différentielle linéaire et $t \rightarrow R_{t_0}^t$ sa résolvante. Nous allons montrer

Théorème 2.5.0.19 [LIOUVILLE] *Soit $v : t \rightarrow \det(R_{t_0}^t)$. Alors la fonction v vérifie l'équation différentielle linéaire*

$$\dot{v} = \text{trace}(A(t))v.$$

En particulier, on a

$$\det(R_{t_0}^t) = e^{\int_{t_0}^t \text{trace}(A(s))ds}$$

On remarque que en particulier on a toujours $\det(R_{t_0}^t) \neq 0$. Autrement dit, $R_{t_0}^t$ est inversible. Nous le savions déjà, pourquoi ?

DÉMONSTRATION : Soit $(e_1, \dots, e_n)$ une base de E . Soit ω une forme multilinéaire alternée sur E de degré maximal. Rappelons que

$$\begin{aligned} \omega(Ae_1, \dots, Ae_n) &= \det(A)\omega(e_1, \dots, e_n) \\ \sum_{i=1}^{i=n} \omega(e_1, \dots, Ae_i, \dots, e_n) &= \text{trace}(A)\omega(e_1, \dots, e_n). \end{aligned}$$

Ainsi soit

$$h(t) = \omega(R_{t_0}^t e_1, \dots, R_{t_0}^t e_n).$$

En dérivant, il vient

$$\begin{aligned}
 \dot{h}(t) &= \frac{d}{dt} \omega(R_{t_0}^t e_1, \dots, R_{t_0}^t e_n) \\
 &= \sum_{i=1}^n \omega(R_{t_0}^t e_1, \dots, \frac{d}{dt} R_{t_0}^t e_i, \dots, R_{t_0}^t e_n) R_{t_0}^t e_i \\
 &= \sum_{i=1}^n \omega(R_{t_0}^t e_1, \dots, A(t) R_{t_0}^t e_i, \dots, R_{t_0}^t e_n) R_{t_0}^t e_i \\
 &= \text{trace}(A(t)) \omega(R_{t_0}^t e_1, \dots, R_{t_0}^t e_n) \\
 &= \text{trace}(A(t)) h(t).
 \end{aligned}$$

Nous concluons en remarquant que $h(t) = \omega(e_1, \dots, e_n) v(t)$. $\square$

Soit en général ω une forme multilinéaire alternée sur E de degré $n = \dim(E)$. Si $(x_1, \dots, x_n)$ sont n solutions de

$$\dot{x} = A(t)x.$$

Le *Wronskien* de la famille de solutions est

$$W(t) = \omega(x_1(t), \dots, x_n(t)).$$

Nous venons également de montrer

Théorème 2.5.0.20 *Le Wronskien $W(t)$ vérifie l'équation différentielle linéaire*

$$\dot{W} = \text{trace}(A(t))W.$$

En particulier, on a

$$W(t) = W(t_0) e^{\int_{t_0}^t \text{trace}(A(s)) ds}$$

Chapitre 3

Le théorème fondamental de Cauchy-Lipschitz

3.1 Hypothèses et énoncé

3.1.1 Fonctions k -lipschitziennes

Soit M et N deux espaces métriques. Soit g une application de M dans N . Soit k un nombre réel. On dit que g est k -lipschitzienne si

$$\forall x, y \in M, \quad d(g(x), g(y)) \leq k.d(x, y).$$

On dira aussi que qu'une fonction f est *localement lipschitzienne*, si tout point possède un voisinage sur lequel f est lipschitzienne.

REMARQUES :

1. Il est immédiat de vérifier que toute fonction k -lipschitzienne est continue.
2. La réciproque n'est pas vraie : il existe des fonctions continues qui ne sont pas lipschitziennes, même localement. Un exemple simple est la fonction f de $\mathbb{R}$ dans lui-même définie par $x \rightarrow \sqrt{|x|}$. Comme

$$\lim_{x \rightarrow 0} \frac{|f(x) - f(0)|}{|x - 0|} = +\infty,$$

la fonction f n'est lipschitzienne pour la métrique usuelle sur aucun voisinage de 0. item

3. Par contre, soit E et F deux espaces vectoriels normés. Soit U un ouvert de E . Soit f définie sur U à valeurs dans F . Si f est C^1 sur U , alors f est localement lipschitzienne. En effet, la différentielle de f dépend continûment du point. Dès lors, tout point x de U possède un voisinage

convexe C – une boule – sur lequel $\|Df\|$ est bornée par un nombre k . D’après le théorème des accroissements finis on a

$$\forall x, y \in C, \quad \|f(x) - f(y)\| \leq k\|x - y\|.$$

EXERCICE :

1. Soit E et F deux espaces vectoriels normés. On suppose que E est dimension finie. Soit U un ouvert de E . Soit f définie sur U à valeurs dans F . Soit f une application C^1 sur U , montrez que la restriction de f à tout compact C de U est k -lipschitzienne. *Indication* : On montre tout d’abord par récurrence que si K est une réunion finie de N boules fermées alors

$$k = N \cdot \sup_{x \in K} \|D_x f\|.$$

convient. On traite ensuite le cas général.

3.1.2 Un premier énoncé.

Soit maintenant E un espace vectoriel normé complet. On rappelle que l’on note $B(x, r)$ la boule de centre x de rayon r . Soit Ω un ouvert de E , I un intervalle ouvert de $\mathbb{R}$ et $f : I \times \Omega \rightarrow E$ une application continue. On considère l’équation différentielle

$$\dot{x} = f(t, x). \quad (3.1)$$

Nous allons démontrer le théorème suivant. Nous choisissons l’énoncé le plus simple et l’affinerons par la suite.

Théorème 3.1.2.1 [CAUCHY-LIPSCHITZ] *On suppose qu’il existe une constante $k > 0$, telle que pour tout t dans I , l’application (à t fixé)*

$$x \rightarrow f(t, x).$$

est k -lipschitzienne. Alors pour tout point y_0 de Ω , pour tout s_0 de I , il existe

- *un voisinage U de y_0 dans Ω ,*
- *un intervalle ouvert I_0 contenant s_0 ,*

tel que pour tout $(t_0, x_0) \in I_0 \times U$, il existe une solution $x(t)$ définie sur I_0 de l’équation (3.1), avec condition initiale (t_0, x_0) .

Enfin si deux solutions x_1 et x_2 définies sur des intervalles ouverts I_1 et I_2 ont les mêmes conditions initiales $(t_0, x_0) \in I_0 \times U$ alors ces solutions coïncident sur un voisinage t_0 .

REMARQUES :

1. Ce théorème énonce dans un premier temps l’existence des solutions au moins sur un court intervalle de temps, qui peut être choisi constant au voisinage de chaque condition initiale. Nous verrons plus tard dans le lemme 3.2.1.2 que chaque solution possède une “durée de vie” naturelle.

2. Le théorème 5.1, démontré plus tard, précise que la solution dépend continûment des conditions initiales .
3. Enfin, il montre que deux solutions de même conditions initiales coïncident au moins sur un court intervalle de temps. Nous améliorerons plus tard ce résultat d'unicité dans le lemme 3.2.1.1.
4. Le même énoncé est vrai avec l'hypothèse f de classe C^1 . En effet, dans cette situation, pour tout (y_0, s_0) , il existe un voisinage I_0 de s_0 un voisinage Ω_0 de y_0 telle que f est k -lipschitzienne sur Ω_0 à t fixé dans I_0 , par le théorème des accroissements finis.

3.1.3 Démonstration du théorème

Nous allons donner une démonstration du théorème Cauchy-Lipschitz. Nous allons introduire comme précédemment une application T d'un espace métrique dans lui-même M . La difficulté par rapport à la démonstration du cas linéaire est de s'assurer que la transformation T est bien définie. C'est l'objet de notre premier lemme.

Notations et rappel

Soit $s_0 \in I$. Nous notons pour tout $\varepsilon > 0$,

$$I_\varepsilon = [s_0 - \varepsilon, s_0 + \varepsilon].$$

Nous notons également,

$$E_\varepsilon = C^0(I_\varepsilon, E),$$

l'espace des applications continues de l'intervalle I_ε dans E . On le munit de la norme

$$\|x\| = \sup_{t \in I_\varepsilon} \|x(t)\|.$$

L'espace vectoriel E_ε est alors un espace de Banach – c'est-à-dire un espace vectoriel normé complet. Nous rappelons que si L est un fermé de E , alors

$$C^0(I_\varepsilon, L) = \{f \in E_\varepsilon \mid f(I_\varepsilon) \subset L\},$$

est un sous-ensemble fermé de E_ε et qui en particulier est complet [2].

Soit $s_0 \in I$. Soit β un nombre positif tel que $I_\beta \subset I$.

Inégalité de sécurité

Soit y_0 un point de Ω . Soit $R > 0$ tel que

$$\overline{B}(y_0, 2R) \subset \Omega.$$

On pose

$$M_\varepsilon = C^0(I_\varepsilon, \overline{B}(y_0, 2R)).$$

Soit

$$U = B(y_0, R).$$

Pour tout $\varepsilon \leq \beta$ et (t_0, x_0) de $I_\varepsilon \times U$, nous considérons l'application

$$T_{(t_0, x_0)} : \begin{cases} M_\varepsilon & \rightarrow E_\varepsilon \\ y & \mapsto T_{(t_0, x_0)}(y) \end{cases} \text{ où } T_{(t_0, x_0)}(y) : t \mapsto x_0 + \int_{t_0}^t f(s, y(s)) ds.$$

Cette application est bien définie. Lorsqu'il n'y aura pas d'ambiguïté, on omettra l'indice (t_0, x_0) .

Notre lemme de "sécurité" est le suivant

Lemme 3.1.3.1 [SÉCURITÉ] *Il existe une constante $\varepsilon_1 \leq \beta$ telle que si $\varepsilon \leq \varepsilon_1$ et $(t_0, x_0) \in I_\varepsilon \times U$ alors*

$$T_{(t_0, x_0)}(M_\varepsilon) \subset M_\varepsilon.$$

DÉMONSTRATION : Posons

$$A = \sup_{(t, x) \in I_{\beta_0} \times B_R} \|f(t, x)\|.$$

De plus, comme f est k -lipschitzienne et I_β est compact, on a

$$A \leq 2kR + \sup_{s \in J_{\beta_0}} \|f(y_0, s)\| < \infty.$$

Posons

$$\varepsilon_1 = \inf\left(\frac{R}{2A}, \beta\right).$$

Soit $\varepsilon \leq \varepsilon_1$. Soit $y \in M_\varepsilon$. Soit $t \in I_\varepsilon$. Alors

$$\begin{aligned} \|T(y)(t) - y_0\| &\leq \|x_0 - y_0\| + \left\| \int_{t_0}^t f(y(s), s) ds \right\| \\ &\leq R + \int_{t_0}^t \|f(y(s), s)\| ds \\ &\leq R + \int_{t_0}^t A ds \\ &\leq R + 2A\varepsilon \leq 2R. \end{aligned}$$

Autrement dit

$$\forall t \in I_\varepsilon, \quad T(y)(t) \in \overline{B}(y_0, 2R).$$

C'est-à-dire $T(y) \in M_\varepsilon$. $\square$

Existence et unicité de la solution

Nous utiliserons les mêmes notations que précédemment. Nous considérons ε_1 donné par le lemme ci-dessus. Soit toujours

$$T_{(t_0, x_0)} : \begin{cases} M_\varepsilon & \rightarrow E_\varepsilon \\ y & \mapsto T(y) \text{ où } T(y) : t \mapsto x_0 + \int_{t_0}^t f(y(s), s) ds. \end{cases}$$

Nous remarquons à nouveau que x est solution de conditions initiales (t_0, x_0) sur I_ε , si et seulement si $T_{(t_0, x_0)}(x) = x$. Comme pour les équations linéaires, le théorème découle du théorème de Picard et du lemme suivant.

Lemme 3.1.3.2 *Il existe $\varepsilon_0 \leq \varepsilon_1$, tel que si $\varepsilon \leq \varepsilon_0$, alors pour tout $(t_0, x_0) \in I_\varepsilon \times U$, on a*

$$\forall x, y \in M_\varepsilon, \quad \|T_{(t_0, x_0)}(x) - T_{(t_0, x_0)}(y)\| \leq \frac{1}{2} \|x - y\|.$$

DÉMONSTRATION : Soit

$$\varepsilon_0 = \inf(\varepsilon_1, \frac{1}{4k}).$$

Nous avons

$$\begin{aligned} \|T(x)(t) - T(y)(t)\| &\leq \left\| \int_{t_0}^t (f(s, x(s)) - f(s, y(s))) ds \right\| \\ &\leq \int_{t_0}^t \|f(s, x(s)) - f(s, y(s))\| ds. \end{aligned}$$

En utilisant le fait que $z \mapsto f(t, z)$ est k -lipschitzienne, il vient

$$\begin{aligned} \|T(x)(t) - T(y)(t)\| &\leq \int_{t_0}^t k \|x(s) - y(s)\| ds \\ &\leq \|x - y\|_\infty \int_{t_0}^t k ds \\ &\leq k\varepsilon \|x - y\|_\infty \end{aligned}$$

Autrement dit, si

$$\varepsilon \leq \varepsilon_0 = \inf(\varepsilon_1, \frac{1}{2k}),$$

il vient

$$\|T(x) - T(y)\|_\infty \leq \frac{1}{2} \|x - y\|_\infty.$$

C'est l'inégalité que nous recherchions $\square$

3.2 Solutions maximales, durée de vie des solutions

Nous allons démontrer quatre lemmes qui nous permettent d'améliorer notre connaissance des solutions. Nous considérons comme d'habitude l'équation

$$\dot{x} = f(t, x), \quad (3.2)$$

où f est définie sur $I \times \Omega \subset E \times R$. On supposera dans cette section que

- soit f est k -lipschitzienne à t fixé,
- soit f est C^1 sur Ω .

3.2.1 Unicité, solutions maximales et prolongement.

Lemme 3.2.1.1 [UNICITÉ ET PROLONGEMENT] *Soit x_1 définie sur l'intervalle ouvert I_1 et x_2 définie sur l'intervalle ouvert I_2 , deux courbes à valeurs dans E . On suppose que x_1 et x_2 sont deux solutions de l'équation 3.2 de même condition initiale (x_0, t_0) . Alors*

- x_1 et x_2 sont égales sur l'intervalle $J = I_1 \cap I_2$.
- Il existe une solution y définie sur $I_1 \cup I_2$ qui coïncide avec x_1 et x_2 sur I_1 et I_2 respectivement.

DÉMONSTRATION : Considérons le sous ensemble F de J :

$$F = \{t \in J / x_1(t) = x_2(t)\}.$$

Les courbes x_1 et x_2 étant continues, l'ensemble F est un fermé de J . Il est non vide car il contient t_0 . Montrons qu'il est ouvert. Soit $s \in F$. Les courbes x_1 et x_2 peuvent être alors vues comme des solutions de l'équation 3.2 de même condition initiale $(s, x_1(s), s) = (s, x_2(s))$. D'après le théorème de Cauchy-Lipschitz, elles coïncident sur un petit voisinage de s . Ainsi F est ouvert. Puisque F est non vide, ouvert et fermé et que J est un intervalle, nous avons $F = J$. Pour la deuxième partie, nous définissons y par $y = x_1$ sur I_1 et $y = x_2$ sur I_2 . Cette définition est sans ambiguïté d'après ce qui précède. $\square$

Lemme 3.2.1.2 [SOLUTION MAXIMALE, DURÉE DE VIE] *Soit $(t_0, x_0) \in I \times \Omega$. Il existe alors une solution dite maximale x définie sur un intervalle dit durée de vie J ayant la propriété suivante : si s appartient à I , et si y est une solution de condition initiale $(s, x(s))$ défini sur un intervalle $\hat{J}$, alors $\hat{J} \subset J$ et pour tout u de $\hat{J}$, $y(u) = x(u)$.*

DÉMONSTRATION : Ceci est une conséquence immédiate de la deuxième partie du lemme d'unicité 3.2.1.1. $\square$

3.2.2 Prolongement, sortie des compacts et explosion en temps fini

Lemme 3.2.2.1 [PROLONGEMENT DES SOLUTIONS] *Soit x une solution de l'équation 3.2 définie sur l'intervalle $J =]a, b[$ où $[a, b] \subset I$. On suppose qu'il existe une suite $\{t_n\}_{n \in \mathbb{N}}$ d'éléments de I convergeant vers b , telle*

$$\lim_{n \rightarrow \infty} x(t_n) = x_0 \in \Omega.$$

Alors il existe $\varepsilon > 0$ tel que x s'étend en une solution définie sur $]a, b + \varepsilon[$.

DÉMONSTRATION : D'après le théorème de Cauchy-Lipschitz, il existe un voisinage U de x_0 et $\varepsilon > 0$, telle que pour tout point (s, z) de $]b - \varepsilon, b + \varepsilon[\times U$, il existe une solution $\bar{x}$ de conditions initiales (s, z) et définie sur l'intervalle $]b - \varepsilon, b + \varepsilon[$. Nous appliquons ce résultat à $(s, z) = (t_n, x(t_n))$ pour n suffisamment grand de telle sorte que $|t_n - b| \leq \varepsilon/2$. En particulier $\bar{x}$ est définie sur $]b - \varepsilon/2, b + \varepsilon/2[$. Comme x est aussi solution de conditions initiales $(t_n, x(t_n))$. Nous en déduisons d'après la deuxième partie du lemme d'unicité que nous pouvons étendre x à $]a, b + \varepsilon/2[$. $\square$

On peut alors donner un nouvel énoncé du lemme sur le prolongement des solutions. Ce nouvel énoncé est particulièrement utile lorsque $I = \mathbb{R}$.

Lemme 3.2.2.2 [SORTIE DES COMPACTS] *On se donne une équation différentielle donnée par une fonction f définie sur $I \times \Omega$ où $I =]a, b[$. On se donne une solution maximale x définie sur l'intervalle $]t^-, t^+[$. On suppose que $t^+ < b$.*

Alors la solution x sort de tout compact de Ω vers le futur. C'est-à-dire, pour tout compact K de Ω , il existe $t_K < t^+$ telle que

$$t > t_K \implies x(t) \notin K.$$

Explosion en temps fini

Nous utiliserons librement la terminologie suivante : nous dirons qu'une courbe c à valeurs dans E – ou une fonction à valeurs dans $\mathbb{R}$ – *explose en un temps fini vers le futur*, si elle est définie sur un intervalle $[a, t_0[$ avec $t_0 < \infty$ et

$$\lim_{t \rightarrow t_0} \|c(t)\| = +\infty.$$

On vérifie aisément qu'une courbe qui explose en un temps fini sort de tout compact. Le temps t_0 s'appelle *temps d'explosion* (vers le futur).

EXERCICE :

1. Quelles sont les solutions maximales de $\dot{x} = -x^2$.

3.3 Retour sur Cauchy-Lipschitz

En combinant les résultats précédents nous obtenons le théorème suivant qui affine le théorème de Cauchy-Lipschitz. Nous avons perdu le résultat sur le contrôle de l'intervalle d'existence des solutions, mais nous obtiendrons un résultat plus – plus tard – sur la dépendance des solutions par rapport aux conditions initiales.

Théorème 3.3.0.3 [EXISTENCE ET UNICITÉ DES SOLUTIONS] *On suppose que f définie sur $I \times \Omega \subset \mathbb{R} \times E$ vérifie l'une des deux hypothèses suivantes*

- *f est C^1 ,*
- *il existe k , telle que pour tout t , l'application $x \mapsto f(t, x)$ est k -lipschitzienne.*

Alors pour toute (t_0, x_0) de $I \times \Omega$, il existe un intervalle ouvert I contenant t_0 , une courbe C^1 solution de l'équation

$$\dot{x} = f(t, x),$$

avec conditions initiales (t_0, x_0) . De plus si y est une solution définie sur un intervalle J de même condition initiale (t_0, x_0) , alors $J \subset I$ et y est la restriction de x à J .

Chapitre 4

Le lemme de Gronwall

Le lemme de Gronwall est un résultat sur les inéquations différentielles. Il permet de contrôler des solutions d'équations différentielles que nous ne savons pas expliciter par des solutions d'équations plus simples dont nous connaissons les solutions.

Il est d'une importance pratique cruciale.

Nous en donnerons également une utilisation théorique en l'utilisant pour démontrer la dépendance continue des solutions par rapport aux conditions initiales.

4.1 Énoncés

Le lemme de Gronwall est un plus un principe qu'un résultat précis : il se présente sous plusieurs formes. Comprendre sa preuve est plus important que de se souvenir des différentes variantes.

4.1.1 Gronwall avec des inégalités strictes

Lemme 4.1.1.1 [GRONWALL] *Soit y une solution de l'équation différentielle $\dot{y} = g(t, y)$ définie sur l'intervalle $I = [t_0, t_1[$. On suppose que g est continue. Soit x une courbe dérivable dans E définie sur I telle que*

$$\begin{aligned} \|x(t_0)\| &\leq y(t_0) \\ \forall t \in I, \quad \|\dot{x}(t)\| &< g(\|x(t)\|, t). \end{aligned}$$

Alors

$$\forall t \in]t_0, t_1[, \quad \|x(t)\| < y(t). \quad (4.1)$$

DÉMONSTRATION :

Dans un premier temps, nous allons montrer qu'il existe $s > t_0$ tel que

$$\forall t \in]t_0, s[, \quad \|x(t)\| < y(t). \quad (4.2)$$

Si $\|x(t_0)\| < y(t_0)$, c'est évident par continuité. Concentrons nous sur le cas $\|x(t_0)\| = y(t_0)$. On écrit un développement limité à l'ordre 1 au voisinage de t_0 . On a

$$\begin{aligned} y(t) &= y(t_0) + \dot{y}(t_0)(t - t_0) + o_y(t - t_0) \\ &= y(t_0) + g(y(t_0), t_0)(t - t_0) + o_y(t - t_0) \\ x(t) &= x(t_0) + \dot{x}(t_0)(t - t_0) + o_x(t - t_0). \end{aligned}$$

Pour $t > t_0$, nous obtenons

$$\begin{aligned} \|x(t)\| &\leq \|x(t_0)\| + \|\dot{x}(t_0)\|(t - t_0) + o_x(t - t_0) \\ &\leq y(t_0) + g(y(t_0), t_0)(t - t_0) - a(t - t_0) + o_x(t - t_0). \end{aligned}$$

Par ailleurs, au voisinage de t_0 et pour $t > t_0$, en posant $a = g(\|x(t_0)\|, t_0) - \|\dot{x}(t_0)\| > 0$, nous avons

$$o_x(t - t_0) > o_y(t - t_0) - a(t - t_0),$$

Nous obtenons ainsi, toujours pour $t > t_0$ et au voisinage de t_0

$$\|x(t)\| < y(t_0) + g(y(t_0), t_0)(t - t_0) + o_y(t - t_0) = y(t).$$

Nous venons de démontrer que l'inégalité (4.1) est vraie au voisinage de t_0 . Il nous reste à montrer qu'elle est vraie sur $]t_0, t_1[$ tout entier.

Raisonnons par l'absurde et supposons qu'il existe s tel que

$$\|x(s)\| \geq y(s).$$

Soit donc

$$t_2 = \inf\{t > t_0, \|x(t)\| \geq y(t)\}.$$

D'après la première partie de notre démonstration, on a $t_2 > t_0$. Commençons par remarquer que par continuité de x et y , on a

$$\|x(t_2)\| = y(t_2).$$

Nous effectuons maintenant un développement limité au voisinage de t_2 , pour $t < t_2$.

$$\begin{aligned} y(t) &= y(t_2) + g(y(t_2), t_2)(t - t_2) + o_y(t - t_2) \\ \|x(t)\| &\geq \|x(t_2)\| - \|\dot{x}(t_2)\|(t - t_2) - o_x(t - t_2) \\ &\geq \|x(t_2)\| + \|\dot{x}(t_2)\|(t - t_2) - o_x(t - t_2) \end{aligned}$$

Un raisonnement analogue au raisonnement précédent montre alors qu'au voisinage de t_2 pour $t < t_2$, on a

$$\|x(t)\| > y(t).$$

Mais ceci est en contradiction avec la définition de t_2 . $\square$

4.1.2 Gronwall classique

Voici la version classique du lemme de Gronwall. C'est un cas particulier du lemme précédent. Nous laissons sa démonstration en exercice.

Lemme 4.1.2.1 [GRONWALL CLASSIQUE] *Soit x une courbe dérivable dans E définie sur I solution de l'équation différentielle*

$$\dot{x} = f(t, x).$$

On suppose que

$$\|f(t, x)\| < a\|x\| + b.$$

Alors pour tout t de I

$$\text{si } a \neq 0, \quad \|x(t)\| \leq \|x(t_0)\|e^{a|t-t_0|} + \frac{b}{a}(e^{a|t-t_0|} - 1) \quad (4.3)$$

$$\text{si } a = 0, \quad \|x(t)\| \leq \|x(t_0)\| + b|t - t_0|. \quad (4.4)$$

REMARQUE : : Dans cet énoncé, on contrôle la solution à la fois pour les temps plus grand que t_0 et pour ceux plus petits que t_0 .

Indication : Pour traiter le cas $t > t_0$, utilisez le lemme de Gronwall 4.1.1.1 et la solution explicite de l'équation

$$\dot{y} = ay + b.$$

Pour traiter le cas, $t < t_0$, considérez la fonction $\bar{x}(t) = x(2t_0 - t)$.

4.1.3 Sur et sous-solutions

Signalons enfin les versions scalaires du lemme de Gronwall. Nous les laissons également en exercice.

Lemme 4.1.3.1 [SOUS-SOLUTION] *Soit y une solution de l'équation différentielle $\dot{y} = g(t, y)$ définie sur l'intervalle $I = [t_0, t_1[$. On suppose que g est continue. Soit x une fonction définie sur I à valeurs dans $\mathbb{R}$ telle que*

$$x(t_0) \leq y(t_0) \quad (4.5)$$

$$\forall t \in I, \quad \dot{x}(t) < g(t, x(t)). \quad (4.6)$$

Alors

$$\forall t \in]t_0, t_1], \quad x(t) < y(t).$$

Indication : reprendre la démonstration de Gronwall.

Nous avons un énoncé symétrique

Lemme 4.1.3.2 [SUR-SOLUTION] *Soit y une solution de l'équation différentielle $\dot{y} = g(t, y)$ définie sur l'intervalle $I = [t_0, t_1[$. Soit x une fonction définie sur I à valeurs dans $\mathbb{R}$ telle que*

$$x(t_0) \geq y(t_0) \quad (4.7)$$

$$\forall t \in I, \quad \dot{x}(t) > g(t, x(t)). \quad (4.8)$$

Alors

$$\forall t \in]t_0, t_1], \quad x(t) > y(t).$$

Indication : appliquer le lemme précédent à $-x$ et $-y$.

EXERCICE :

Donnez des énoncés de sur et sous-solutions pour $t < t_0$.

4.2 Contrôle de la durée de vie des solutions

Le lemme de Gronwall combiné au lemme de sortie des compacts 3.2.2.2 permet de contrôler la durée de vie des solutions dans de nombreux cas. Nous allons en donner deux exemples classiques. A nouveau, il faut plus se souvenir de la méthode que du résultat.

4.2.1 Utilisation de Gronwall pour montrer qu'une durée de vie est infinie

Lemme 4.2.1.1 *Soit f une fonction définie sur $I \times E$. On suppose qu'il existe des fonctions continues a et b définies sur I*

$$\|f(t, x)\| < a(t)\|x\| + b(t).$$

Alors les solutions maximales de

$$\dot{x} = f(t, x),$$

sont définies sur I tout entier.

REMARQUE : Le même résultat – avec la même preuve – est vrai lorsque l'on suppose $\|f(t, x)\| < g(t, x)$ si l'on sait que les solutions de $\dot{y} = g(t, y)$ sont définies sur I tout entier.

DÉMONSTRATION : Soit $(t_0, x_0) \in I \times E$. Soit $I =]\alpha, \beta[$. On considère l'équation scalaire affine $\dot{y} = a(t)y + b(t)$. Nous savons résoudre explicitement cette équation. Ses solutions maximales sont définies pour tout temps. Soit y la solution de conditions initiales $(\|x_0\|, t_0)$.

Soit maintenant x une solution maximale de l'équation (4.2.1.1) définie sur l'intervalle $J =]t^-, t^+[$ avec conditions initiales (t_0, x_0) . D'après le lemme de Gronwall classique on a

$$\forall t > t_0, \|x(t)\| \leq y(t).$$

En particulier, si $t^+ < \beta$

$$\lim_{t \rightarrow t^+} \|x(t)\| \leq y(t^+) < +\infty.$$

Ainsi,

$$R = \sup_{t \in [t_0, t_+]} \|x(t)\| < \infty.$$

Autrement dit la solution maximale ne sort pas – vers le futur – de la boule centrée en 0 et de rayon R . Nous obtenons la contradiction en utilisant le lemme 3.2.2.2.

Un raisonnement analogue montre que $t_- = -\alpha$. $\square$

EXERCICE : : complétez la preuve pour les temps plus petits que t_0 .

Corollaire 4.2.1.2 *Soit f une fonction bornée sur $\mathbb{R} \times E$. Alors les solutions maximales de*

$$\dot{x} = f(t, x),$$

sont définies sur $\mathbb{R}$ tout entier.

4.2.2 Utilisation de Gronwall pour montrer qu'une durée de vie est finie

Nous allons traiter un exemple sous la forme d'un exercice rapide. On considère le champ de vecteur $(2y^2 - 1, 1 - 2x^2)$ sur $\mathbb{R}^2$ associé au système différentiel

$$\begin{cases} \dot{x} &= 2y^2 - 1 \\ \dot{y} &= -1 - 2x^2. \end{cases}$$

Nous allons montrer que les solutions de conditions initiales $(t_0, (x_0, y_0))$ avec $x_0 \neq y_0$ explosent en un temps fini. On considère la fonction $v = x - y$. Pour une solution on a $\dot{v} = 2x^2 + 2y^2$. On remarque que l'on a l'inégalité

$$\dot{v} \geq v^2.$$

Or nous savons que toute solution v_0 de l'équation différentielle

$$\dot{v} = v^2, \tag{4.9}$$

explose en un temps fini t^+ :

$$\lim_{t \rightarrow t^+} v_0(t) = +\infty.$$

Raisonnons par l'absurde et supposons qu'il existe une solution (x, y) définie sur $\mathbb{R}$ tout entier. Alors en particulier, pour tout $t_1 < \infty$, nous avons

$$\lim_{t \rightarrow t_1} v(t_1) < \infty.$$

Comme d'après le lemme des sur-solutions, on a

$$v_0(t) \leq v(t),$$

pour v_0 solution de l'équation (4.9) avec condition initiale $v(t_0) = (x_0, y_0)$, nous obtenons la contradiction.

EXERCICE : Que se passe-t-il si $x_0 = y_0$?

Résumons la méthode que nous avons utilisée :

- nous avons choisi une fonction continue V sur E (dans notre cas $V = x - y$),
- nous avons montré que si $x(t)$ est une solution alors $v(t) = V(x(t))$ vérifie une inéquation différentielle

$$\dot{v} \geq h(t, v),$$

telle que la solution w de

$$\dot{w} = h(t, w),$$

de condition initiale $(t_0, V(x(t_0)))$ explose en un temps fini.

Nous en déduisons que v – et donc x par continuité de V – explose elle aussi en un temps fini inférieur au temps d'explosion de w . En particulier, la durée de vie de x est finie et plus courte que celle de w .

Cette méthode s'applique dans de nombreux cas.

4.3 Contrôle de la divergence de solutions

Nous allons pouvoir contrôler la façon dont deux solutions de conditions initiales proches se comportent au bout d'un temps fini. Nous nous donnons comme d'habitude l'équation différentielle

$$\dot{x} = f(t, x).$$

On suppose que f est k -lipschitzienne et définie sur $I \times \Omega$ où que E est dimension fini et f est C^1 . Notre but est de montrer le lemme suivant.

Lemme 4.3.0.1 [CONTRÔLE DES SOLUTIONS] *Soit x une solution définie sur l'intervalle $[a, b]$. Soit $\beta > 0$.*

Il existe alors $\varepsilon > 0$, tel que si $t_0 \in [a, b]$ et si (x_1, t_1) vérifie

$$\|x_1 - x(t_0)\| \leq \varepsilon, \tag{4.10}$$

$$t_1 \in [a, b] \text{ et } |t_1 - t_0| \leq \varepsilon, \tag{4.11}$$

alors

- la solution y de conditions initiales (x_1, t_1) est également définie sur l'intervalle $[a, b]$,
- de plus, les solutions x et y sont proches :

$$\forall t \in [a, b], \quad \|x(t) - y(t)\| \leq \beta.$$

REMARQUE :

1. Nous obtenons lors de la preuve des constantes explicites.
2. Bien sûr plus $[a, b]$ grandit, plus ε devient petit : ce qui signifie que deux solutions de conditions initiales proches peuvent toujours s'éloigner au bout d'un certain temps. Trouvez un exemple simple de ce type de phénomène !

DÉMONSTRATION : Soit enfin y une solution de conditions initiales (x_1, t_1) . On note par J son intervalle de définition et l'on suppose que $J \subset [a, b]$. Soit

Une première inégalité : Nous allons montrer tout d'abord l'inégalité

$$\forall t \in J, \quad \|x(t) - y(t)\| \leq \|x(t_1) - y(t_1)\| e^{k(b-a)} \quad (4.12)$$

Cela suit immédiatement de la version de Gronwall classique et du fait que f est k -lipschitzienne. Posons en effet, $z(t) = x(t) - y(t)$ sur J . Comme f est k -lipschitzienne, nous avons, pour tout t de J :

$$\begin{aligned} \|\dot{z}(t)\| &= \|\dot{x}(t) - \dot{y}(t)\| \\ &= \|f(t, x(t)) - f(t, y(t))\| \\ &\leq k\|x(t) - y(t)\| = k\|z(t)\|. \end{aligned}$$

Par le lemme de Gronwall classique 4.1.2.1, nous avons pour tout t de J

$$\|z(t)\| \leq \|z(t_1)\| e^{k|t-t_1|}.$$

Nous allons affiner cette inégalités

Quelques constantes : Introduisons quelques constantes numériques dépendant de notre problème.

1. Pour tout r , soit K_r le compact défini par

$$K_r = \{y \in \Omega / \exists t : t \in [a, b], \quad \|y - x(t)\| \leq r\}.$$

Soit donc $R > 0$ tel que

$$K_R \subset \Omega$$

(Montrez en exercice qu'un tel R existe)

2. Soit

$$\varepsilon_1 = \frac{1}{4e^{k(b-a)}} \inf(\beta, R).$$

3. Soit ε_2 tel que

$$|s - t_0| \leq \varepsilon_2 \implies \|x(s) - x(t_0)\| \leq \varepsilon_1.$$

L'existence de ε_2 vient de la continuité de x . Le lecteur attentif et connaissant l'inégalité des accroissements finis vérifiera qu'on peut prendre

$$\varepsilon_2 = \frac{\varepsilon_1}{\sup_{s \in [a, b]} (\|f(x(s), s)\|)}.$$

4. Posons enfin

$$\varepsilon = \inf(\varepsilon_1, \varepsilon_2).$$

Une meilleure inégalité : Nous déduisons de notre première étape l'inégalité suivante : si y est définie sur l'intervalle J de conditions initiales (t_1, x_1) avec

$$\|x(t_0) - x_1\| \leq \varepsilon \text{ et } |t_0 - t_1| \leq \varepsilon$$

alors

$$\forall t \in J, \quad \|x(t) - y(t)\| \leq \inf\left(\frac{\beta}{2}, \frac{R}{2}\right) \quad (4.13)$$

En effet

$$\|x(t_1) - y(t_1)\| \leq \|x(t_0) - y(t_1)\| + \|x(t_0) - x(t_1)\| \leq \varepsilon + \varepsilon_2 \leq 2\varepsilon_1.$$

Enfin, par construction on a bien

$$2\varepsilon_1 e^{k(b-a)} \leq \inf\left(\frac{\beta}{2}, \frac{R}{2}\right).$$

L'inégalité 4.13 suit donc de l'inégalité 4.12.

Durée de vie de y : Montrons maintenant que y est définie sur $[a, b]$. Nous savons d'après de l'inégalité 4.13, que pour tout t de J . On a

$$y(t) \in K_R.$$

Comme K_R est un compact de Ω , nous pouvons par le lemme 3.2.2.1 prolonger y sur $[a, b]$ tout entier.

Conclusion : L'inégalité (4.13) nous dit alors

$$\forall t \in [a, b], \quad \|x(t) - y(t)\| \leq \frac{\beta}{2}.$$

C'est ce que nous voulions démontrer. $\square$

Chapitre 5

Le flot d'une équation différentielle

Nous considérons toujours l'équation

$$\dot{x} = f(t, x).$$

où f est définie sur $I \times \Omega \subset \mathbb{R} \times E$. Nous voulons démontrer que les solutions dépendent continûment des conditions initiales. De manière pour le moment imprécise sur les domaines de définition, nous voulons montrer que l'application qui à (t_0, x_0, t) associe la valeur $x(t)$ de la solution au temps t avec condition initiale (t_0, x_0) est continue.

Nous allons introduire pour cela une notion fondamentale, celle de *flot*, qui généralise celle de résolvante.

5.1 Flot et dépendance continue par rapport aux conditions initiales

Énonçons maintenant le théorème fondamental de ce paragraphe et la définition du flot. Pour tout (t_0, x_0) , nous noterons I_{t_0, x_0} l'intervalle de définition de la solution maximale x de condition initiale (t_0, x_0) .

Nous noterons alors

$$\varphi_{t_0}^t(x_0) = x(t).$$

Théorème 5.1.0.2 *L'ensemble*

$$U = \{(t_0, x_0, t) \in I \times \Omega \times \mathbb{R} / t \in I_{t_0, x_0}\},$$

est ouvert. L'application

$$\left\{ \begin{array}{ll} U & \rightarrow \Omega \\ (x, t, s) & \mapsto \varphi_t^s(x) \end{array} \right. ,$$

est continue.

Nous appellerons φ_{t_0} le *flot* de l'équation différentielle. L'ouvert U est le *domaine de définition du flot*. La deuxième partie du théorème 5.1.0.2 affirme donc que le flot est continu.

Nous montrerons également la proposition

Proposition 5.1.0.3 *Nous avons, pourvu que chacun des termes soit bien défini,*

$$\varphi_t^s \circ \varphi_s^u = \varphi_t^u, \quad \varphi_t^t = \text{Id}$$

De plus si l'équation est autonome on a,

$$\varphi_t^s = \varphi_0^{t-s}, \quad \varphi_0^t \circ \varphi_0^s = \varphi_0^{t+s}.$$

DÉMONSTRATION : la première partie de la proposition utilise la remarque suivante. Si $t \rightarrow x(t)$ est la solution de conditions initiales (u, x_0) c'est aussi - par unicité - la solution de conditions initiales $(s, x(s))$ Autrement dit on a

$$\varphi_u^t(x_0) = x(t) = \varphi_s^t(x(s)) = \varphi_s^t \circ \varphi_u^s(x_0),$$

ce qui démontre la première partie de la proposition.

Si maintenant l'équation est autonome et si $t \rightarrow x(t)$ est la solution de conditions initiales (u, x_0) , alors $y : t \rightarrow x(t+u)$ est solution de condition initiale $(0, x_0)$, on a donc bien

$$\varphi_0^{t-u}(x_0) = y(t-u) = x(t) = \varphi_u^t(x_0).$$

La dernière équation suit alors de la toute première. $\square$

Démonstration du théorème

Montrons tout d'abord que U est ouvert. Soit (t_0, x_0, s_0) in U . On supposera que $s_0 \geq t_0$, l'autre cas se traitant de manière symétrique. Par définition il existe une solution $t \rightarrow x(t)$ de conditions initiales (t_0, x_0) définies sur l'intervalle $[t_0, s_0]$. Nous pouvons alors prolonger légèrement la solution et la supposer définie sur $[t_0, s_0 + \alpha]$. Appliquons le lemme 4.3.0.1 avec $[a, b] = [t_0 - \alpha, s_0 + \alpha]$ il existe donc $\varepsilon > 0$ telle que la solution de conditions initiales (x_1, t_1) est définie sur $[t_0 - \alpha, s_0 + \alpha]$ si

$$\|x_1 - x(t_0)\| \leq \varepsilon, \tag{5.1}$$

$$t_1 \in [t_0 - \alpha, t_1 + \alpha] \text{ et } |t_1 - t_0| \leq \varepsilon. \tag{5.2}$$

Autrement dit si

$$\sup(\|x_1 - x(t_0)\|, |t_1 - t_0|, |t_1 - t_0|) \leq \varepsilon_1 = \inf(\varepsilon, \alpha).$$

Alors la solution de conditions initiales (x_1, t_1) est définie sur $[t_1, s_1]$. C'est-à-dire $(x_1, t_1, s_1) \in U$. L'ensemble U est donc bien ouvert.

Par ailleurs quelque soit β , pour $\varepsilon \leq \varepsilon_2$ suffisamment petit on a, toujours d'après le lemme 4.3.0.1.

$$\|x(s_1) - y(s_1)\| \leq \frac{\beta}{2}.$$

Par continuité de x , il existe enfin η , tel que

$$|s_1 - t_1| \leq \eta \implies \|x(s_1) - y(s_1)\| \leq \frac{\beta}{2}.$$

Ainsi

$$\sup(\|x_1 - x(t_0)\|, |t_1 - t_0|, |t_1 - t_0|) \leq \inf(\varepsilon_1, \varepsilon_2),$$

implique

$$\|x(t_1) - y(s_1)\| \leq \|x(t_1) - x(s_1)\| + \|x(s_1) - y(s_1)\| \leq \beta.$$

Nous venons de montrer la continuité de

$$(x, t, s) \rightarrow \varphi_t^s(x).$$

□

5.2 Dépendance par rapport à un paramètre

Nous nous intéressons à une équation dépendant d'un paramètre :

$$\dot{x} = f_\lambda(t, x).$$

Où

$$f : (\lambda, t, x) \rightarrow f_\lambda(t, x),$$

est définie sur un ouvert $U \times I \times \Omega$. Nous allons montrer

Théorème 5.2.0.4 *On suppose que f est continue par rapport à toutes les variables et qu'elle vérifie l'une des deux hypothèses suivantes*

- f lipschitzienne par rapport à la variable (λ, x) à t fixé.
- E est dimension fini et f est C^1 .

Il existe alors un ouvert $W \subset F \times \mathbb{R} \times E \times \mathbb{R}$, et une application continue

$$\Phi : \begin{cases} W & \rightarrow E \\ (\lambda, x, s, t) & \mapsto \Phi(\lambda, x, s, t) \end{cases},$$

telle que

$$t \mapsto \Phi(\lambda, x, s, t)$$

est la solution maximale de l'équation

$$\dot{x} = f_\lambda(t, x).$$

REMARQUES :

1. Il existe un théorème plus fort qui demande seulement que f_λ soit lipschitzienne par rapport à x . Il se démontre en utilisant un théorème de Picard avec paramètres.
2. Cet énoncé dit que les solutions (pendant un temps fini) dépendent continûment des paramètres.

DÉMONSTRATION : Nous nous ramenons immédiatement au cas de la dépendance continues des conditions initiales, en considérant l'équation

$$\begin{cases} \dot{x} &= f_\lambda(t, x) \\ \dot{\lambda} &= 0; \end{cases}$$

□

5.3 Application : Gronwall avec des inégalités larges

Lemme 5.3.0.5 [GRONWALL AVEC INÉGALITÉ LARGE] *Soit y une solution de l'équation différentielle $\dot{y} = g(t, y)$ définie sur l'intervalle $I = [t_0, t_1[$. On suppose que g est lipschitzienne à t fixé. Soit x une courbe dérivable dans E définie sur I telle que*

$$\begin{aligned} \|x(t_0)\| &\leq y(t_0) \\ \forall t \in I, \quad \|\dot{x}(t)\| &\leq g(t, \|x(t)\|). \end{aligned}$$

Alors

$$\forall t \in [t_0, t_1], \quad \|x(t)\| \leq y(t).$$

Indication : Utilisez le lemme de Gronwall avec la fonction $g_\varepsilon = g + \varepsilon$, puis faites tendre ε vers 0, en utilisant le théorème sur la dépendance continue par rapport aux paramètres. On remarque que l'on a besoin d'une hypothèse de régularité plus forte sur g que pour Gronwall strict.

On peut d'ailleurs remarquer en exercice que cet énoncé contient l'unicité de la solution de l'équation $\dot{y} = g(t, y)$. On a donc besoin d'une hypothèse précise sur g meilleure que la continuité.

Nous avons les énoncés correspondant pour les sous et sur solutions.

Lemme 5.3.0.6 [SOUS-SOLUTION LARGE] *Soit y une solution de l'équation différentielle $\dot{y} = g(t, y)$ définie sur l'intervalle $I = [t_0, t_1[$. On suppose que g est lipschitzienne à t fixé. Soit x une fonction définie sur I à valeurs dans $\mathbb{R}$ telle que*

$$x(t_0) \leq y(t_0) \tag{5.3}$$

$$\forall t \in I, \quad \dot{x}(t) \leq g(t, x(t)). \tag{5.4}$$

Alors

$$\forall t \in [t_0, t_1], \quad x(t) \leq y(t).$$

Lemme 5.3.0.7 [SUR-SOLUTION LARGE] *Soit y une solution de l'équation différentielle $\dot{y} = g(t, y)$ définie sur l'intervalle $I = [t_0, t_1[$. On suppose que g est lipschitzienne à t fixé. Soit x une fonction définie sur I à valeurs dans $\mathbb{R}$ telle que*

$$x(t_0) \geq y(t_0) \quad (5.5)$$

$$\forall t \in I, \quad \dot{x}(t) \leq g(t, x(t)). \quad (5.6)$$

Alors

$$\forall t \in [t_0, t_1], \quad x(t) \leq y(t).$$

5.4 Complément : Le flot comme solution d'une équation différentielle

Rappelons que la résolvante elle-même est la solution d'une équation différentielle sur espace construit à partir de E , à savoir $\text{End}(E)$ l'ensemble des endomorphismes de E . Nous allons procéder de la même manière pour le flot, sous la forme d'un exercice. Dans la section suivante, nous élaborerons sur une situation analogue.

- Soit donc K un fermé de E de Ω . Nous introduisons les objets suivants
- L'espace $\mathcal{E}$ des applications continues bornées de K dans E

$$\mathcal{E} = C^0(K, E).$$

Nous rappelons que $\mathcal{E}$ est un espace vectoriel normé complet muni de la norme

$$\|X\|_{C^0} = \sup_{x \in K} \|X(x)\|.$$

- Le sous-ensemble $\mathcal{O}$ des applications continues de K à valeurs dans Ω

$$\mathcal{O} = C^0(K, \Omega).$$

On vérifie aisément que l'ensemble $\mathcal{O}$ est un ouvert de $\mathcal{E}^k$.

- L'application F de $I \times \mathcal{O}$ dans $\mathcal{E}$ qui à (t, X) associe

$$\tilde{F}(t, X) : \begin{cases} K & \rightarrow \mathcal{E} \\ y & \mapsto f(t, X(y)) \end{cases}.$$

EXERCICE :

1. L'application F est continue et de plus, pour tout t de I ,

$$Z \rightarrow F(t, Z),$$

est k -lipschitzienne.

2. En déduire que l'équation différentielle

$$\dot{X} = F(t, X),$$

admet une solution X_{t_0} avec conditions initiale (t_0, Id) .

3. Montrez alors que pour tout x_0 de K , $t \rightarrow X_{t_0}(t)(x_0)$ est solution de $\dot{x} = f(t, x)$ avec condition initiale (t_0, x_0) . Quelle est la relation entre X_{t_0} et le flot ?
4. Méditez l'analogie entre ce que vous venez de faire et la construction de la résolvante d'une équation linéaire.

5.5 Dépendance C^k par rapport aux conditions initiales

Nous allons démontrer le résultat suivant qui affine le théorème 5.1.0.2 sur la dépendance continue des conditions initiales.

Théorème 5.5.0.8 *Soit f une fonction de classe C^k définie sur $I \times \Omega$ avec $k \geq 1$. Alors, le flot $\varphi_{t_0}^t$ de l'équation différentielle*

$$\dot{x} = f(t, x),$$

est de classe C^k .

Une proposition préliminaire locale

La première partie du raisonnement est locale en temps et en espace. Nous allons montrer

Proposition 5.5.0.9 *Pour tout x_0 , pour tout t_0 , il existe un voisinage $]a, b[$ de t_0 un voisinage K de x , tel que si $t, s \in]a, b[$ alors φ_s^t est de classe C^k sur K .*

DÉMONSTRATION : Nous allons raisonner comme dans la section précédente. Soit donc K un voisinage fermé inclus dans Ω sur lequel les dérivées de f jusqu'à l'ordre k sont bornées. Nous introduisons les objets suivants

- L'espace $\mathcal{E}^k$ des applications de K dans E de classe C^k et dont toutes les dérivées jusqu'à l'ordre k sont bornées.

$$\mathcal{E}^k = C^k(K, E).$$

Nous rappelons que $\mathcal{E}$ est un espace vectoriel normé complet muni de la norme

$$\|X\|_{C^k} = \sup_{x \in K, 0 \leq p \leq k} \|D^p X(x)\|.$$

On a noté $D^p X$ la dérivée p -ième de l'application X .

- L'ensemble $\mathcal{O}$ des applications continues de K à valeurs dans Ω

$$\mathcal{O} = C^0(K, \Omega).$$

On vérifie aisément que l'ensemble $\mathcal{O}$ est un ouvert de $\mathcal{E}^k$.

- L'application F de $I \times \mathcal{O}$ dans $\mathcal{E}^k$ qui à (t, X) associe

$$F(t, X) : \begin{cases} K & \rightarrow \mathcal{E}^k \\ y & \mapsto f(t, X(y)). \end{cases}$$

L'application F est bien définie et de plus, pour tout t de I ,

$$Z \rightarrow F(t, Z),$$

est A -lipschitzienne pour une certaine constante A indépendante de t . Montrons, ce dernier point en détail. Tout d'abord, si h est définie de K dans E , si x et y sont de classe C^p à valeurs dans K , nous avons par la formule de dérivée des compositions et une petite récurrence que

$$\|D^p(h \circ x - h \circ y)\| = 2^p \sup_{q \leq p} \|D^q h\| \cdot \|D^{p-q} x - D^{p-q} y\|. \quad (5.7)$$

Nous en déduisons que F est effectivement A -lipschitzienne pour

$$A = 2^k \sup_{t \in I} \|D^k f(t, \cdot)\|_{C^k}.$$

Par le théorème de Cauchy-Lipschitz 3.1.2.1, il existe un voisinage $]a, b[$ de t_0 , telle que l'équation différentielle

$$\dot{X} = F(t, X),$$

admet pour tout s de $]a, b[$ une solution $X_s : t \rightarrow X_{t_0}^t$ avec conditions initiales (s, Id) définie sur l'intervalle $]a, b[$.

Nous allons maintenant montrer que X_s^t coïncide avec le flot. Par définition X_s^t est une application de K dans Ω . Posons alors

$$x_s(t) = X_s^t(x_0).$$

On a alors

$$\begin{aligned} x_s(s) &= X_s^s(x_0) = \text{Id}(x_0) = x_0. \\ \dot{x}_s(t) &= \dot{X}_s^t(x_0) = F(t, X_s^t(x_0)) = f(t, X_s^t(x_0)) = f(t, x_s(t)). \end{aligned}$$

Ainsi, $x_s(t)$ est la solution de $\dot{x} = f(t, x)$ de conditions initiales (s, x_0) définie sur l'intervalle $]a, b[$. Autrement dit, pour $t \in]a, b[$, nous avons $\varphi_s^t = X_s^t$ sur K .

Ainsi pour tout x_0 , pour tout t_0 , il existe un voisinage $]a, b[$ de t_0 un voisinage K de x , tel que si $t, s \in]a, b[$ alors φ_s^t est de classe C^k sur K . $\square$

Fin de la démonstration du théorème 5.5.0.8

Nous allons maintenant conclure : soit $s \rightarrow x(s)$ la solution de condition initiale (t_0, x_0) de $\dot{x} = f(t, x)$. On suppose qu'elle est définie sur l'intervalle $[t_0, t_1]$. Nous voulons démontrer qu'il existe un voisinage K de x , tel que $\varphi_{t_0}^{t_1}$ est de classe C^k sur F .

D'après ce qui précède pour tout $s \leq t_1$, il existe un intervalle ouvert U_s contenant s , un voisinage K_s de $x(t)$, tel que si $t, u \in U_s$ alors φ_s^u est de classe C^k sur K_s .

Recouvrons l'intervalle compact $[t_0, t_1]$, par une union finie de ces voisinages $U_i = U_{s_i}$ de telle sorte que

- $U_i \cap U_{i+1} \neq \emptyset$,
- $t_0 \in U_1$.

Posons

$$V_p = \bigcup_{i=1}^{i=p} U_i.$$

Nous allons montrer par récurrence sur p , qu'il existe un voisinage W_p de x_0 tel que pour tout t dans V_p , $\varphi_{t_0}^t$ est de classe C^k sur W_p .

Par hypothèse, ceci est vrai pour $p = 1$. Supposons la propriété ci-dessus vraie pour p .

Soit $u_p \in U_p \cap U_{p+1}$. Nous savons que

1. $\varphi_{t_0}^t$ est de classe C^k sur W_p ,
2. pour s dans U_{p+1} , $\varphi_{u_p}^s$ est de classe C^k sur K_{p+1} .

D'après la proposition 5.1.0.3

$$\varphi_{t_0}^s = \varphi_{u_p}^s \circ \varphi_{t_0}^{u_p},$$

Nous en déduisons que pour tout s de U_{p+1} , $\varphi_{t_0}^s$ est de classe C^k sur

$$W_{p+1} = W_p \cap (\varphi_{t_0}^{u_p})^{-1}(K_{p+1}).$$

ceci est a fortiori vrai pour s dans V_p . Nous avons donc terminé la démonstration de cette récurrence et celle du théorème. $\square$

Chapitre 6

Etude géométrique des équations différentielles

6.1 Equations autonomes en dimension 2

Nous allons expliquer dans ce chapitre une série de remarques géométriques permettant d'obtenir des informations qualitatives sur les équations scalaires, c'est-à-dire les équations de la forme

$$\dot{x} = f(t, x),$$

où f est une fonction définie de $]a, b[\times I$ dans $\mathbb{R}$ et I un intervalle de $\mathbb{R}$. Pour simplifier l'exposition de ce chapitre nous nous restreindrons au cas $I = \mathbb{R}$. Le lecteur est invité à titre d'exercice à trouver les hypothèses et énoncés correspondant au cas général. Dans cette introduction et une dernière section, nous expliquerons également quelques idées géométriques analogues concernant plus généralement les équations autonomes en dimension 2. Rappelons que nous avons vu dans la proposition 1.1.2.2 qu'une équation scalaire définit naturellement un système d'équations autonome en dimension 2 dont les solutions s'expriment en les solutions de l'équation initiale. Ce système s'écrit

$$\begin{cases} dx &= f(y, x) \\ dy &= 1. \end{cases} \quad (6.1)$$

6.1.1 Portrait de phase

Pour les équations autonomes en dimension 2, le portrait de phase est planaire. Rappelons qu'une *orbite* ou *trajectoire* est une courbe solution de l'équation différentielle considérée. Le *portrait de phase* est alors la représentation graphique – dans le plan – d'un certain nombre de trajectoires significatives. Dans le cas de l'équation autonome associée à une équation scalaire, les trajectoires sont des graphes dans le plan des (t, x) par rapport à l'axe des t .

6.1.2 Les trajectoires ne se coupent pas

Voici une première traduction graphique de notre théorème fondamental, celui de Cauchy-Lipschitz : *par tout point du domaine de définition de l'équation différentielle passe une courbe solution et une seule*. En particulier, *deux trajectoires ne peuvent se couper*. Voici deux applications de cette observation.

EXEMPLE :

1. Considérons une équation autonome de la forme

$$\begin{cases} \dot{x} &= x.f(x, y) \\ \dot{y} &= y.g(x, y). \end{cases}$$

Nous remarquons que toute solution de cette équation dont les conditions initiales sont sur les axes restent sur les axes (pourquoi?). Autrement dit les axes sont des réunions d'orbites. Ainsi, toute orbite dont les conditions initiales sont dans un quadrant reste dans ce même quadrant.

2. On se donne une trajectoire périodique non constante c de période T et un ouvert U borné tel que $\partial U = c[0, T]$ d'un champ de vecteurs défini sur $\mathbb{R}^2$. Alors toute solution de conditions initiales dans U est définie sur $] -\infty, +\infty[$ et est à valeurs dans U ; ceci suit du lemme de sortie des compacts 3.2.2.2 et de notre observation. En effet, soit x une solution maximale de condition initiales dans U et définie sur $]a, b[$. Comme pour tout t , $x(t) \notin \partial U$ – les orbites ne se croisent pas – on en déduit par le théorème des valeurs intermédiaires que pour tout t , $x(t) \in U$. On conclut par le lemme de sortie des compacts 3.2.2.2.

En général, les trajectoires jouent donc un rôle de barrières étanches : elles ne peuvent être traversées par d'autres orbites.

6.1.3 Points fixes

Remarquons enfin que les orbites constantes correspondent aux zéros de la fonction définissant le système d'équation autonome. Nous n'en dirons pas plus pour le moment et réservons au prochain chapitre une étude plus fine de ces orbites constantes ou points fixes. Remarquons tout de même le résultat évident suivant

Proposition 6.1.3.1 *Soit x un point de Ω . On suppose que x n'est pas un zéro de X , alors $t \rightarrow \phi_t(x)$ est injective au voisinage de zéro.*

DÉMONSTRATION : en effet, par définition du flot

$$\left. \frac{d}{dt} \right|_{t=0} \phi_t(x) = X(x) \neq 0.$$

Un développement limité permet alors de conclure. $\square$

6.2 Barrières et entonnoirs pour les équations scalaires

Nous nous restreignons dans ce paragraphe aux équations scalaires. Nous nous donnerons donc une fonction f que, pour simplifier, nous supposons définie dans $]a, b[\times \mathbb{R}$ et de classe C^1 .

$$\dot{x} = f(t, x), \quad (6.2)$$

Nous allons décrire d'autres courbes qui vont jouer le rôle de barrières : les trajectoires peuvent les traverser dans un sens mais pas dans un autre. Elles jouent un grand rôle pour décrire pratiquement le portrait de phase.

6.2.1 Barrières

Nous dirons que le graphe d'une fonction α , définie sur $]c, d[\subset]a, b[$ – où que α elle-même – est une *barrière inférieure* si

$$\forall t \in]c, d[, \quad \dot{\alpha}(t) < f(t, \alpha(t)).$$

Autrement dit, une barrière inférieure est le graphe d'une sous-solution. La proposition suivante affirme qu'on ne peut approcher une barrière inférieure par en-dessus. Ce n'est bien sûr qu'une reformulation du lemme de Gronwall pour les sous-solutions.

Proposition 6.2.1.1 *Soit α une barrière inférieure définie sur $]c, d[$. Soit $t \rightarrow x(t)$ une solution de l'équation 6.2 définie sur $]u, v[\subset]c, d[$. Soit $t_0 \in]u, v[$ tel que $\alpha(t_0) = x(t_0)$. Alors*

- Pour tout $t_0 < t < v$, on a $x(t) > \alpha(t)$.
- Pour tout $u < t < t_0$, on a $x(t) < \alpha(t)$.

Nous avons également une notion symétrique, à savoir : le graphe d'une fonction α , définie sur $]c, d[\subset]a, b[$, – où que α elle-même – est une *barrière supérieure* si

$$\forall t \in]c, d[, \quad \dot{\alpha}(t) > f(t, \alpha(t)).$$

Nous observons à nouveau qu'une barrière supérieure est le graphe d'une sur-solution. Nous avons de même

Proposition 6.2.1.2 *Soit α une barrière supérieure définie sur $]c, d[$. Soit $t \rightarrow x(t)$ une solution de l'équation 6.2 définie sur $]u, v[\subset]c, d[$. Soit $t_0 \in]u, v[$ tel que $\alpha(t_0) = x(t_0)$. Alors*

- Pour tout $t_0 < t < v$, on a $x(t) < \alpha(t)$.
- Pour tout $u < t < t_0$, on a $x(t) > \alpha(t)$.

6.2.2 Entonnoirs et entonnoirs étranglés

Un *entonnoir* est un ouvert du plan U , borné vers le haut par une barrière supérieure β et vers le bas par une barrière inférieure α , définies sur le même intervalle $]c, d[$ et telles que $\alpha < \beta$:

$$U = \{(t, x) \mid t \in]c, d[, \alpha(t) < x < \beta(t)\}.$$

Nous avons alors

Lemme 6.2.2.1 [COINCÉS DANS UN ENTONNOIR] *Soit U un entonnoir défini sur $]c, d[$. Soit $(x(t))$ la solution maximale de conditions initiales $(t_0, x_0) \in U$. Alors x est définie sur l'intervalle $[t_0, d[$ et on a*

$$\forall t \in [t_0, d[, (t, x(t)) \in U.$$

DÉMONSTRATION : Soit x une solution maximale de conditions initiales (t_0, x_0) définie sur $]u, v[$. Posons $v_0 = \inf(v, d)$. D'après nos résultats sur les barrières, pour tout $v_0 > t > t_0$, on a

$$x(t) \in U.$$

Supposons $v_0 < d$, alors $x([t_0, v_0])$ reste inclus dans le compact

$$K = \{(t, y) \mid t_0 \leq t \leq v, \alpha(t) \leq y \leq \beta(t)\} \subset]a, b[\times \mathbb{R}.$$

Le lemme de sortie des compacts 3.2.2.2 permet de montrer que $v_0 = d$ et donc de conclure. $\square$

On dira qu'un entonnoir tel que ci-dessus est *étranglé* si de plus

$$\lim_{t \rightarrow d} \alpha = \lim_{t \rightarrow d} \beta = y_0.$$

Les entonnoirs étranglés fonctionnent comme des pièges :

Lemme 6.2.2.2 [PIÉGÉS DANS UN ENTONNOIR] *Soit U un entonnoir étranglé défini sur $]c, d[$. Soit $(x(t))$ la solution maximale de conditions initiales $(t_0, x_0) \in U$. Alors x est définie sur l'intervalle $[t_0, d[$, on a $\forall t \in [t_0, d[, (t, x(t)) \in U$ et de plus*

$$\lim_{t \rightarrow d} x(t) = y_0.$$

DÉMONSTRATION : c'est évident. $\square$

Remarquons que c ou d peuvent très bien être infinis.

6.2.3 Anti-entonnoirs et anti-entonnoirs étranglés

Anti-entonnoirs et anti-entonnoirs étranglés

Un *anti-entonnoir* U est un ouvert du plan, bornée vers le bas par une barrière supérieure β et vers le haut par une barrière inférieure α , définies toutes les deux sur le même intervalle $]c, d[$ et telles que $\alpha < \beta$:

$$U = \{(t, x) \mid t \in]c, d[, \alpha(t) < x < \beta(t)\}.$$

Remarquons qu'un anti entonnoir fonctionne comme un entonnoir vers la passé : si une courbe aboutit en un point de U en t_0 , elle est dans U pour les temps plus petits que t_0 .

Il n'y a bien sûr aucune raison maintenant que toutes les solutions issues de U restent coincées dans l'anti-entonnoir vers le futur. En fait, il existe même toujours des solutions qui sortent de cet anti-entonnoir, soit vers le haut, soit vers le bas comme nous le verrons prochainement. Le théorème suivant nous dit cependant qu'il existe toujours au moins une solution qui reste dans l'anti-entonnoir : c'est celle qui n'a pas réussi à choisir entre sortir vers le haut ou sortir vers le bas.

Théorème 6.2.3.1 [EXISTENCE D'UNE SOLUTION PIÉGÉE] *Soit U un anti-entonnoir défini sur $]c, d[$. Il existe alors une solution x définie sur $]c, d[$ et pour tout $t \in]c, d[$ on a $(t, x(t)) \in U$.*

DÉMONSTRATION : Soit $t_0 \in]c, d[$. On considère φ le flot de notre équation. Autrement dit la courbe

$$t \rightarrow \varphi_{t_0}^t(x),$$

est la solution de l'équation de conditions initiales (t_0, x_0) . On considère

$$\begin{aligned} I &= \{x \in \mathbb{R} \mid \exists s, s > t_0, \varphi_{t_0}^s(x) > \alpha(s)\}. \\ J &= \{y \in \mathbb{R} \mid \exists s, s > t_0, \varphi_{t_0}^s(y) < \beta(s)\}. \end{aligned}$$

PREMIÈRE ÉTAPE : Nous voulons montrer : I et J sont disjoints. Pour cela, il suffit de montrer que

$$\forall x \in I, \quad s > t_0 \implies \varphi_{t_0}^s(x) > \beta(s). \quad (6.3)$$

$$\forall x \in J, \quad s > t_0 \implies \varphi_{t_0}^s(x) < \alpha(s). \quad (6.4)$$

Démontrons l'assertion 6.4, la deuxième assertion ayant une preuve symétrique. Raisonnons par l'absurde et supposons qu'il existe s_1 tel que

$$\varphi_{t_0}^{s_1}(x) < \beta(s).$$

Nous en déduisons par le lemme des sous-solutions que

$$\forall s > s_1, \varphi_{t_0}^s(x) < \beta(s).$$

Nous savons par ailleurs qu'il existe s_2 tel que

$$\varphi_{t_0}^{s_2}(x) > \alpha(s).$$

De la même manière, nous avons

$$\forall s > s_2, \varphi_{t_0}^s(x) > \beta(s).$$

Nous obtenons la contradiction en remarquant que pour $s \geq \sup(s_1, s_2)$, nous avons

$$\varphi_{t_0}^s(x) > \beta(s) \text{ et } \varphi_{t_0}^s(x) < \alpha(s).$$

Nous venons de démontrer les assertions 6.4 et 6.4.

DEUXIÈME ÉTAPE : Nous allons montrer *I et J sont des ouverts non vides.* I et J sont non vides car par le lemme des sur et sous-solutions

$$\alpha(t_0) \in I \text{ et } \beta(t_0) \in J. \quad (6.5)$$

Montrons que I est ouvert. Soit x un élément de I . Il existe alors s tel que $\phi_{t_0}^s(x) > \alpha(s)$. Par continuité, de $x \rightarrow \phi_{t_0}^s(x)$, pour y au voisinage de x , on a encore $\phi_{t_0}^s(y) > \alpha(s)$. L'ensemble I – et par un raisonnement symétrique J – est ouvert.

CONCLUSION : On considère

$$z_0 = \inf\{x \in I \mid x > \beta(t_0)\}.$$

D'après la deuxième étape, z_0 est bien défini et n'appartient ni à I ni à J . Soit $z(t)$ la solution de conditions initiales (t_0, z_0) . D'après ce qui précède

$$\forall s \geq t_0, \quad z(s) \in U.$$

Par un raisonnement désormais classique invoquant le lemme de sortie des compacts, $z(t)$ est définie sur $[t_0, d[$. Par ailleurs un anti-entonnoir est un entonnoir vers le passé et nous en déduisons par le lemme 6.2.2.1 que z est définie sur $]c, t_0]$ et vérifie

$$\forall s \leq t_0, \quad z(s) \in U.$$

Nous venons de démontrer notre résultat car z est bien une solution coincée dans l'anti-entonnoir. $\square$

REMARQUE : Dans le cas où f est définie au voisinage en d , ce théorème découle immédiatement du lemme sur les entonnoirs. Pourquoi ?

Anti-entonnoirs étranglés et dispersifs

On dira à nouveau qu'un anti-entonnoir comme ci-dessus est *étranglé* en y_0 si on a

$$\lim_{t \rightarrow d} \alpha = \lim_{t \rightarrow d} \beta = y_0.$$

Nous dirons qu'il est *dispersif* si de plus au voisinage de (d, y_0) on a

$$\partial_x f(t, x) > 0.$$

On remarquera que f n'est pas nécessairement définie en (d, y_0) . Nous avons alors le résultat suivant particulièrement intéressant dans ce dernier cas.

Théorème 6.2.3.2 [UNICITÉ D'UNE SOLUTION PIÉGÉE] *Soit U un anti-entonnoir défini sur $]c, d[$ étranglé et dispersif. Soit $t_0 \in]c, d[$. Il existe alors un unique x_0 tel que si $x(t)$ est la solution de condition initiale (t_0, x_0) alors x est définie sur $]t_0, d[$ et pour tout $t \in [t_0, d[$ on a $(t, x(t)) \in U$. On a alors*

$$\lim_{t \rightarrow d} x(t) = y_0.$$

DÉMONSTRATION : Tout d'abord comme l'anti-entonnoir est étranglé, nous remarquons qu'une solution x de l'équation différentielle de conditions initiales (t_0, x_0) restant dans l'anti-entonnoir doit vérifier

$$\lim_{t \rightarrow d} x(t) = y_0.$$

Soit donc maintenant une autre solution $\bar{x}$ de condition initiale $(t_0, \bar{x}_0)$ restant également dans l'anti-entonnoir. Nous en déduisons donc

$$\lim_{t \rightarrow d} x(t) - \bar{x}(t) = 0. \quad (6.6)$$

Nous raisonnons par l'absurde et supposons que $x_0 < \bar{x}_0$. Les trajectoires ne pouvant se croiser on a donc

$$\forall t > t_0, x(t) < \bar{x}(t).$$

Mais par ailleurs, pour tout t et s avec $t < s$ de $]t_0, d[$ on a

$$\bar{x}(t) - x(t) = \bar{x}(s) - x(s) + \int_t^s f(u, \bar{x}(u)) du - \int_t^s f(u, x(u)) du. \quad (6.7)$$

Pour t dans $]c, d[$, posons

$$V_t = U \cap]t, d[\times \mathbb{R}.$$

Comme l'anti-entonnoir est étranglé et dispersif, pour t_1 suffisamment proche de d , V_{t_1} est inclus dans l'ouvert des points (s, y) pour lequel $\partial_y f(s, y) \geq 0$. Nous en déduisons que si $u > t_1$ on a

$$f(u, \bar{x}(u)) - f(u, x(u)) = \int_{x(u)}^{\bar{x}(u)} \partial_y f(u, v) dv \geq 0.$$

Combinant ceci avec l'équation (6.7) nous en déduisons que pour $s > t > t_1$, on a

$$\bar{x}(t) - x(t) \geq \bar{x}(s) - x(s).$$

Mais ceci est incompatible avec l'équation (6.6). Nous obtenons une contradiction et le résultat suit. $\square$

Futur et passé

Nous n'avons considéré dans ce paragraphe que les barrières supérieures, inférieures, entonnoirs et anti-entonnoirs vers le futur. Les mêmes notions et les résultats analogues existent vers le passé. Nous laissons au lecteur le soin de les découvrir et les écrire.

6.3 Equations autonomes

Nous revenons maintenant au cas général des équations autonomes en dimension 2. Une telle équation est définie par un *champ de vecteurs* $F(x, y) = (X(x, y), Y(x, y))$ défini sur un ouvert Ω et à valeurs dans $\mathbb{R}^2$. Le système correspondant est

$$\begin{cases} \dot{x} &= X(x, y) \\ \dot{y} &= Y(x, y). \end{cases}$$

6.3.1 Rappels sur les courbes planes

Nous allons rappeler brièvement quelques définitions sur les courbes planes.

Ensembles et courbes de niveau

Soit f une fonction définie d'un ouvert de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$. Un point x est un *point régulier* si $d_x f \neq 0$, et un *point critique* sinon. Soit a un nombre réel, l'*ensemble de niveau* a de f est l'ensemble $E_a = f^{-1}\{a\}$. Si le réel a est tel que E_a n'est constitué que de points réguliers, alors a est une *valeur régulière* de f . Dans le cas contraire, a est une *valeur critique*.

Si a , est une valeur régulière, E_a est une *courbe de niveau*. La *droite affine tangente* en un point $z = (x_0, y_0)$ de la courbe de niveau est la droite affine passant par y et parallèle à la droite $\text{Ker}(df(y))$.

Elle est donc d'équation

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

Nous dirons qu'un vecteur X est *transverse* en un point z d'une courbe de niveau, s'il n'est pas parallèle à la droite tangente.

Courbes

Nous dirons en général qu'un ensemble c est une *courbe*, tout point de c possède un voisinage U tel que $U \cap c$ est une courbe de niveau pour une certaine fonction f . On montre alors que la droite tangente est indépendante du choix de cette fonction f .

On admet la proposition suivante

Proposition 6.3.1.1 [COURBES ET GRAPHES] *Un ensemble c est une courbe si et seulement si tout point $z = (x_0, y_0)$ de c possède un voisinage U dans lequel on peut décrire c comme un graphe soit par rapport à l'axe des x , soit par rapport de l'axe des y . Autrement dit, il existe un voisinage U tel que*

- *soit il existe une fonction φ définie sur un voisinage I de x_0 , telle que $c \cap U = \{(x, \varphi(x)) \mid x \in I\}$.*
- *soit il existe une fonction ψ définie sur un voisinage I de y_0 , telle que $c \cap U = \{(\psi(y), y) \mid y \in I\}$.*

De cette proposition, on déduit aisément le corollaire :

Proposition 6.3.1.2 *Soit ξ une application continue et injective définie d'une intervalle $]a, b[$ à valeurs dans $\mathbb{R}^2$. On suppose qu'il existe une courbe ξ telle que $x([a, b]) \subset c$, alors $\xi([a, b])$ est un ouvert de c ; C'est-à-dire il existe un ouvert Ω de $\mathbb{R}^2$ tel que*

$$x([a, b]) = \Omega \cap c.$$

DÉMONSTRATION : C'est un résultat local, nous pouvons donc supposer que c est le graphe par rapport à l'axe des x d'une fonction ϕ . Nous avons donc $\xi(t) = (x(t), \phi(y(t)))$. Nous déduisons du fait que ξ est continue injective que $x(t)$ elle-même est continue injective. En particulier $x([a, b])$ est un intervalle ouvert J de bornes $x(a)$ et $x(b)$. Ainsi

$$\xi([a, b]) = c \cap I \times \mathbb{R}.$$

c'est ce que nous voulions montrer. $\square$

On montre enfin le résultat suivant que nous admettrons également

Proposition 6.3.1.3 *Soit X un vecteur transverse en $z = (x_0, y_0)$ à une courbe c . Soit $\epsilon > 0$. Considérons le segment $I_\epsilon = z +]-\epsilon, \epsilon[X$. Il existe alors $\epsilon > 0$ tel que*

$$I_\epsilon \cap c = \{z\}.$$

6.3.2 Barrières et régionnement

On utilise très souvent en dimension 2 un principe de régionnement : on coupe le plan en zones délimitées par des courbes et on cherche à comprendre comment entrent et sortent les trajectoires de ces zones. Soyons plus précis.

Soit U un ouvert de $\mathbb{R}^2$, dont la frontière $\partial U = \bar{U} \setminus U$ est une réunion disjointe d'un ensemble fini de courbes $c_1, \dots, c_n$ et de points $x_1 \dots x_p$, vérifiant

$$\bar{c}_i \subset c_i \bigcup_i \{x_i\}.$$

On dira que la courbe c_i est une *barrière à l'entrée* pour le champ X si

- X est transverse à c_i ,
- X pointe vers l'extérieur de U : c'est-à-dire pour tout x de c_i , il existe $\epsilon > 0$ tel que $x +]0, \epsilon]X(x) \notin U$.

On dira symériquement que la courbe c_i est une *barrière à la sortie* pour le champ X si

- X est transverse à c_i ,
- X pointe vers l'intérieur de U : c'est-à-dire pour tout x de c_i , il existe $\epsilon > 0$ tel que $x +]0, \epsilon]X(x) \in U$.

EXERCICE : Montrez que si X est transverse à c_i alors soit X pointe vers l'intérieur de U , soit il pointe vers l'extérieur de U .

Nous avons alors

Proposition 6.3.2.1 [RÉGIONNEMENT] *Soit U un ouvert de $\mathbb{R}^2$, dont la frontière $\partial U = \bar{U} \setminus U$ est une réunion disjointe d'un ensemble fini de courbes $c_1, \dots, c_n$ et de points $x_1 \dots x_p$, vérifiant*

$$\bar{c}_i \subset c_i \bigcup_i \{x_i\}.$$

On suppose qu'il existe des ensembles I_e et I_s $I_e \cup I_s = \{1, \dots, n\}$ tels que pour tout i de I_e c_i est une barrière à l'entrée et que pour tout j de I_s c_j est une barrière à la sortie.

Nous avons alors les résultats suivants

1. *Comportement au bord :*

- si $x \in c_i$ avec $i \in I_e$, alors pour t dans un voisinage suffisamment petit de 0, pour $t > 0$ on a $\varphi_t(x) \notin U$, et pour $t < 0$ on a $\varphi_t(x) \in U$.
- Symétriquement : si $x \in c_j$ avec $j \in I_s$, alors pour t dans un voisinage suffisamment petit de 0, pour $t > 0$ on a $\varphi_t(x) \in U$, et pour $t < 0$ on a $\varphi_t(x) \notin U$.

2. *Orbite entrante ou sortante :*

- Soit x un point de U , supposons qu'il existe $t > 0$ tel que $\phi_t(x) \notin U$, alors il existe u , $0 < u < t$ tel que $\phi_u(x) \in \bar{c}_i$ avec $i \in I_e$.
- Symétriquement : soit x un point de U , supposons qu'il existe $t < 0$ tel que $\phi_t(x) \notin U$, alors il existe u , $t < u < 0$ tel que $\phi_u(x) \in \bar{c}_i$ avec $i \in I_s$.

6.3.3 Courbes de niveaux et orbites

Soit donc X un champ de vecteur défini sur un ouvert Ω . Considérons l'équation autonome associée

$$\dot{x} = X(x).$$

Soit φ_t son flot. Nous allons rappeler quelques définitions.

Intégrale première du mouvement

On dit qu'une fonction H définie sur l'ouvert de définition Ω du champ de vecteurs et à valeurs réelles est *invariante* ou que c'est une *intégrale première du mouvement* si, pour tout x la fonction

$$t \rightarrow H \circ \varphi_t(x),$$

est constante. Autrement dit, une intégrale première du mouvement est constante le long des orbites du flot. On remarque aussi

Proposition 6.3.3.1 *L'ensemble $\mathcal{C}$ des points critiques d'une intégrale première est invariant par le flot. Il est donc réunion d'orbites.*

DÉMONSTRATION : en effet

$$d_x H = d_x (H \circ \varphi_t) = d_{\varphi_t(x)} H \circ D_x \varphi_t.$$

Comme $D_x \varphi_t$ est inversible, nous avons

$$d_x H = 0 \implies d_{\varphi_t(x)} H = 0.$$

□

Nous allons examiner ce qui se passe en dehors des points critiques de H .

Courbe de niveau d'une intégrale première

Si H est une intégrale première du champ de vecteurs X , chaque ensemble de niveau est stable par le flot. C'est donc une réunion d'orbites. Dans le cas des courbes de niveaux — c'est à dire en dehors des points critiques de H — la situation s'améliore. Soyons plus précis :

- Soit X un champ de vecteurs défini sur un ouvert Ω de $\mathbb{R}^2$. Soit Z l'ensemble des zéros de X .
- Soit H une intégrale première de X définie sur Ω . Soit $\mathcal{C}$ l'ensemble de ses points critiques.
- Soit $\hat{\Omega} = \Omega \setminus (\mathcal{C} \cup Z)$. D'après les paragraphes précédents, $\hat{\Omega}$ est invariant par le flot.

On remarque alors que $E_c = H^{-1}(c) \cup \hat{\Omega}$ est une courbe : c'est la courbe de niveau c de H restreinte à $\hat{\Omega}$. Nous allons voir maintenant que les composantes connexes de E_c — lorsque c varie — sont alors les orbites de X tracées dans $\hat{\Omega}$, plus précisément :

Proposition 6.3.3.2 *Soit x_0 un point de E_c . Soit $I =]a, b[$, l'intervalle de définition de la solution maximale ξ de conditions initiales $(0, x_0)$. Alors*

1. *l'application $\xi : t \rightarrow x(t)$ est une surjection de I sur une composante connexe γ de E_c .*
2. *si γ est compacte, ξ est une orbite périodique : il existe T tel que pour tout t , $\xi(t + T) = \xi(t)$.*
3. *si γ n'est pas compacte, alors ξ est une bijection.*

Cette proposition ne nous dit rien sur les orbites de X dans l'ensemble $\mathcal{C}$ des points critiques. C'est bien sûr normal : une fonction constante est toujours une intégrale première, elle ne peut apporter aucune information nouvelle sur le flot, ceci est cohérent avec le fait que tout point est point critique de la fonction constante.

DÉMONSTRATION : D'après les propositions 6.3.1.2 et 6.1.3.1, $\varphi_{]a, b[}(x_0)$ est un ouvert de E_c . Par ailleurs, soit z un point de E_c dans l'adhérence de $\xi(I)$. Nous en déduisons par le même raisonnement qu'il existe $\varepsilon > 0$, tel $\varphi_{]-\varepsilon, \varepsilon[}(z)$ est un ouvert de E_c . Par définition de ce qu'est une adhérence, $\varphi_{]a, b[}(x_0)$ et $\varphi_{]-\varepsilon, \varepsilon[}(z)$ ont une intersection non vide. Ainsi, z et x sont sur la même orbite et dès lors $z \in \varphi_I(x)$. Nous venons de montrer que $\varphi_{]a, b[}(x)$ est à la fois ouvert et fermé dans E_c . Comme $\varphi_{]a, b[}(x)$ est connexe, c'est une composante connexe de E_c . Ceci démontre le premier point.

Supposons que ξ ne soit pas injective, alors il existe t et T tel que $\varphi_t + T(x) = \varphi_t(x_0)$. Dès lors d'après les propriétés du flot, dès que ceci à un sens, on a pour tout s , $\varphi_s + T(x_0) = \varphi_s(x_0)$. Nous en déduisons que

$$\varphi_{]a, b[}(x_0) \subset \varphi_{[0, T]}(x_0).$$

Ainsi l'orbite de x_0 est compacte. Ceci démontre le troisième point.

Réciproquement, supposons que l'orbite de x soit compact. D'après le lemme de sortie des compacts $I =]-\infty, +\infty[$. Soit $\{t_n\}_{n \in \mathbb{N}}$ une suite croissante et tendant vers l'infini. Nous pouvons alors supposer – après extraction d'une sous-suite – que $\varphi_{t_n}(x_0)$ converge $\varphi_u(x_0)$. Le même raisonnement que pour le premier point nous dit alors qu'il existe $\varepsilon > 0$, tel $\varphi_{]-\varepsilon+u, \varepsilon+u[}(x_0)$ est un ouvert de E_c . Par définition de l'adhérence ; il existe N tel que $t_N > u + \varepsilon$ et

$$\varphi_{t_N}(x_0) \in \varphi_{]-\varepsilon+u, \varepsilon+u[}(x_0).$$

En particulier, il existe $v < u + \varepsilon$, tel que

$$\varphi_{t_N}(x_0) = \varphi_v(x_0).$$

Ainsi ξ n'est pas injective. Nous avons achevé de démontrer le deuxième point et la démonstration de notre proposition. $\square$

Le cas des systèmes Hamiltoniens

Soi H une fonction définie sur $\mathbb{R}^2$, le système Hamiltonien associé est le système

$$\begin{cases} \dot{x} &= -\partial_y H \\ \dot{y} &= \partial_x H. \end{cases}$$

On remarque – en exercice – que la fonction H elle-même est une intégrale première du mouvement. Par ailleurs, les zéros du champ de vecteur associé à l'équation sont précisément les points critique de H . Nous en déduisons que les portrait de phase de l'équation est exactement le tracé des courbes de niveaux de la fonction H . Le lecteur est invité - sous forme d'exercice – à comprendre comment s'orientent les courbes de niveaux pour donner les orbites du flot.

Chapitre 7

Stabilité des points d'équilibre de champs de vecteurs.

7.1 Champ de vecteur et équations autonomes

Nous étudierons désormais les équations autonomes. Nous nous donnons donc un champ de vecteur X défini d'un ouvert Ω de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^n$. Nous considérons l'équation

$$\dot{x} = X(x).$$

Dans ce cas ci, nous savons que si ϕ_s^t est le flot de l'équation, nous avons

$$\phi_s^t = \phi_0^{t-s}, \quad \phi_0^t \circ \phi_0^s = \phi_0^{t+s}.$$

Nous appellerons alors *flot de l'équation autonome* ou *flot du champ de vecteurs* l'application

$$\varphi_t = \phi_0^t.$$

Nous avons lorsque chacun des termes est défini :

$$\varphi_{t+s} = \varphi_t \circ \varphi_s.$$

Pour conclure remarquons que le flot du champ de vecteur X est caractérisé par

$$\begin{cases} \varphi_0(x) &= x \\ \frac{d}{dt} \varphi_t(x) &= X(\varphi_t(x)). \end{cases}$$

ou sous forme intégrale

$$\phi_t(x) = x + \int_0^t X(\phi_s(x)) ds.$$

7.2 Points d'équilibre

Soit donc X un champ de vecteur et φ son flot.

Point d'équilibre

Un *point d'équilibre* est un point x_0 tel que $X(x_0) = 0$. En particulier,

$$\forall t \in \mathbb{R}, \quad \varphi_t(x_0) = x_0.$$

REMARQUES :

1. Si réciproquement, on $\varphi_t(x_0) = x_0$ pour t variant dans un intervalle ouvert, alors x_0 est un point fixe. Ceci s'obtient en dérivant $\varphi_t(x_0)$ par rapport à t dans cet ouvert.
2. Si on a

$$\lim_{t \rightarrow \infty} \varphi_t(y_0) = x_0,$$

Alors x_0 est un point fixe. En effet, par continuité du flot on a

$$\forall s > 0, \varphi_s(x_0) = \lim_{t \rightarrow \infty} \varphi_s(\varphi_t(y_0) = x_0) = \lim_{t \rightarrow \infty} \varphi_{t+s}(y_0) = x_0.$$

Point d'équilibre stable

Un point d'équilibre est *stable* si, pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que

$$\|x - x_0\| \leq \alpha \text{ et } t \geq 0 \implies \|\varphi_t(x) - x_0\| \leq \varepsilon;$$

En particulier, nous en déduisons que la trajectoire future de tout point de $B(x_0, \delta)$ reste dans un compact et donc que cette trajectoire future est définie sur tout $\mathbb{R}^+$ par le lemme de prolongement (ou de sortie des compacts).

On dit qu'un point d'équilibre est *instable* si il n'est pas stable.

EXERCICE :

1. Montrez que $(0, 0)$ est point d'équilibre stable de $X(x, y) = (-2y, 3x)$. Utilisez pour cela la fonction $H(x, y) = 3x^2 + 2x^2$.
2. Montrez que $(0, 0)$ est un point d'équilibre instable de $X(x, y) = (-2x, 3y)$.

Point d'équilibre asymptotiquement stable

Un point d'équilibre x_0 est *asymptotiquement stable* si il est stable et si de plus il possède un voisinage U tel que

$$\forall x \in U, \quad \lim_{t \rightarrow \infty} \varphi_t(x) = x_0.$$

1. Montrez que $(0, 0)$ n'est pas un point d'équilibre asymptotiquement stable de $X(x, y) = (-2y, 3x)$.
2. Montrez que $(0, 0)$ est un point d'équilibre asymptotiquement stable de $X(x, y) = (-x, -y)$.

Bassin d'attraction

Soit x_0 un point d'équilibre asymptotiquement stable. Le *bassin d'attraction* de x_0 est l'ensemble

$$B = \{y \in \Omega, \lim_{t \rightarrow \infty} \varphi_t(x) = x_0.\}$$

Bien sûr, les points de B ont par définition une trajectoire future définie pour tout temps.

Nous avons alors

Proposition 7.2.0.3 *Le bassin d'attraction d'un point asymptotiquement stable est un ouvert.*

DÉMONSTRATION : Soit donc x_0 un point d'équilibre asymptotiquement stable. En particulier, il existe un ouvert U contenant x_0 tel que pour tout y de U , nous avons

$$\lim_{t \rightarrow \infty} \varphi_t(y) = x_0.$$

Soit maintenant B le bassin d'attraction de x_0 . Soit $z \in B$. Par définition, nous avons

$$\lim_{t \rightarrow \infty} \varphi_t(z) = x_0.$$

Dès lors, il existe t_0 tel que $\varphi_{t_0}(z) \in U$. Par la continuité de φ_{t_0} , il existe un voisinage V de z tel que $\varphi_{t_0}(V) \subset U$. Soit maintenant $u \in V$. Nous avons alors

$$\lim_{t \rightarrow \infty} \varphi_t(u) = \lim_{s \rightarrow \infty} \varphi_{s+t_0}(u) = \lim_{s \rightarrow \infty} \varphi_s(\varphi_{t_0}(u)) = x_0.$$

Dès lors, $V \subset B$. Nous venons de démontrer que B est ouvert. $\square$

7.3 Fonction de Liapounov

Soit x_0 un point d'équilibre du champ de vecteur X . Nous dirons qu'une fonction L définie au voisinage U de x_0 est une *fonction de Liapounov* pour X en x_0 si

1. L admet un minimum local strict en x_0 ,
2. Si $x \neq x_0$, alors la fonction $t \rightarrow L \circ \varphi_t$ est strictement décroissante pour $t > 0$ sur l'intervalle sur lequel elle est définie.

Il est utile de remarquer que les conditions suivantes impliquent que L est une fonction de Liapounov

1. L est de classe C^2 ,
2. $D_{x_0}^2 L$ est définie positive,
3. Au voisinage de x_0 , si $x \neq x_0$, $dL(X(x)) < 0$.

Nous avons

Théorème 7.3.0.4 *Si x_0 admet une fonction de Liapounov, alors x_0 est un point d'équilibre asymptotiquement stable.*

DÉMONSTRATION : Nous supposons pour simplifier que $L(x_0) = 0$. Nous savons alors qu'il existe un voisinage W de x_0 tel que

$$x \in W, x \neq x_0 \implies L(x) > 0.$$

Nous supposons que W est compact et inclus dans Ω . Soit $\alpha > 0$, tel que $B(x_0, \alpha) \subset W$. L'ensemble $K = W \setminus B(x_0, \alpha)$ est compact : c'est l'intersection d'un fermé et d'un compact. Soit alors

$$A = \inf\{L(y), y \in K\}.$$

Comme K est compact et ne contient pas x_0 , on a $A > 0$. Nous considérons ensuite

$$U_n = \{x \in B(x_0, \alpha) \mid L(x) < \frac{A}{n}\}.$$

La fonction L étant continue, U_n est ouvert. Par construction, il est inclus dans K

Première étape : durée de vie des solutions. Soit $x \in U_n$. Supposons la solution maximale de conditions initiales $(x, 0)$ définie sur l'intervalle $I =]t_1, t_0[$. Montrons maintenant que $t_0 = +\infty$ et que

$$\forall t > 0, \varphi_t(x) \in U_n \subset W.$$

Supposons tout d'abord $t_0 < \infty$, il existe alors t tel que $\varphi_t(x) \notin K$ par le lemme de sortie des compacts. Par le théorème des valeurs intermédiaires, il existe alors $s > 0$ tel que $\|\varphi_s(x_0)\| = \alpha$. Mais alors, $L(\varphi_s(x_0)) \geq A$. Nous obtenons la contradiction, car L décroît le long des orbites :

$$L(\varphi_s(x)) \leq L(\varphi_0(x)) \leq \frac{A}{n}.$$

Nous venons de montrer la première partie de notre assertion. Nous venons également de démontrer que

$$\forall x \in U_n, \forall s > 0, \varphi_s(x) \in B(x_0, \alpha).$$

En particulier, comme

$$\forall x \in U_n, \forall s > 0, L(\varphi_s(x)) \leq L(x) \leq \frac{A}{n}.$$

Nous venons de montrer

$$\forall s > 0, \varphi_s(U_n) \subset U_n.$$

Deuxième étape : la taille des ensembles U_n tend vers zéro. Posons $\varepsilon_n = \sup_{x \in U_n} \|x - x_0\|$, nous avons

$$\lim_{n \rightarrow \infty} \varepsilon_n = 0.$$

Tout d'abord, il existe une suite de points $\{y_n\}_{n \in \mathbb{N}}$ de W , telle que

$$\|y_n - x_0\| = \varepsilon_n \text{ et } L(y_n) \leq \frac{A}{n}.$$

Il suffit de prendre pour y_n un point qui réalise le maximum de la fonction $z \rightarrow \|z - x_0\|$ dans l'adhérence - compacte - de U_n . Nous pouvons maintenant supposer - après une extraction de sous-suite - que la suite $\{y_n\}_{n \in \mathbb{N}}$ converge vers un point y_0 . Par continuité de L , $L(y_0) = 0$. Comme x_0 est un minimum strict de L sur W , nous avons $y_0 = x_0$. Ceci entraîne bien que

$$\lim_{n \rightarrow \infty} \varepsilon_n = \lim_{n \rightarrow \infty} \|y_n - x_0\| = 0.$$

Troisième étape : x_0 est stable. Nous en déduisons que x_0 est un point d'équilibre stable. En effet soit $\varepsilon > 0$, soit alors n tel que $\varepsilon_n < \varepsilon$. Soit enfin δ tel que

$$B(x_0, \delta) \subset U_n.$$

Alors

$$\|x - x_0\| \leq \delta \implies x \in U_n \implies \forall t > 0, \quad \varphi_t(x) \in U_n \subset B(x_0, \varepsilon).$$

Quatrième étape : x_0 est asymptotiquement stable. Nous savons enfin que pour tout point x de U_1 , $t \rightarrow L \circ \varphi_t(x)$ est décroissante. Soit donc

$$a = \inf_{t \rightarrow \infty} L \circ \varphi_t(x) = \lim_{t \rightarrow \infty} L \circ \varphi_t(x).$$

Supposons maintenant que $\varphi_t(x)$ ne converge pas vers x_0 . Il existe alors, par compacité de W une sous-suite $\{t_n\}_{n \in \mathbb{N}}$ telle que $\{\varphi_{t_n}(y)\}_{n \in \mathbb{N}}$ converge vers un point y_0 différent de x_0 . Par continuité de L , nous avons

$$L(y_0) = a.$$

Comme $y \neq x_0$, par la définition de ce qu'est une fonction de Liapounov, il existe $T > 0$ tel que

$$L \circ \varphi_T(y) = b < a.$$

La continuité du flot nous donne alors

$$L \circ \varphi_T(y) = \lim_{n \rightarrow \infty} L \circ \varphi_{t_n+T}(y) = a,$$

et la contradiction. $\square$

7.4 Stabilité et valeurs propres

Nous allons nous donner un critère permettant de repérer – grâce à une fonction de Liapounov – les points d'équilibre asymptotiquement stables.

Théorème 7.4.0.5 [STABILITÉ ET SPECTRE] *Soit X un champ de vecteur. On suppose que au voisinage de x_0 , on peut écrire le développement limité*

$$X(x) = A \cdot (x - x_0) + X_1(x).$$

où A est linéaire et $\lim_{x \rightarrow x_0} \frac{X_1(x)}{\|x - x_0\|} = 0$. Autrement dit, on suppose que X est différentiable en x_0 et l'on pose $A = D_{x_0}X$.

On suppose de plus que pour toute valeur propre λ de A , on a $\Re(\lambda) < 0$. Dans ces conditions, x_0 admet une fonction de Liapounov et est un point d'équilibre asymptotiquement stable.

DÉMONSTRATION : Nous supposons pour simplifier que $x_0 = 0$. Nous rappelons que d'après le théorème 2.2.0.10, il existe $\alpha > 0$ et $T > 0$ tel que

$$\forall t > T, \quad \|\exp(tA)x\| \leq \|x\|e^{-\alpha t}. \quad (7.1)$$

Nous introduisons une fonction L définie par

$$L(x) = \int_0^\infty \|\exp(sA)x\|^2 ds.$$

D'après l'inégalité 7.1, la fonction L est bien définie. De plus, si $x \neq 0$, on a

$$L(x) \neq 0 = L(0).$$

Donc 0 est un minimum local strict de L . Comme L est une limite de forme quadratiques, c'est elle-même une forme quadratique. Un développement limité nous donne donc

$$L(x + \varepsilon u) = L(x) + 2\varepsilon \int_0^\infty \langle \exp(sA)x, \exp(sA)u \rangle ds + \varepsilon^2 L(u).$$

Nous en déduisons que

$$D_x L(u) = 2\varepsilon \int_0^\infty \langle \exp(sA)x, \exp(sA)u \rangle ds.$$

Nous déduisons tout d'abord de l'inégalité (7.1) qu'il existe C telle que

$$|D_x L(u)| \leq C \|u\| \cdot \|x\|.$$

Par ailleurs,

$$D_x L(A \cdot x) = 2 \int_0^\infty \langle \exp(sA)x, \exp(sA)Ax \rangle ds \quad (7.2)$$

$$= \int_0^\infty \frac{d}{ds} \|\exp(sA)x\|^2 ds \quad (7.3)$$

$$= \lim_{s \rightarrow \infty} (\|\exp(sA)x\|^2 - \|x\|^2) \quad (7.4)$$

$$= -\|x\|^2. \quad (7.5)$$

Revenons au champ X , nous avons alors

$$D_x L(X) = D_x L(A \cdot x) + D_x L(X_1(x)) = \|x\|^2 \left(-1 + \frac{D_x L(X_1(x))}{\|x\|^2} \right).$$

Mais pour x suffisamment petit, nous avons

$$\frac{\|X_1(x)\|}{\|x\|^2} \leq \frac{1}{2C},$$

et donc,

$$\frac{|D_x L(X_1(x))|}{\|x\|^2} \leq \frac{1}{2}.$$

Ainsi

$$d_x L(X(x)) < 0, \text{ pour } x \neq 0.$$

Ceci entraîne que L décroît strictement le long des orbites de X différentes de l'orbite constante x_0 . Nous venons de terminer la preuve du théorème. $\square$

7.5 Un exemple : la dynamique d'un champ de gradient

Nous allons traiter un exemple simple celui du champ de vecteur donné par le gradient d'une fonction f définie sur un ouvert Ω .

$$X(x) = \nabla_x f = (\partial_{x_1} f, \dots, \partial_{x_n} f).$$

Nous allons résumer ces propriétés sous la proposition suivante. Il s'agit plus d'une illustration des techniques que nous avons vues précédemment que d'un résultat à retenir.

Proposition 7.5.0.6 *Soit f une fonction C^∞ définie sur un ouvert Ω . Soit $X = \nabla f$ le champ de gradient de f . Alors*

1. *la fonction f croît le long de toutes les orbites de f , elle croît strictement le long des orbites non constantes.*
2. *Les points d'équilibres de X sont les points critiques de f .*

3. Un point d'équilibre isolé est asymptotiquement stable si et seulement si c'est un maximum local de f .
4. On suppose que pour tout r , l'ensemble $V_r = f^{-1}[c, +\infty[$ est compact. Alors, toute solution maximale est définie sur $[0, +\infty[$.
5. Soit $\{\varphi\}_{t \in \mathbb{R}}$ le flot du gradient de f . Soit $x_0 \in \Omega$. Soit $\{t_n\}_{n \in \mathbb{N}}$ est une suite de réels tendant vers l'infini telle que $\{\varphi_{t_n}\}_{n \in \mathbb{N}}$ converge vers y_0 . Alors a est un point critique de f . Si de plus y_0 est isolé parmi les points critiques alors $\lim_{t \rightarrow \infty} \varphi_t(x) = y_0$.

Signalons que la dynamique d'un flot de gradient est très spéciale : par exemple, elle ne possède aucune orbite périodique.

DÉMONSTRATION : les deux premiers points sont évidents. Du deuxième point, on déduit que si y_0 est un point d'équilibre stable alors c'est un maximum local. De même, si y_0 est un point maximum local alors c'est un point d'équilibre car $-f$ est une fonction de Lyapounov pour le champ de gradient toujours d'après les deux premières propriétés. Le quatrième point suit du lemme de sortie des compacts 3.2.2.2.

Le cinquième point est le plus délicat. La fonction $t \rightarrow f(\varphi_t(x_0))$ est croissante. Par continuité de f , on a

$$f(y_0) = \lim_{t \rightarrow \infty} f(\varphi_t(x_0)).$$

Montrons maintenant que y_0 est un point critique. On sait que $\lim_{n \rightarrow \infty} \varphi_{1+t_n}(x_0) = \varphi_1(y_0)$. Le même raisonnement montre que

$$f(\varphi_1(y_0)) = \lim_{t \rightarrow \infty} f(\varphi_{1+t}(x_0)) = f(y_0).$$

Comme f croît strictement le long des orbites non constantes du gradient, on déduit que y_0 est une orbite constante et donc un point critique de f .

Supposons maintenant que y_0 est isolé parmi les points critiques. Montrons alors

$$\lim_{t \rightarrow \infty} \varphi_t(x) = y_0.$$

Supposons le contraire, il existe alors $\varepsilon > 0$ et une suite $\{s_n\}_{n \in \mathbb{N}}$ tendant vers l'infini telle $s_n \geq t_n$ et

$$d(\varphi_{s_n}(x), y_0) \geq \varepsilon.$$

On peut choisir ε tel que la boule de centre y_0 de rayon ε ne contienne que y_0 comme point critique.

En utilisant le théorème des valeurs intermédiaires, appliqué à la fonction

$$t \rightarrow d(\varphi_t(x), y_0),$$

sur l'intervalle $[t_n, s_n]$, on en déduit l'existence d'une suite $\{u_n\}_{n \in \mathbb{N}}$ tendant vers l'infini telle que

$$d(\varphi_{u_n}(x), y_0) = \frac{\varepsilon}{2}.$$

Par compacité de la boule de centre a de rayon ε , on en déduit qu'après extraction d'une sous-suite on peut supposer qu'il existe b avec $d(b, y_0) = \frac{\varepsilon}{2}$.

$$\lim_{n \rightarrow \infty} \varphi_{u_n}(x) = b \neq a.$$

On obtient par la première partie que b est un point critique d'où la contradiction. $\square$

Chapitre 8

Linéarisation

Nous allons travailler sur une question un peu plus globale : comprendre le comportement au voisinage d'une orbite donnée.

Nous supposons désormais que le champ de vecteur est C^1 . D'après cette hypothèse et le théorème 5.5.0.8, le flot est C^1 .

8.1 La différentielle du flot

8.1.1 Équation linéarisée

On se donne un champ de vecteur X et une orbite $t \rightarrow \gamma(t)$ de X définie sur un intervalle I . Nous considérons l'application

$$A : \begin{cases} I & \rightarrow & \text{End}(E) \\ t & \mapsto & D_{\gamma(t)}X \end{cases}.$$

L'équation linéarisée le long de γ est l'équation linéaire

$$\dot{Y} = A(t)Y.$$

8.1.2 La différentielle

Nous montrons

Théorème 8.1.2.1 [DIFFÉRENTIELLE DU FLOT] *La différentielle du flot φ_t en x est la résolvante R_0^t de l'équation linéarisée le long de la solution $\gamma : s \rightarrow \varphi_s(x)$.*

DÉMONSTRATION : Fixons x , et rappelons, en utilisant la forme intégrale que

$$\varphi_t(x) = x + \int_0^t X(\phi_s(x)) ds.$$

Soit donc ε un réel positif, nous avons alors

$$\varphi_t(x + \varepsilon u) = x + \int_0^t X(\phi_s(x + \varepsilon u)) ds.$$

Nous en déduisons, en dérivant par rapport à ε

$$\begin{aligned} D_x \varphi_t(u) &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \varphi_t(x + \varepsilon u) \\ &= u + \int_0^t \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} X(\varphi_s(x + \varepsilon u)) ds \\ &= u + \int_0^t D_{\varphi_s(x)} X(D_x \varphi_s(u)) ds \end{aligned} \tag{8.1}$$

Autrement dit, en posant $A(s) = D_{\varphi_s(x)} X$

$$\left. \frac{d}{dt} \right|_{t=s} D_x \varphi_t(u) = A(s) D_x \varphi_s(u);$$

Donc, pour tout u , $D_x \varphi_t(u)$ est la solution de l'équation linéarisée le long de γ de condition initiale u . Comme $D_x \varphi_0 = \text{Id}$, nous en déduisons que $D_x \varphi_t = R_0^t$.
 $\square$

Chapitre 9

Difféomorphisme et changement de coordonnées

Pour résoudre une équation différentielle ou plus modestement chercher des informations qualitatives à son sujet, il est souvent utile de chercher les coordonnées dans lesquelles elle s'exprime le plus simplement. Nous allons donner la construction théorique – utilisant la notion de difféomorphisme – et les règles pratiques de cette recherche de bonnes coordonnées.

Nous nous intéresserons à un problème local : pouvons nous écrire, en choisissant correctement les coordonnées, le champ de vecteurs sous une forme canonique ou, suivant la terminologie usuelle, *forme normale*. Nous montrerons dans un premier temps que en dehors d'un point d'équilibre le champ peut être écrit comme un champ constant.

Pour le moment, insistons tout de même sur le fait que connaître une forme normale, c'est à dire l'allure du portrait de phase au voisinage d'un point, n'apporte que des informations très partielles : nous ne savons pas ce qui se passe quand l'orbite sort du voisinage en question, nous ne savons par exemple pas du tout si elle va y revenir.

9.1 Changement de coordonnées

Soit E et F deux espaces vectoriels. Soit O un ouvert de E et U un ouvert de F . Un difféomorphisme C^k de O sur U est une bijection ψ de O sur U de classe C^k et dont l'inverse est également de classe C^k .

Dans la pratique, E et F sont souvent identifiés avec $\mathbb{R}^p$ et un difféomorphisme représente un *changement de coordonnées*. En effet, l'identification de E avec $\mathbb{R}^k$ est donnée par une application linéaire

$$X = (x_1, \dots, x_p),$$

où les x_k sont des *coordonnées cartésiennes* de E . De même, celle de F avec $\mathbb{R}^p$

est donnée par une application linéaire

$$Y = (y_1, \dots, y_p),$$

Le difféomorphisme $\psi = (\psi_1, \dots, \psi_p)$ de O avec U , nous donne alors en plus du système cartésien “nature” de coordonnées sur O à savoir les fonctions $(x_1, \dots, x_p)$, un *nouveau système de coordonnées* à savoir les fonctions $(\psi_1 = y_1 \circ \psi, \dots, \psi_p = y_p \circ \psi)$. Lorsque l’on fait des calculs, on utilise souvent l’abus de langage qui consiste à parler des coordonnées $(y_1, \dots, y_p)$ plutôt que de la notation plus lourde $(y_1 \circ \psi, \dots, y_p \circ \psi)$. Alternativement, on a sur U deux systèmes de coordonnées $(y_1, \dots, y_p)$ et $(x_1 = x_1 \circ \psi^{-1}, \dots, x_p = x_p \circ \psi^{-1})$. Ces deux systèmes de coordonnées (sur U ou O) sont alors reliés par les formules

$$y_i = \psi_i(x_1, \dots, x_p).$$

Donnons un exemple bien connu pour illustrer cette discussion, celui du changement de coordonnées polaire/cartésienne. On considère E et F des espaces vectoriels de dimension 2. On note les (r, θ) coordonnées (cartésiennes !) de E , r et θ sont des fonctions linéaires sur E . On considère $O =]0, \infty[\times]-\pi, \pi[$. On considère ensuite F muni de coordonnées cartésiennes (x, y) et $U = F \setminus \{(x, 0), x \in]-\infty, 0[$. Le difféomorphisme ψ est alors donnée par

$$\psi(r, \theta) = (r \cos(\theta), r \sin(\theta)).$$

Autrement dit on a en utilisant l’abus de langage que nous avons signalé des coordonnées (r, θ) sur U dites *polaires* reliées aux coordonnées cartésiennes (x, y) par les formules

$$x = r \cos(\theta), \quad y = r \sin(\theta).$$

9.2 Image d’un champ de vecteur par un difféomorphisme

On se donne un champ de vecteur $Z = (Z_1, \dots, Z_p)$ défini sur un ouvert O de $\mathbb{R}^p$. On se donne un difféomorphisme ψ de O sur un ouvert U de $\mathbb{R}^p$. L’*image du champ de vecteur Z par ψ* est le champ de vecteur

$$\psi_*(Z) : y \rightarrow D_{\psi^{-1}(y)}\psi(Z(y)).$$

Commençons par montrer que le flot de $\psi_*(Z)$ se comprend très facilement à l’aide de celui de Z . Ceci nous permet de relier les solutions des équations différentielles associées dans les différentes coordonnées

Proposition 9.2.0.2 *Si $\gamma(t)$ est une orbite de Z , alors $\psi(\gamma(t))$ est une orbite de $\psi_*(Z)$. De manière équivalente, si ϕ_t est le flot de Z , alors celui de $\psi_*(Z)$ est $\tilde{\phi}_t = \psi \circ \phi_t \circ \psi^{-1}$.*

DÉMONSTRATION : en effet, nous avons $\tilde{\phi}_0 = \text{id}$. De plus, posons $\gamma(t) = \phi_t \circ \psi^{-1}(x)$ alors

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} \tilde{\phi}_t(x) &= D_{\gamma(0)} \psi \left(\left. \frac{d}{dt} \right|_{t=0} \gamma(t) \right) \\ &= D_{\psi^{-1}(x)} Z(\psi^{-1}(x)) \\ &= \psi_*(Z)(x). \end{aligned}$$

Nous allons maintenant voir que $\psi_*(Z)$ est relié au changement de coordonnées et comment s'écrit l'équation différentielle associée.

Proposition 9.2.0.3 *Posons $x = \psi^{-1}(y_1, \dots, y_p)$. L'équation différentielle associée au champ de vecteur $\psi_*(Z)$ est donnée par*

$$\begin{cases} \dot{y}_1 &= \sum_i \partial_i \psi_1(x) Z_i(x) \\ &\dots \\ \dot{y}_p &= \sum_i \partial_i \psi_p(x) Z_i(x) \end{cases}$$

Ceci est évident. Observons comment fonctionnent les calculs sur un exemple : on considère l'équation différentielle (E)

$$\begin{cases} \dot{x} &= -\frac{x}{x^2+y^2} - y \\ \dot{y} &= -\frac{y}{x^2+y^2} + x. \end{cases}$$

Utilisons les coordonnées polaires $x = r \cos(\theta)$, $y = r \sin(\theta)$. On obtient en dérivant

$$\begin{cases} \dot{x} &= \dot{r} \cos(\theta) - r \dot{\theta} \sin(\theta) \\ \dot{y} &= \dot{r} \sin(\theta) + r \dot{\theta} \cos(\theta). \end{cases}$$

En inversant ce système, il vient

$$\begin{cases} \dot{r} &= \frac{1}{r}(x\dot{x} + y\dot{y}) \\ \dot{\theta} &= \frac{1}{r}(\dot{y} \cos(\theta) - \dot{x} \sin(\theta)). \end{cases}$$

On peut maintenant écrire l'équation (E) dans ces coordonnées sous la forme du système suivant, système dont la résolution est immédiate

$$\begin{cases} \dot{r} &= -\frac{1}{r} \\ \dot{\theta} &= 1. \end{cases}$$

Les solutions sont $(r(t), \theta(t)) = (\sqrt{a-2t}, t+b)$. En revenant aux coordonnées cartésiennes on obtient

$$(x(t), y(t)) = (\cos(t+b)\sqrt{a-2t}, \sin(t+b)\sqrt{a-2t}).$$

.

EXERCICE :

1. Montrez que pour tout t , $(\phi_t)_* X = X$.
2. Montrez que $(\psi \circ \phi)_*(Z) = \psi_*(\phi_*)(Z)$.

9.3 Forme normale : le flot en dehors d'un point équilibre

En dehors, d'un point d'équilibre, la situation est simple localement : *il existe des coordonnées dans lesquelles le champ de vecteurs est constant*. Soyons plus précis :

Théorème 9.3.0.4 [BOÎTE DE FLOT] *Soit X un champ de vecteur C^1 . On suppose que $X(x_0) \neq 0$. Il existe alors un voisinage U , de x_0 , un difféomorphisme ψ de U sur un ouvert de $\mathbb{R}^p$, tel que*

$$\phi_*(X) = (1, 0, \dots, 0).$$

En particulier, le flot φ_t de X est donné pour t petit et x dans U

$$\varphi_t(x) = \psi^{-1}(\psi(x) + (t, 0 \dots)).$$

Attention, ce théorème est trompeur : dans les nouvelles coordonnées, l'équation différentielle est très simple à résoudre, mais pour produire ces coordonnées, le théorème utilise la résolution de l'équation différentielle !

DÉMONSTRATION : Pour simplifier les notations nous supposons que $x_0 = 0$. Comme $X(x_0) \neq 0$, nous considérons une base $\{e_1, \dots, e_p\}$ de E telle que $e_1 = X(x_0)$. Nous utilisons cette base pour identifier E avec $\mathbb{R}^p$. D'après le théorème de Cauchy-Lipschitz, il existe $\varepsilon > 0$, tel que si $x = (x_1, \dots, x_p)$ vérifie $\sup |x_i| < \varepsilon$, alors $\varphi_t(x)$ est défini pour $t \in]-\varepsilon, \varepsilon[$. Nous considérons alors l'application

$$\theta : \begin{cases} O =]-\varepsilon, \varepsilon[\rightarrow E \\ (x_1, x_2, \dots, x_p) \mapsto \varphi_{x_1}(0, x_2, \dots, x_p) \end{cases}.$$

Remarquons tout d'abord que pour tout x de O

$$\begin{aligned} D_x \theta(e_1) &= \left. \frac{d}{ds} \right|_{s=0} \varphi_{x_1+s}(0, x_2, \dots, x_p) \\ &= \left. \frac{d}{ds} \right|_{s=0} \varphi_s(\varphi_{x_1}(0, x_2, \dots, x_p)) \\ &= \left. \frac{d}{ds} \right|_{s=0} \varphi_s(\theta(x)) \\ &= X(\theta(x)). \end{aligned} \tag{9.1}$$

Calculons la différentielle de θ en 0. Nous remarquons que si $i \neq 1$

$$\begin{aligned} D_0 \theta(e_i) &= \left. \frac{d}{ds} \right|_{s=0} \varphi_0(0, x_2, \dots, x_i + s, \dots, x_p) \\ &= \left. \frac{d}{ds} \right|_{s=0} (0, x_2, \dots, x_i + s, \dots, x_p) \\ &= e_i. \end{aligned} \tag{9.2}$$

Combiné avec le résultat précédent, nous en déduisons que $D_0\theta = \text{Id}$. D'après le théorème d'inversions locale θ est un difféomorphisme au voisinage de 0. D'après la relation 9.1, $\theta_*(e_1) = X$. Nous en déduisons le théorème en posant $\psi = \theta^{-1}$. $\square$

Chapitre 10

L'application premier retour et les orbites périodiques

Après notre étude des points fixes, nous allons continuer l'étude des ensembles invariants par le flot en étudiant les *orbites périodiques*. Nous nous poserons une question identique à celle que nous posons pour les points fixes : une orbite périodique est-elle attractante ? les points proches de cette orbite vont-ils s'accumuler sur cette orbite ?

Dans tout ce chapitre, nous considérerons un champ de vecteur X sur un espace affine E de dimension 2. Nous noterons $m \rightarrow \varphi_t(m)$ ou $m \rightarrow \varphi(t, m)$ le flot de X . Nous considérons une *orbite périodique* γ de période T , avec $T > 0$, c'est à dire une courbe γ définie de $\mathbb{R}$ dans E et telle que

- $\forall s, \gamma(s + T) = \gamma(s)$,
- $\forall s, \dot{\gamma}(s) = X(\gamma(s))$.

Nous identifions E avec $\mathbb{R}^2$, en choisissant $\gamma(0)$ comme origine, et une base (e_1, e_2) telle que $e_2 = X(\gamma(0))$.

Nous appellerons *transversale*, l'axe des x , c'est à dire la droite passant par $0 = x_0 = \gamma(0)$ et parallèle à e_1 .

10.1 L'application premier retour de Poincaré

L'idée est très simple : partant de l'origine le flot revient en un temps T à l'origine. si nous prenons ensuite un point proche de l'origine x_0 sur la transversale, il va revenir lui aussi, non pas sur lui-même, mais sur un point proche et toujours sur la transversale. L'application dont nous venons d'évoquer la construction qui envoie des points de la transversale proche de l'origine sur la transversale, s'appelle application premier retour. Nous allons maintenant être

plus précis et rigoureux. Nous notons D_x l'axe des x

$$D_x = \{(x, 0), \quad x \in \mathbb{R}\}.$$

Nous noterons en général (m^1, m^2) les coordonnées du point m . Le flot ϕ_t s'écrit alors (ϕ_t^1, ϕ_t^2) . On a $X = (X^1, X^2)$ et par hypothèses $X^1(x_0) = 0$ et $X^2(x_0) = 1$.

Théorème 10.1.0.5 [PREMIER RETOUR] *Il existe $\varepsilon > 0$, une fonction C^1 , T définie de $] -\varepsilon, \varepsilon[$ dans $\mathbb{R}$ et telle que $T(0) = T$, telle que pour tout $x \in] -\varepsilon, \varepsilon[$, nous avons*

$$\varphi_{T(x)}^2(x, 0) = 0,$$

De plus T est unique au voisinage de 0.

L'application G telle que

$$G(x) = \varphi_{T(x)}^1(x, 0)$$

s'appelle *application de premier retour*. L'application $T(x)$ s'appelle *temps de retour*.

DÉMONSTRATION : la démonstration va être une application simple du théorème des fonctions implicites. Écrivons le flot $\varphi_t = (\varphi_t^1(x), \varphi_t^2(x))$. Nous cherchons donc une fonction T définie au voisinage de 0 telle

$$\varphi_{T(x)}^2(x, 0) = 0.$$

Nous considérons la fonction définie sur $\mathbb{R}^2$ au voisinage de $(T, 0)$ par

$$F(t, x) = \varphi_t^2(x, 0).$$

Nous remarquons que $F(T, 0) = 0$ et que

$$\frac{\partial}{\partial t} \Big|_{(T, 0)} F(t, x) = \frac{d}{dt} \Big|_0 \varphi_t^2(0, 0) = X^2(0, 0) = 1.$$

Le théorème des fonctions implicites nous permet donc de conclure. $\square$

10.1.1 Dérivée de l'application premier retour

Soit donc G l'application premier retour, nous allons montrer

Proposition 10.1.1.1 *Nous avons*

$$\dot{G}(0) = \det(D_{\gamma(0)}\varphi_T),$$

EN PARTICULIER, LA DÉRIVÉE DE L'APPLICATION PREMIER RETOUR NE DÉPEND QUE DE L'ORBITE ET PAS DES CHOIX FAITS DANS LA CONSTRUCTION

DÉMONSTRATION : Nous savons tout d'abord que

$$D_{x_0}\varphi_T(X(x_0)) = X(x_0).$$

Ceci s'obtient en dérivant par rapport à t l'équation $\varphi_t(x_0) = \varphi_T\varphi_t(x_0)$. Autrement dit dans la base (e_1, e_2) la matrice de $D_{x_0}\varphi_T$ s'écrit

$$\begin{pmatrix} a & 0 \\ c & 1 \end{pmatrix}$$

Nous remarquons donc que

$$\det(D_{x_0}\phi_T) = a = \partial_1\varphi^1(T, 0). \quad (10.1)$$

Maintenant par construction, nous avons

$$G(s) = \varphi_{T(s)}^1(s, 0).$$

Posons $H(t, s) = \varphi_t^1(s, 0)$, alors

$$\dot{G}(0) = \frac{dT}{ds}(T)\partial_1H(T, 0) + \partial_2H(T, 0). \quad (10.2)$$

Or

$$\begin{cases} \partial_1H(T, 0) &= \partial_t|_{t=T}\varphi_t^1(x_0) = \partial_t|_{t=0}\varphi_t^1(x_0) = X^1(x_0) = 0, \\ \partial_2H(T, 0) &= \partial_s|_{s=0}\varphi_T^1(s, 0) = \partial_1\varphi_T^1(0, 0). \end{cases}$$

Nous en déduisons finalement en combinant les équations (10.1) et (10.2) que

$$\dot{G}(0) = \partial_1\varphi_T^1(0, 0) = \det(D_{x_0}\varphi_T).$$

□

Nous en déduisons le corollaire suivant

Corollaire 10.1.1.2 *Nous avons*

$$\log(\dot{G}(0)) = \int_0^T \text{trace}(D_{\gamma(s)}X)ds.$$

DÉMONSTRATION : Par le théorème 8.1.2.1, la différentielle du flot est la résolvante de l'équation linéarisée. Le corollaire se déduit alors du théorème de Liouville 2.5.0.19. □

10.2 Orbite périodique stable

Soit $\gamma(t)$ une orbite périodique de période T . Nous dirons qu'elle est *attractrice* si pour tout ouvert U contenant $\gamma[0, T]$, il existe un ouvert V tel que

- Nous avons les inclusions $\gamma[0, T] \subset V \subset U$
- Pour tout x de V ,

$$\lim_{t \rightarrow \infty} \inf_{s \in [0, T]} \|\varphi_t(x) - \gamma(s)\| = 0.$$

De manière plus imagée l'orbite de tout point de V va se rapprocher de l'orbite périodique.

Le théorème qui clôt ce chapitre est un analogue du théorème 7.4.0.5

Théorème 10.2.0.3 [STABILITÉ DES ORBITES PÉRIODIQUES] *Soit X un champ de vecteur défini sur le plan. Soit γ une orbite périodique de période T . Supposons*

$$\int_0^T \text{trace}(D_{\gamma(s)}X)ds < 0.$$

Alors, l'orbite périodique γ est stable.

Remarquons au passage que la quantité intégrée et bien connue sous le nom de *divergence* : $\text{trace}(DX)ds = \text{div}(X)$.

Nous n'allons pas donner une preuve rigoureuse de ce théorème, nous allons seulement en démontrer une version faible.

DÉMONSTRATION : Nous allons seulement démontrer un résultat plus faible. Choisissons nos coordonnées comme au début de ce chapitre de telle sorte que $\gamma(0) = (0, 0)$, alors il existe $\varepsilon > 0$ tel que si m est un point de la transversale (l'axe des x) ε -proche de $\gamma(0)$ alors il existe une suite $\{t_n\}$ tendant vers l'infini tel que

$$\lim_{n \rightarrow \infty} \varphi_{t_n}(m) = \gamma(0).$$

Ceci justifie bien le fait que l'orbite de x se rapproche de γ .

Soit G l'application premier retour. D'après le corollaire 10.1.1.2, la dérivée de l'application premier retour est de valeur absolue strictement plus petite que 1. Nous en déduisons qu'il existe ε strictement positif, et k strictement plus petit que 1 tel que

$$\forall s \in]-\varepsilon, \varepsilon[, |G(x)| \leq k|x|.$$

Par continuité du temps de retour, on peut choisir de plus ε suffisamment petit pour que

$$\forall s \in]-\varepsilon, \varepsilon[, \frac{T}{2} \leq T(s) \leq \frac{3T}{2}.$$

Dès lors, si $x \in]-\varepsilon, \varepsilon[$, alors

$$\lim_{k \rightarrow \infty} G^k(x) = 0,$$

Or $G^k(x) = \varphi_{T_k}(x)$ avec

$$T_k = T(G^{k-1}(x)) + T_{k-1} \geq \frac{kT}{2}.$$

On a donc bien $T_k \rightarrow \infty$, et $\lim_{k \rightarrow \infty} \phi_{t_k}(x) = x_0$.

Pour démontrer rigoureusement le théorème, il faut arriver à contrôler ce qui se passe en dehors de la transversale. Ce n'est pas très difficile, mais cela ajoute des détails techniques qui ne sont pas très intéressants à ce stade. $\square$

Un exemple

Nous allons traiter le système

$$\begin{cases} \dot{x} &= -y + x(1 - (x^2 + y^2)) \\ \dot{y} &= x + y(1 - (x^2 + y^2)) \end{cases}$$

Nous vérifions que $\gamma(t) = (\cos(t), \sin(t))$ est une orbite périodique du flot de période 2π . Nous allons montrer qu'elle est stable.

Pour cela nous devons écrire la différentielle du champ de vecteur. On a

$$D_{x,y}X = \begin{pmatrix} 1 - y^2 - 3x^2 & -1 - 2xy \\ 1 - 2xy & 1 - x^2 - 3y^2 \end{pmatrix}.$$

Ainsi

$$\text{trace} D_{x,y}X = 2 - 4(x^2 + y^2).$$

Le long de γ , nous avons

$$\int_0^{2\pi} D_{\gamma(s)}X ds = -4\pi < 0.$$

Nous pouvons donc appliquer le théorème 10.2.0.3 pour conclure à la stabilité de l'orbite γ .

Chapitre 11

Cols

11.1 Instabilité et spectre

Nous allons compléter l'étude des points fixes et de leur stabilité (cf Théorème 7.4.0.5) et démontrer le théorème suivant

Théorème 11.1.0.4 *Soit X un champ de vecteur ayant un zéro en x_0 . On suppose que $D_{x_0}X$ possède une valeur propre de partie réelle positive. Alors x_0 est instable.*

Une construction préliminaire

On choisit des coordonnées sur E de telle sorte que $x_0 = 0$. Soit V la somme des espaces caractéristiques associé aux valeurs propres de module plus grand que 1. Soit $B = \exp(D_0X)$. Par hypothèse, V n'est pas réduit à $\{0\}$. Soit W la somme des espaces caractéristiques associé aux valeurs propres de module inférieur ou égal à 1. On a

$$V \oplus W = E.$$

On choisit une métrique telle que cette décomposition est orthogonale. Soit π la projection orthogonale sur V . Nous rappelons qu'il existe alors un entier p positif et des réels μ et ν avec $\mu > \nu \geq 1$ tels que si $A = B^p = \exp(pD_0X)$.

$$\forall x \in V, \quad \|A(x)\| \geq \mu\|x\|, \quad (11.1)$$

$$\forall x \in W, \quad \|A(x)\| \leq \nu\|x\|. \quad (11.2)$$

Notons π la projection orthogonale sur V . Soit alors

$$C_\rho = \{x \in E, \|\pi(x)\| > \rho\|(1 - \pi)(x)\|\}. \quad (11.3)$$

On remarque que si $x \in C_\rho$, alors

$$\|x\| \leq (1 + \rho)\|\pi(x)\|. \quad (11.4)$$

Un peu d'algèbre linéaire

Montrons alors

Proposition 11.1.0.5 *Il existe des nombres réels $\beta > 0$ et $\lambda > 1$, tels que*

$$x \in C_\beta \implies \|A(x)\| \geq \lambda \|x\|. \quad (11.5)$$

De plus, soit $\alpha = \frac{\mu}{\nu}\beta$. Alors

$$x \in C_\beta \implies A(x) \in C_\alpha. \quad (11.6)$$

DÉMONSTRATION : Prenons en effet β tel que $\lambda = \frac{\mu}{1+\beta} > 1$. Alors, si $x \in C_\beta$, nous obtenons bien

$$\begin{aligned} \|A(x)\| &\geq \|\pi(A(x))\| \\ &\geq \|A(\pi(x))\| \\ &\geq \mu \|\pi(x)\| \\ &\geq \lambda \|x\|. \end{aligned}$$

La première inégalité vient de ce que π est une projection orthogonale, la deuxième que A commute avec π , la troisième de l'inégalité (11.1), la dernière de l'inégalité (11.4) et de la définition de λ .

La deuxième partie de la proposition est évidente. $\square$

Linéarisation

Par le théorème 8.1.2.1 sur la différentielle du flot, nous avons $A = D_0\varphi_p$. Par définition de la différentielle, nous en déduisons que pour tout $\varepsilon > 0$, il existe une boule B_ε centrée sur 0 telle que

$$\forall x \in B, \quad \|\varphi_p(x) - A(x)\| \leq \varepsilon \|x\|. \quad (11.7)$$

Nous montrons maintenant

Proposition 11.1.0.6 *Soit β donné par la Proposition 11.1.0.5. Il existe des nombres réels κ et ε , tels que $\kappa > 1$ et*

$$x \in B_\varepsilon \cap C_\beta \implies \varphi_p(x) \in C_\beta \text{ et } \|\varphi_p(x)\| \geq \kappa \|x\|.$$

DÉMONSTRATION : Nous savons que si x appartient à C_β alors d'après les inégalités (11.4) et (11.1)

$$\|x\| \leq (1 + \beta) \|\pi(x)\| \leq \frac{1 + \beta}{\mu} \|A(\pi(x))\| = \frac{1 + \beta}{\mu} \|\pi(A(x))\|. \quad (11.8)$$

Posons $\xi = \frac{1+\beta}{\mu}$. Si x appartient à $B_\varepsilon \cap C_\beta$, alors, d'après les inégalités (11.7) et (11.8), nous avons

$$\begin{aligned} \|\pi(\varphi_p(x))\| &\geq \|\pi(A(x))\| - \varepsilon\|x\| \\ &\geq (1 - \varepsilon\xi)\|\pi(A(x))\| \end{aligned}$$

D'a la proposition 11.1.0.5, $A(x) \in C_\alpha$ et donc

$$\begin{aligned} \|(1 - \pi)(\varphi_p(x))\| &\leq \|(1 - \pi)(A(x))\| + \varepsilon\|x\| \\ &\leq \alpha\|\pi(A(x))\| + \varepsilon\|x\| \\ &\leq (\alpha + \varepsilon\xi)\|\pi(A(x))\| \end{aligned}$$

Puisque $\alpha > \beta$, pour ε suffisamment petit, on a

$$\frac{\alpha - \varepsilon\xi}{1 + \varepsilon\xi} > \beta.$$

Ainsi

$$\|\pi(\varphi_p(x))\| \geq \beta\|(1 - \pi)(\varphi_p(x))\|$$

Ceci signifie que $\varphi_p(x)$ appartient à C_β . Enfin, prenons ε suffisamment petit pour que $\kappa = \lambda - \varepsilon > 1$, alors

$$\|\varphi_p(x)\| \geq \|A(x)\| - \varepsilon\|x\| \geq (\lambda - \varepsilon)\|x\| \geq \kappa\|x\|.$$

Nous avons fini de démontrer la proposition. $\square$

Preuve du théorème 11.1.0.4

DÉMONSTRATION : La proposition 11.1.0.6 interdit la stabilité. Supposons en effet que x est stable. Alors par définition, il existe un voisinage U de 0 tel que pour tout entier k positif $\varphi_k(U) \in B_\varepsilon$.

Une simple récurrence et la proposition 11.1.0.6 montre que

$$x \in U \cap C_\beta \implies \phi_{kp}(X) \in C_\beta \cap B_\varepsilon \text{ et } \|\varphi_{kp}(x)\| \geq \kappa^p\|x\|.$$

Ainsi

$$\lim_{p \rightarrow \infty} \|\varphi_{kp}(x)\| = \infty.$$

Ceci nous fournit notre contradiction. $\square$

11.2 Séparatrices pour les cols

Nous nous plaçons désormais dans un espace affine E de dimension 2.

11.2.1 Col ou point selle

Soit X un champ de vecteur ayant un zéro en x_0 . Nous dirons que X est un *point selle* ou un *col* si $D_{x_0}X$ a deux valeurs propres de signe opposé. Nous remarquons que x_0 est un col si et seulement si $\det D_{x_0}X < 0$.

11.2.2 Séparatrice

Soit x_0 un point selle. Nous dirons qu'une orbite non constante $c(t)$ de X définie sur $[0, \infty[$ est une *séparatrice stable* si

$$\lim_{t \rightarrow +\infty} c(t) = x_0.$$

Nous dirons qu'une orbite non constante $c(t)$ de X définie sur $] -\infty, 0[$ est une *séparatrice instable* si

$$\lim_{t \rightarrow -\infty} c(t) = x_0.$$

Nous avons alors la proposition.

11.3 Existence de quatre séparatrices

Notre but est de montrer

Théorème 11.3.0.1 [SÉPARATRICES ET COLS] *Soit X un champ de vecteur dans un espace affine de dimension 2. Soit x_0 un point selle de X . Alors x_0 possède exactement quatre séparatrices, deux stables et deux instables. De plus, prolongé par continuité en x_0 , les deux séparatrices stables (resp. instables) sont tangentes à la direction propre de $D_{x_0}X$ de valeur propre positive (resp. négative) mais dans des sens opposés.*

Nous allons donner dans le prochain paragraphe une description plus précise du comportement des orbites dans un voisinage du point col.

Coordonnées

Nous choisissons des coordonnées sur E de telle sorte que $x_0 = 0$, $A = D_0X$ soit diagonale avec sa première valeur propre positive. Notons λ la valeur propre positive de A et $-\mu$ sa valeur propre négative.

Nous notons à $X = (P, Q)$ les coordonnées de X . Pour tout a et b , nous considérons l'ouvert

$$Q_1(a) = \{(x, y) \in E, a > x > 0, |y| < x\}$$

On note

$$\begin{aligned} I_1(a) &= \{(x, y) \in E, a = x, |y| < x\} \\ \Delta_{12} &= \{(x, x) \in E, a > x > 0\} \\ \Delta_{41} &= \{(x, -x) \in E, a > x > 0\} \end{aligned}$$

Nous montrerons le lemme suivant

Lemme 11.3.0.2 [COMPORTEMENT DANS UN QUADRANT] *Soit b quelconque. Pour a suffisamment petit, le quadrant $Q_1(a)$ contient une unique séparatrice instable c_0 . De plus, c_0 est le graphe au-dessus de l'axe des x d'une fonction $y(x)$ définie sur $[0, a]$ et qui vérifie*

$$\frac{dy}{dx} = \frac{Q(x, y)}{P(x, y)}$$

Finalement, soit c une orbite de X passant par $Q_1(a)$ et différente de c_0 . Alors c entre dans $Q_1(a)$ par l'une des diagonales Δ_{12} ou Δ_{41} et sort au bout d'un temps fini par $I_1(a)$

Ce lemme et les énoncés analogues pour les autres quadrants décrit complètement le comportement des orbites au voisinage d'un point selle.

11.3.1 Une équation différentielle associée

Posons $G(x, y) = \frac{Q(x, y)}{P(x, y)}$. Nous considérons alors l'équation différentielle (non autonome) en x définie sur $\mathcal{O} = \{(x, y)/P(x, y) \neq 0\}$.

$$\frac{dy}{dx} = G(x, y) \tag{11.9}$$

Notre but est de montrer

Lemme 11.3.1.1 *Il existe a tel que pour*

- si y est une solution de l'équation (11.9), vérifiant $|y(x_0)| \leq x_0 < a$, alors y est défini sur l'intervalle $[x_0, a]$ et vérifie*

$$\forall s \in]x_0, a], \quad |y(s)| < s.$$

- Il existe une unique solution y_0 définie sur $]0, a]$ et vérifiant*

$$\forall s \in]0, a], \quad |y(s)| < s.$$

Commençons par démontrer la proposition suivante

Proposition 11.3.1.2 *Il existe $a > 0$ tel que $Q_1(a) \subset \mathcal{O}$. Autrement dit, il existe $a > 0$ tel que, pour tout (x, y) de $Q = Q_1(a)$, on a*

$$P(x, y) > 0 \tag{11.10}$$

De plus

$$\forall x \in [0, a], \quad G(x, x) < 0 \text{ et } G(x, -x) > 0. \tag{11.11}$$

Enfin, on a

$$\partial_y G(x, y) < 0.$$

DÉMONSTRATION : Dans nos coordonnées, nous avons les développements limités

$$\begin{aligned} P(x, y) &= \lambda x + \|(x, y)\| \varepsilon_1(x, y) \\ Q(x, y) &= -\mu y + \|(x, y)\| \varepsilon_2(x, y) \end{aligned}$$

Nous savons que sur Q on a $\|(x, y)\| \leq 2|x|$. L'assertion 11.11 suit. Pour la deuxième assertion, remarquons que $\partial_y G(x, y)$ est du signe de

$$C(x, y) = \partial_y Q(x, y) \cdot P(x, y) - \partial_y P(x, y) Q(x, y).$$

Or

$$\begin{aligned} \partial_y P(x, y) &= \varepsilon_3(x, y) \\ \partial_y Q(x, y) &= -\mu + \varepsilon_4(x, y), \end{aligned}$$

En regroupant les termes, il vient

$$C(x, y) = -\mu \lambda \|x\| + \|(x, y)\| \varepsilon_5(x, y).$$

Le même raisonnement que précédemment nous donne que $C(x, y)$ est strictement négative pour sur $Q_1(a)$ pour a suffisamment petit. $\square$

Démonstration du lemme 11.3.1.1

DÉMONSTRATION : La proposition précédent nous dit que l'équation $\frac{dy}{dx} = G(y, x)$ est bien définie sur le quadrant $Q_1(a)$ et que de plus ce quadrant est un anti-entonnoir pour cette équation. Le lemme 11.3.1.1 suit alors du théorème 6.2.3.2 sur les anti-entonnoirs étranglés et dispersifs. $\square$

11.3.2 Retour aux orbites du champ de vecteur et preuve du lemme 11.3.0.2

Nous faisons le lien avec les orbites du champ de vecteur X .

Proposition 11.3.2.1 *Soit c une orbite du champ X de condition initiale (x_0, y_0) dans $Q_1(a)$. Il existe alors $s > 0$, tel que*

1. $c[0, s] \in Q_1(a)$
2. $c(s) \in I(a)$
3. $c[0, s]$ est le graphe par rapport de l'axe des x d'une fonction y solution de l'équation différentielle (11.9).

DÉMONSTRATION : Soit $]u, v[$ l'intervalle de définition de c . Soit $[0, t[\subset]u, v[$ l'intervalle maximal telle que $c[0, t[\in Q_1(a)$. Comme $Q_1(a)$ est borné et d'adhérence incluse dans le domaine de définition de X , on a $t < v$ et $c(t) \notin Q_1(a)$.

Écrivons $c = (c_1, c_2)$. Par définition on a $\dot{c}_1 = P(c_1, c_2)$. Ainsi d'après l'assertion (11.10), $\dot{c}_1 > 0$. La fonction $t \rightarrow c_1$ est donc un difféomorphisme de $[0, t]$ vers $[x_0, c_1(t)]$. Ainsi, la courbe c est le graphe de l'application $y = c_2 \circ c_1^{-1}$. On obtient par ailleurs

$$\frac{dy}{dx} = \frac{dc_2}{dt} \left(\frac{dc_1}{dt} \right)^{-1} = \frac{P(x, y)}{Q(x, y)} = G(x, y).$$

Le reste de la proposition suit du lemme 11.3.1.1. $\square$

Nous en déduisons la proposition suivante dont le lemme 11.3.0.2 se déduit.

Proposition 11.3.2.2 *Soit α l'unique point de I . Soit c_0 est l'orbite du champ de vecteur X de condition initiale (a, α) alors*

- $\forall t < 0$ on a $c(t) \in Q_1(a)$.
- c est définie sur l'intervalle $] -\infty, 0]$
- Enfin

$$\lim_{t \rightarrow -\infty} c(t) = 0.$$

De plus, soit c est orbite de X passant par $Q_1(a)$ et différente de c_0 , elle entre dans $Q_1(a)$ par l'une des diagonales Δ_{12} ou Δ_{41} et sort au bout d'un temps fini par $I_1(a)$

11.3.3 Unicité des séparatrices

Pour compléter la preuve du théorème, il nous faut démontrer l'unicité des séparatrices. Celle ci se déduit de la proposition suivante, qui utilise elle même la proposition 11.3.2.1.

Proposition 11.3.3.1 *Soit c une orbite du champ de vecteur de condition initiale dans $Q_1(a)$. Alors il existe un intervalle $[0, s]$ tel que*

$$\forall t \in [0, s], \quad c(t) \in Q_1(a). \quad (11.12)$$

$$\|c(s)\| \geq a \quad (11.13)$$

Dès lors, si c est une séparatrice stable, alors il existe t_0 tel que

$$\forall t > t_0, \quad c(t) \notin Q_1(a).$$

L'unicité des séparatrices se déduit aisément de cette proposition

11.3.4 Position des tangentes

Il nous faut encore comprendre la position des tangentes en l'origine. Considérons donc pour tout b positif

$$Q_1(b, \theta) = \{(x, y) / 0 < x < b, |y| < \theta x\}.$$

Pour finir de démontrer le théorème en ce qui concerne la position des tangentes, il suffit de démontrer la proposition

Proposition 11.3.4.1 *Pour tout $\theta > 0$, il existe b , tel que la solution de l'équation (11.9) définie sur $[0, b]$ est incluse dans $Q_1(b, \theta)$*

DÉMONSTRATION : Ceci se déduit d'un simple changement de coordonnées considérons les coordonnées

$$(\tilde{x}, \tilde{y}) = (x, \theta^{-1}y).$$

Ces coordonnées sont tout aussi valables que celles que nous avons choisies. En particulier, nous pouvons appliquer le lemme 11.3.0.2 dans ces coordonnées et en déduire qu'il existe s et b tel que pour tout $t < s$ nous avons pour

$$c_0(t) \in \tilde{Q}_1(b) = \{(\tilde{x}, \tilde{y}), \quad 0 < \tilde{x} < b, |\tilde{y}| < \tilde{x}\}.$$

Comme

$$\tilde{Q}_1(b) = Q_1(b, \theta),$$

nous en déduisons la proposition. $\square$

Deuxième partie

Introduction à la Mécanique Classique

Chapitre 12

Fonctions invariantes

Pour étudier un système dynamique, il est utile de chercher tout d'abord à examiner les objets qu'il laisse invariants que ce soit des fonctions, mesures, points, ensemble etc.

Nous nous attacherons tout d'abord aux fonctions. Dans la terminologie de la mécanique classique, de telles fonctions s'appellent *intégrales premières du mouvement* et elles jouent un grand rôle. Nous avons vu dans la proposition 2.1, que l'existence de fonctions invariantes permet de réduire le nombre de paramètres dans l'équation en se restreignant aux surfaces de niveaux. L'exemple de la mécanique Newtonienne est bien connu : la conservation de l'énergie pour les systèmes sans frottements facilite grandement la recherche des solutions. Dans le meilleur des cas, celui des *systèmes complètement intégrables* de la mécanique classique, nous verrons que ces intégrales premières permettent effectivement d'intégrer le mouvement c'est-à-dire de calculer les trajectoires dans des coordonnées adéquates.

Nous expliquerons quelques méthodes permettant de chercher des intégrales premières du mouvement, particulièrement dans le cadre des mécaniques lagrangienne et hamiltonienne, où elles apparaissent liées aux symétries du système. Nous verrons également comment utiliser ces intégrales premières pour décrire efficacement le système.

Dans ce premier chapitre, nous précisons quelques définitions et remarques élémentaires.

12.1 Définitions

Nous revenons sur une notion déjà introduite pour l'approfondir.

Soit donc X un champ de vecteur défini sur M , où M est un ouvert de $\mathbb{R}^n$ ou, plus généralement encore une sous-variété, et ϕ_t son flot. Une fonction f , définie sur M , est dite *invariante* par le flot si pour tout x , la fonction définie sur $\mathbb{R}$ par $t \mapsto f(\phi_t(x))$ est constante. On dit aussi que f est une *intégrale première du mouvement*.

REMARQUES :

1. Pour le moment, on n'a supposé aucune régularité à f ; par la suite il sera intéressant de considérer le cas où la régularité de f est la plus faible possible à savoir mesurable. Si par contre f est de classe C^1 , on montre aisément que f est invariante si et seulement si $df(X) = 0$.
2. On considère les équations de la mécanique newtonienne " $f = m\gamma$ " dans le cas d'une force dérivant d'un potentiel U . Nous pouvons écrire cette équation du second ordre sous la forme

$$\begin{cases} \frac{d}{dt}\dot{q} &= -m \frac{\partial}{\partial q}U(q) \\ \frac{d}{dt}q &= \dot{q}. \end{cases}$$

Montrez alors que l'énergie totale

$$E = \frac{1}{2}m\dot{q}^2 + U(q),$$

est une intégrale première du mouvement.

3. On a déjà vu que l'Hamiltonien H est une intégrale première du mouvement pour les équations de Hamilton.

$$\begin{cases} \frac{d}{dt}p &= -\frac{\partial}{\partial q}H, \\ \frac{d}{dt}q &= \frac{\partial}{\partial p}H. \end{cases}$$

4. Soit x_0 un point fixe du flot ϕ^t . On suppose que le bassin d'attraction de x_0

$$B = \{y, \lim_{t \rightarrow \infty} \phi_t(y) = x_0\},$$

est ouvert. Montrez alors que les seules fonctions continues invariantes définies au voisinage de x_0 sont constantes. Qu'en est-il des fonctions mesurables ? sont-elles presque partout constantes ?

5. Plus généralement, soit γ une orbite bornée de du flot ϕ_t ; On suppose que le bassin d'attraction de γ

$$B = \{y, \lim_{t \rightarrow \infty} d(\phi_t(y), \gamma) = 0\},$$

est ouvert. Montrez alors que les seules fonctions continues invariantes définies au voisinage de γ sont constantes

Chapitre 13

Mécanique Lagrangienne

Les équations d'Euler-Lagrange décrivent le mouvement d'une particule obéissant à un principe de moindre action. Le théorème d'Emmy Noether montrera comment les symétries des données du problème donnent naissance à des intégrales premières du mouvement.

13.1 Lagrangien

13.1.1 Equations d'Euler-Lagrange

Soit U un ouvert d'un espace vectoriel E muni des coordonnées $q = (q_i)_{i=1\dots n}$. Dans la mécanique lagrangienne, on considère l'*espace des phases* $U \times E$, interprété comme l'espace des positions-vitesses. Le premier facteur du produit correspond aux *positions* et les coordonnées correspondantes sont notées $q = (q_i)_{i=1,\dots,n}$ comme précédemment. Le deuxième facteur correspond aux *vitesses* et les coordonnées correspondantes sont notées cette fois-ci $\dot{q} = (\dot{q}_i)_{i=1,\dots,n}$. Le *Lagrangien* du système est une fonction $L = L(q, \dot{q})$ définie sur l'espace des phases. Les *équations d'Euler-Lagrange* associées sont données par

$$\begin{cases} \frac{d}{dt} \frac{\partial}{\partial \dot{q}} L(q, \dot{q}) &= \frac{\partial}{\partial q} L(q, \dot{q}) \\ \frac{d}{dt} q &= \dot{q}. \end{cases}$$

13.1.2 Points critiques de l'action

Il est important de comprendre la provenance de ces équations. Elles proviennent de l'importation en mécanique du principe de Fermat en optique géométrique : *la lumière suit le plus court chemin*.

Précisons ces idées. Donnons-nous un Lagrangien L . L'*action* correspondant à L associe à tout arc de courbe $c :]a, b[\rightarrow U$ le nombre

$$S(c) = \int_c L(c, \frac{d}{dt}c) dt.$$

Une *variation à extrémités fixées* de la courbe c est une famille de courbes $c^s : [a, b] \rightarrow U$ dépendant d'un paramètre réel s telle que

- $(s, t) \mapsto c^s(t)$ est de classe C^1 ,
- $c^0 = c$,
- $c^s(a) = c(a)$, $c^s(b) = c(b)$.

Une courbe c est *point critique de l'action* si pour toute variation à extrémités fixées c^s on a

$$\frac{d}{ds} \Big|_{s=0} S(c^s).$$

On montre alors :

Théorème 13.1.2.1 *La courbe c est point critique de l'action, si et seulement si la courbe $(c, \frac{d}{dt}c)$, tracée dans l'espace des phases, est solution de l'équation d'Euler-Lagrange.*

DÉMONSTRATION : Considérons $c^s = (c_1^s, \dots, c_n^s)$ une variation de la courbe c à extrémités fixées. Pour alléger nos notation, nous écrirons $\int f dt$ à la place de $\int f(t) dt$. Nous avons

$$\begin{aligned} \frac{d}{ds} S(c_s) &= \int_a^b \frac{\partial}{\partial s} L(c^s, \frac{d}{dt} c^s) dt \\ &= \sum_i \left(\int_a^b \frac{\partial}{\partial q_i} L(c^s, \frac{d}{dt} c^s) \frac{\partial}{\partial s} c_i^s dt \right. \\ &\quad \left. + \int \frac{\partial}{\partial \dot{q}_i} L(c^s, \frac{d}{dt} c^s) \frac{\partial^2}{\partial s \partial t} c_i^s dt \right). \end{aligned}$$

En intégrant par partie le dernier terme, on obtient

$$\begin{aligned} \frac{d}{ds} S(c^s) &= \sum_i \left(\int_a^b \frac{\partial}{\partial q_i} L(c^s, \frac{d}{dt} c^s) \frac{\partial}{\partial s} c_i^s dt \right. \\ &\quad \left. - \int_a^b \frac{d}{dt} \frac{\partial}{\partial \dot{q}_i} L(c^s, \frac{d}{dt} c^s) \frac{\partial}{\partial s} c_i^s dt \right) \\ &\quad + \sum_i \left[\frac{\partial}{\partial \dot{q}_i} L(c^s, \frac{d}{dt} c^s) \frac{\partial}{\partial s} c_i^s \right]_a^b. \end{aligned}$$

Si nous posons $w(t) = \frac{\partial}{\partial s} \Big|_{s=0} c^s(t)$, et si nous considérons une variation à extrémités fixées, nous obtenons

$$\frac{d}{ds} \Big|_{s=0} S(c_s) = \sum_i \int_a^b \left(\frac{\partial}{\partial q_i} L(c, \frac{d}{dt} c) - \frac{d}{dt} \frac{\partial}{\partial \dot{q}_i} L(c, \frac{d}{dt} c) \right) w_i dt.$$

De cette formule, nous en déduisons que toute solution des équations d'Euler-Lagrange est point critique de l'action.

Par ailleurs, nous pouvons remarquer que si w vérifie $w(a) = w(b) = 0$, il existe une variation c^s à extrémités fixées telle que $w(t) = \frac{\partial}{\partial s}|_{s=0} c^s(t)$. Nous en déduisons que si c est point critique de l'action, alors c est solution des équations d'Euler-Lagrange. En effet, si une fonction f est telle que $\int_a^b f g dt = 0$ pour toute fonction g telle que $g(a) = g(b) = 0$ alors $f = 0$.

□

En Physique, on s'intéresse tout particulièrement aux minima de l'action, mais connaître les points critiques est une information importante. Il existe bien sûr des techniques plus sophistiquées permettant d'étudier les minima et maxima locaux de l'action, elles dépassent cependant le cadre de cette introduction.

13.1.3 Equation de Newton

Montrez que l'on retrouve les équations de la mécanique de Newton en considérant les équations d'Euler-Lagrange associées au Lagrangien

$$L = \frac{1}{2} m \|\dot{q}\|^2 - U(q).$$

13.1.4 Sous-variétés

Le cadre du problème s'étend naturellement aux sous-variétés V d'un espace vectoriel E . Un Lagrangien est alors une fonction définie sur l'espace tangent à $TV \subset V \times E$.

Souvent, on s'intéressera à des Lagrangiens restreints de Lagrangiens définis sur E tout entier. Des cas particuliers de Lagrangiens intéressants sont alors l'énergie (cinétique) $L = \frac{1}{2} \|\dot{q}\|^2$ ou la longueur $L = \|\dot{q}\|$. On cherche à nouveau les points critiques de l'action en ne s'intéressant qu'aux courbes tracées sur V .

Pour trouver ces points critiques, une stratégie consiste à choisir une paramétrisation de la sous-variété, à écrire le Lagrangien dans ces coordonnées, puis les équations d'Euler-Lagrange, comme cela va être fait dans l'exercice suivant.

13.1.5 Exemple et exercice : surface de révolution

On se donne une courbe γ tracée dans le plan de coordonnées (x, z) , donnée par $\gamma(t) = (r(t), z(t))$. On suppose que γ est paramétrée par l'arc. On considère la surface S obtenue par rotation autour de l'axe des z .

1. Donner une paramétrisation de la surface S , c'est-à-dire ici une immersion ϕ de $\mathbb{R}^2$, dont les coordonnées sont notées (r, θ) , dans $\mathbb{R}^3$ et dont l'image est S .
2. On restreint le Lagrangien $L = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$ à S . Autrement dit on considère l'action donnée pour les courbes $(r(t), \theta(t))$ tracées dans $\mathbb{R}^2$ par

$$S = \int \|D\phi(\frac{dr}{dt}, \frac{d\theta}{dt})\|^2.$$

Quelle est l'expression de L dans les coordonnées (r, θ) ?

3. Ecrire les équations d'Euler-Lagrange dans ces coordonnées. Que se passe-t-il dans le cas du cylindre ($r = \text{constante}$) ?

13.2 Symétries

Un paradigme de la physique est que toute symétrie est censée faciliter l'étude d'un système. Dans le cas de la mécanique lagrangienne, l'existence de symétries se traduit par celle d'intégrales premières.

Rappelons tout d'abord qu'un difféomorphisme $\varphi = (\varphi_i)_{i=1,\dots,n}$ d'un ouvert U d'un espace vectoriel U , donne naissance à un difféomorphisme

$$T\varphi : (q, \dot{q}) \mapsto (q, D_q\varphi(\dot{q})),$$

de l'espace des phases $U \times E$. Un difféomorphisme φ est une *symétrie du Lagrangien* si $L \circ T\varphi = L$.

Théorème 13.2.0.1 (E. NOETHER) *Soit φ^s un groupe à un paramètre s de symétries de L . Alors, la fonction*

$$I(q, \dot{q}) = \sum_i \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}) \frac{d}{ds} \Big|_{s=0} \varphi_i^s(q),$$

est une intégrale première du mouvement.

DÉMONSTRATION : On considère $(q(t), \dot{q}(t))$ une courbe tracée dans l'espace des phases et solution des équations d'Euler Lagrange. Nous obtenons alors

$$\begin{aligned} \frac{d}{dt} I(q, \dot{q}) &= \frac{d}{dt} \sum_i \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}) \frac{d}{ds} \varphi_i^s(q), \\ &= \sum_i \left(\frac{d}{dt} \left(\frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}) \right) \frac{d}{ds} \varphi_i^s(q) + \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}) \frac{d}{dt} \frac{d}{ds} \varphi_i^s(q) \right). \end{aligned}$$

On utilise les équations d'Euler-Lagrange dans le premier terme de droite pour obtenir

$$\frac{d}{dt} I(q, \dot{q}) = \sum_i \left(\frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}) \frac{d}{ds} \varphi_i^s(q) + \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}) \frac{d}{ds} \frac{d}{dt} \varphi_i^s(q) \right)$$

On reconnaît dans le dernier terme de droite une formule de composition des

dérivées. On a donc

$$\begin{aligned}\frac{d}{dt}I(q, \dot{q}) &= \frac{d}{ds}L(\varphi^s(q), \frac{d}{dt}\varphi^s(q)) \\ &= \frac{d}{ds}L(\varphi^s(q), D_q\varphi^s(\frac{d}{dt}q)) \\ &= \frac{d}{ds}L \circ T\varphi^s(q, \frac{d}{dt}q) \\ &= 0.\end{aligned}$$

On a bien montré que la fonction I est constante sur les solutions des équations d'Euler-Lagrange $\square$

13.2.1 Exemple et exercice, mouvement dans un champ central

Dans cet exercice, nous allons montrer comment on calcule et utilise des intégrales premières du mouvement.

On se donne dans ce qui suit un potentiel $U(q)$ sur $\mathbb{R}^n$ invariant par rotation et on considère le Lagrangien

$$L = \frac{1}{2}\|\dot{q}\|^2 - U(q).$$

1. *Intégrales premières.* Montrez que les fonctions

$$I_{ij} = \dot{q}_i q_j - \dot{q}_j q_i$$

sont des intégrales premières du mouvement. *Indication* : considérez les symétries données par les rotations dans les plans de coordonnées (i, j) . Pour $n = 3$, retrouvez l'invariance du moment cinétique, $I = q \wedge \dot{q}$. Montrez en général que si $q(t_0) \notin P$, que si P est un plan qui contient $q(t_0)$ et $\dot{q}(t_0)$, alors ce plan contient $q(t)$ et $\dot{q}(t)$ pour t au voisinage de t_0 . En déduire en général que les orbites ne passant pas par l'origine sont planaires. Que se passe-t-il pour les orbites passant par l'origine ?

2. On considère à partir de maintenant le cas de la dimension 2. On utilise les coordonnées polaires (r, θ) . Montrez que le moment cinétique $M = \dot{\theta}r^2$ est une intégrale première du mouvement.
3. On introduit la fonction *énergie potentielle effective*

$$V = U + \frac{M^2}{2r^2}.$$

L'introduction de cette fonction va permettre de réécrire le système comme un système en dimension 1. Plus précisément, montrez alors que l'énergie totale s'écrit

$$E = \frac{\dot{r}^2}{2} + V(r),$$

et que les équations du mouvement sont

$$\frac{d^2}{dt^2}r = -\frac{\partial}{\partial r}V.$$

4. *Remarque.* La présence d'une seule intégrale première nous a permis ici de réduire la dimension de l'espace des phases de 4 à 2, alors que l'on s'attendrait à un gain d'un seul paramètre. Ceci est dû à des symétries "cachées" dont l'interprétation est plus claire dans le cadre de la géométrie symplectique, cadre sous-jacent à la mécanique.
5. *Orbites circulaires.* Montrez que les orbites partant d'un point à distance r_0 , pour lesquelles $E = V(r_0)$ et $\frac{\partial}{\partial r}V = 0$, sont circulaires.
6. *Allure grossière des orbites.* Fixons une orbite. De la conservation du moment, déduire que l'angle θ varie de manière monotone le long de cette orbite. De la conservation de l'énergie, en déduire que r oscille entre deux valeurs r_{min} et r_{max} pour lesquelles $E = V(r)$, en autorisant le cas où ces bornes ne sont pas atteintes, auxquels cas r est monotone. Les points pour lesquels $r = r_{max}$ (resp. $r = r_{min}$) s'appellent *apocentres* (resp. *péricentres*). Si le centre est la Terre, le Soleil, ou la Lune, on parle d'*apogée*, d'*apohélie*, d'*aposélénie* etc.

13.2.2 Exemple et exercice, mouvement dans un champ central (bis)

Le but de cet exercice nettement plus difficile est de montrer que les seuls potentiels pour lesquelles les trajectoires sont toutes périodiques sont les potentiels $U = \frac{k}{r}$ potentiel de la gravitation, $U = kr^2$ potentiel de l'oscillateur harmonique (ressort). On déduit en particulier de la première loi de Kepler, *les orbites des planètes sont elliptiques*, et de l'hypothèse d'invariance par rotation du potentiel, la formule précise du potentiel de gravitation.

On suppose donc que toutes les orbites bornées sont périodiques et que $\frac{p}{\partial r}U$ est positif.

1. Montrez que l'angle entre un apocentre et péricentre successif s'écrit

$$\Theta = \int_{r_{min}}^{r_{max}} \frac{Mdr}{r^2 \sqrt{2(E - V(r))}}.$$

Si toutes les orbites sont périodiques Θ est commensurable à π . Par ailleurs, si l'on fait l'hypothèse raisonnable, assurée par des conditions de non dégénérescence de V , en dehors des orbites circulaires, que Θ dépend continument des orbites, en dehors des orbites circulaires, on en déduit que Θ est constant. Nous supposons par la suite que Θ est constant.

2. Montrez qu'il existe des orbites circulaires pour n'importe quel rayon r .
3. Montrer qu'au voisinage d'une orbite circulaire

$$\Theta \approx \Theta_c = \pi \sqrt{\frac{\frac{d}{dt}U}{3 \frac{d}{dt}U + r \frac{d^2}{dt^2}U}}.$$

Il sera utile d'utiliser le changement de variable $r = M/x$.

4. Toujours en utilisant ce dernier changement de variables, montrez que Θ_c est constant seulement pour les potentiels $U = ar^\alpha$ (pour $\alpha \geq -2, \alpha \neq 0$) et $U = b \log(r)$, et qu'alors $\Theta_c = \pi/\sqrt{\alpha+2}$ (le cas logarithmique correspond à $\alpha = 0$)
5. On suppose que $\lim_{r \rightarrow +\infty} U(r) = +\infty$ Montrez alors que

$$\lim_{E \rightarrow +\infty} \Theta(E, M) = \frac{\pi}{2}.$$

Indication On utilisera le changement de variables $y = r_{\min}/r$ pour écrire

$$\Theta = \int_{y_{\min}}^1 \frac{dy}{\sqrt{2(Z(1) - Z(y))}}, \quad Z(y) = \frac{y^2}{2} + \frac{M^2}{r_{\min}^2} U\left(\frac{r_{\min}^2}{My^2}\right).$$

13.2.3 Problème à deux corps

On considère le *problème à deux corps*. On se donne donc deux corps de masse m et M respectivement dans $\mathbb{R}^n$, ayant les positions q et Q , les vitesses $\dot{q}$ et $\dot{Q}$. Le Lagrangien du système est donné par

$$L = \frac{M\|\dot{Q}\|^2}{2} + \frac{m\|\dot{q}\|^2}{2} + \frac{k(m+M)}{\|q-Q\|}.$$

1. Montrez que le Lagrangien est invariant par translation. En déduire en utilisant le théorème de Noether, que *la vitesse du centre de gravité* $m\dot{q} + M\dot{Q}$ est une intégrale première du mouvement.
2. On suppose que cette vitesse est nulle, et on utilise donc des coordonnées dans lesquelles l'origine est le centre de gravité $m\dot{q} + M\dot{Q}$. Fixons i, j et considérons les rotations $R^{ij}(\theta)$ d'angle θ dans le plan de coordonnées i et j . Quelles sont les intégrales premières associées à l'invariance du système par ces rotations ? En déduire que le mouvement est planaire, à savoir que les deux corps pesants restent dans un plan fixe.
3. On suppose donc que $n = 2$. L'espace des positions du système est donc décrit par la position du petit corps que l'on repère en coordonnées polaires (r, θ) . Montrez que le Lagrangien du système est celui d'une particule dans un champ central

$$L = \frac{a\dot{r}^2}{2} + \frac{ar^2\dot{\theta}^2}{2} + \frac{b}{r}.$$

Donner également l'écriture de l'intégrale première I associée à l'invariance par rotation.

Chapitre 14

Mécanique Hamiltonienne

Nous allons montrer comment la connaissance de suffisamment d'intégrales premières permet de décrire les solutions des équations d'Hamilton, une version duale des équations d'Euler-Lagrange, dans le cadre des systèmes complètement intégrables. Un outil de nature "algébrique", le crochet de Poisson, facilitera cette étude. Ce même crochet de Poisson permet aussi de rechercher les intégrales premières. Il nous aidera aussi à comprendre les changements de variables - la seule manière d'intégrer des équations différentielles - adaptés au problème.

14.1 Equations d'Hamilton

Rappelons les équations de Hamilton. Elles sont liées à une fonction H , le *Hamiltonien*, définie sur $\mathbb{R}^{2n}$. Si les fonctions coordonnées de $\mathbb{R}^{2n}$ sont (q_i, p_i) , elles s'écrivent

$$\begin{cases} \frac{d}{dt}p &= -\frac{\partial}{\partial q}H \\ \frac{d}{dt}q &= \frac{\partial}{\partial p}H. \end{cases}$$

On introduit le champ de vecteur associé à cette équation différentielle appelé aussi *champ hamiltonien*, ou *gradient hamiltonien* ou encore *gradient symplectique*

$$X_H = \left(\frac{\partial}{\partial p}H, -\frac{\partial}{\partial q}H \right).$$

14.1.1 Transformation de Legendre

Signalons rapidement le lien entre les équations de Lagrange et celles d'Hamilton.

Dans certains cas, par exemple pour les Lagrangiens du type

$$L = T - U = \frac{m}{2} \|\dot{q}\|^2 - U(q),$$

la *transformation de Legendre*, donnée par

$$F : (q, \dot{q}) \mapsto (q, p) = (q, \frac{\partial}{\partial \dot{q}} L(q, \dot{q})).$$

est un difféomorphisme. Ce difféomorphisme va nous permettre de relier, sous la forme d'un exercice, les équations de Lagrange et celles d'Hamilton.

Exercice : Lagrange et Hamilton

On suppose que la transformée de Legendre F est un difféomorphisme. Posons $H(q, p) = p\dot{q} - L(q, \dot{q})$, pour $p = \frac{\partial}{\partial \dot{q}} L(q, \dot{q})$. Par hypothèse, la fonction H est alors parfaitement définie.

1. Montrez tout d'abord que

$$\frac{\partial H}{\partial p}(q, \frac{\partial L}{\partial \dot{q}}) = \dot{q},$$

2. Montrez ensuite

$$\frac{\partial H}{\partial q}(q, \frac{\partial L}{\partial \dot{q}}) = -\frac{\partial L}{\partial q}(q, \dot{q}).$$

3. Montrez enfin que la courbe $(q(t), \dot{q}(t))$ est solution des équations de Lagrange si et seulement si la courbe

$$(q(t), p(t)) = F(q(t), \dot{q}(t))$$

est solution des équations d'Hamilton.

4. Montrez ensuite que l'invariance au cours du temps de $E = T + U$ dans les équations de Lagrange s'interprète comme l'invariance de H dans les équations de Hamilton.

Les deux formulations sont donc - dans les cas intéressants - équivalentes. L'intérêt de la formulation de Hamilton provient de ce qu'elle possède plus de symétries. Nous allons voir de manière plus précise que le formalisme hamiltonien est plus susceptible de calculs "formels" à l'aide du crochet de Poisson. Ce formalisme possède également plus de changements de variables pertinents.

Application : le problème à deux corps

On reprend le Lagrangien apparaissant dans le problème à deux corps [13.2.3](#).

$$L = \frac{a\dot{r}^2}{2} + \frac{ar^2\dot{\theta}^2}{2} + \frac{b}{r}.$$

En utilisant la transformation de Legendre, écrire le Hamiltonien correspondant en fonction des coordonnées $(r, \theta, p_r, p_\theta)$. Montrez alors que l'invariance de I pour la formulation lagrangienne, induit celle de p_θ pour la formulation hamiltonienne.

14.2 Crochet de Poisson

Cet outil fondamental permet d'étudier la mécanique hamiltonienne. Soient f et g deux fonctions définies sur $\mathbb{R}^{2n}$, définissons leur *crochet de Poisson*, noté $\{f, g\}$ comme

$$\{f, g\} = \sum_i \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i}.$$

Remarquons que le choix du signe est soumis à des conventions variables suivant les auteurs, voire à l'intérieur d'un même ouvrage ...

Voici quelques propriétés algébriques importantes du crochet de Poisson.

Proposition 14.2.0.1 *Nous avons*

1. *Antisymétrie* $\{f, g\} = -\{g, f\}$,
2. *Dérivation* $\{fh, g\} = f\{h, g\} + h\{f, g\}$,
3. *Identité de Jacobi* $\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0$,
4. $\{f, g\} = -df(X_g) = dg(X_f)$
5. f est une intégrale première du système associé à l'hamiltonien H si et seulement si $\{f, H\} = 0$.
6. $X_{\{f, g\}} = [X_f, X_g]$. En particulier, si $\{f, g\}$ est une fonction constante alors les flots de X_f et X_g commutent.

DÉMONSTRATION : Les propriétés (i) et (ii) sont des vérifications immédiates. L'identité de Jacobi (iii) est plus longue à vérifier mais est purement technique. la propriété (iv) est également immédiate et entraîne (v) immédiatement. La propriété (vi) découle de l'identité de Jacobi. En effet, de celle-ci on en déduit que pour toute fonction h

$$\{h, \{f, g\}\} = \{\{h, f\}, g\} - \{\{h, g\}, f\}.$$

et donc

$$dh(X_{\{f, g\}}) = -X_g X_f . h + X_f X_g . h = dh([X_f, X_g]).$$

Comme ceci est vrai pour tout h , nous en déduisons bien le résultat. $\square$

14.2.1 Algèbre des intégrales premières

Nous en déduisons des propriétés algébriques de l'ensemble des intégrales premières. Formellement les trois premières propriétés donnent à l'ensemble des fonctions la structure d'une *Algèbre de Poisson*. La dernière nous assure que l'ensemble des intégrales premières est une sous-algèbre, ce qui se traduit dans notre cas par la propriété suivante : si f et g sont des intégrales premières alors non seulement fg , $f + g$ sont des intégrales premières mais aussi $\{f, g\}$.

Le crochet de Poisson permet d'algébriquer la recherche d'intégrales premières. Par exemple, on commence par vérifier que l'on a

$$\{p_i, p_j\} = \{q_i, q_j\} = 0, \quad \{p_i, q_j\} = \delta_{ij}.$$

Nos règles de calcul nous permettent alors de calculer $\{p_i^n, q_j^m\}$ et donc le crochet de Poisson de tout polynôme en les variables p et q . Physiquement, les fonctions sont des polynômes ... ou au pire des séries formelles

$$f = \sum_{i_1, \dots, i_l, j_1, \dots, j_k} a_{i_1, \dots, i_l, j_1, \dots, j_k} p_{i_1} \dots p_{i_l} q_{j_1} \dots q_{j_k}.$$

Le problème de recherche d'intégrale première devient alors algébrique : la recherche de solutions d'équations linéaires en les $a_{i_1, \dots, i_l, j_1, \dots, j_k}$. Remarquons tout de même qu'il n'en devient pas nécessairement plus facile, mais il peut au moins se résoudre algorithmiquement de manière approximative.

Exercice : le problème à deux corps

On considère le Hamiltonien du problème à deux corps

$$H = \frac{p_r^2}{2a} + \frac{p_\theta^2}{2ar^2} + \frac{b}{r}.$$

Montrez directement que $\{H, p_\theta\} = 0$. Retrouvez ainsi le fait que p_θ est une intégrale première du mouvement.

14.2.2 Evolution des observables

Le crochet de Poisson va nous fournir une manière "intrinsèque" de décrire les équations de Hamilton. Si f est une fonction, ϕ_t le flot du gradient hamiltonien, on remarque alors que

$$\frac{d}{dt} f \circ \phi_t = \{H, f\}.$$

Ce que l'on écrit de manière abrégée, sous la forme

$$\dot{f} = \{H, f\}.$$

L'information contenue dans toutes ces équations permet de retrouver les équations d'Hamilton. Si en effet ϕ_t est un groupe à un paramètre de difféomorphismes, qui satisfait que pour toute fonction f ,

$$\dot{f} = \{H, f\},$$

On vérifie aisément en exercice que ϕ_t est le flot du gradient hamiltonien de H .

Grâce à cette remarque, on arrive donc à définir les équations de Hamilton sans utiliser les coordonnées p et q . Ceci ouvre la voie à l'extension de la définition des équations de Hamilton sur les sous-variétés et plus généralement sur les variétés. Extension qui nous mènerait trop loin pour cette introduction.

Remarquons que cette dernière équation décrit l'information physique que l'on cherche à avoir sur le système : l'évolution de la fonction, ou observable, ou mesure effectuée sur le système. Enfin, la ressemblance formelle de cette équation avec l'équation de Shrödinger n'est pas un hasard. Le processus qui permet - dans les bons cas et difficilement formalisable - de remonter du système classique au système quantique sous-jacent utilise des manipulations algébriques sur cette équation et l'algèbre de Poisson des fonctions.

14.3 Coordonnées symplectiques

Dans la résolution des équations différentielles un rôle royal est réservé au changement de variables : on cherche toujours les variables dans lesquelles l'équation se ramène à une forme connue, ou plus simple.

Dans les équations d'Hamilton, le cas où le Hamiltonien ne dépend que des variables q est facile à résoudre. Nous allons tout d'abord examiner les coordonnées symplectiques, c'est-à-dire d'autres variables que les coordonnées de départ p et q , dans lesquelles les équations d'Hamilton, et plus précisément le crochet de Poisson, s'expriment simplement. Ce seront les changements de variables adaptés à la mécanique Hamiltonienne.

14.3.1 Variables en involution

Occupons nous tout d'abord des variables de type q et cherchons à les généraliser.

Soit U un ouvert de $\mathbb{R}^{2n}$. Nous dirons que les l fonctions Q_i définies sur U sont en *involution* si les crochets de Poisson $\{Q_i, Q_j\}$ sont nuls.

Nous dirons que ces l variables sont *non dégénérées* si les formes dQ_i sont linéairement indépendantes. Montrons la proposition suivante.

Proposition 14.3.1.1 *Soit Q_i des variables en involution. Soit H une fonction telle que $H = h(Q_1, \dots, Q_l)$. Alors $\{H, Q_i\} = 0$.*

DÉMONSTRATION : Ceci découle de manipulations sur le crochet de Poisson

$$\begin{aligned}
 \{H, Q_j\} &= \sum_i \frac{\partial H}{\partial p_i} \frac{\partial Q_j}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial Q_j}{\partial p_i} \\
 &= \sum_{i,k} \frac{\partial h}{\partial Q_k} \frac{\partial Q_k}{\partial p_i} \frac{\partial Q_j}{\partial q_i} - \frac{\partial h}{\partial Q_k} \frac{\partial Q_k}{\partial q_i} \frac{\partial Q_j}{\partial p_i} \\
 &= \sum_k \frac{\partial h}{\partial Q_k} \{Q_j, Q_k\} \\
 &= 0.
 \end{aligned}$$

□

14.3.2 Variables conjuguées

Construisons des variables de type p , associées à nos variables de type q .
Des variables en involution permettent d'en définir d'autres :

Théorème 14.3.2.1 *Soit Q_i , n variables non dégénérées en involution définies sur $U \subset \mathbb{R}^{2n}$. Pour tout x de U , il existe alors n fonctions P_i , définies au voisinage de x , et telles que $\{P_i, P_j\} = 0$, $\{P_i, Q_j\} = \delta_{ij}$.*

Les fonctions P_i s'appellent *variables conjuguées* aux fonctions Q_i . La mécanique classique parle parfois de variables *angles* pour les Q_i et *actions* pour les P_i .

Le théorème nous assure de l'existence de variables conjuguées. Par contre, la méthode de construction présente dans la preuve n'est pas nécessairement très utile. Une autre approche est donnée par la méthode d'Hamilton-Jacobi que nous exposerons à la fin de ce chapitre

DÉMONSTRATION : Nous allons démontrer la première partie de ce résultat par récurrence et nous supposons donc que nous avons construit P_i $i = 1, \dots, k$ telles que $(dQ_1, \dots, dQ_n, dP_1, \dots, dP_k)$ soient linéairement indépendants et de plus

$$\{Q_i, Q_j\} = \{P_i, P_j\} = 0, \quad \{P_i, Q_j\} = \delta_{ij}.$$

Notons φ_t^i les flots des $X_{f_i} = X_i$, et φ_{n+j} ceux des $X_{g_j} = X_{n+j}$. Nous savons que tous ces difféomorphismes commutent (cf (vi) de 14.2.0.1). De plus, la condition de non-dégénérescence assure que pour tout y de U , l'espace vectoriel

$$V_y = \sum_i \mathbb{R}X_i(y)$$

est de dimension $n + k$.

Nous pouvons trouver une sous-variété W de dimension $n - k$ - par exemple un ouvert d'un sous-espace vectoriel - passant par x , et telle que pour tout y ,

$$T_y W \oplus V_y = \mathbb{R}^{2n}. \quad (14.1)$$

Considérons alors l'application Ψ définie de la manière suivante.

$$\begin{cases} W \times \mathbb{R}^{n+k} & \rightarrow \mathbb{R}^{2n} \\ (w, s_i, \dots, s_{n+k}) & \mapsto \varphi_{s_1}^1 \circ \dots \circ \varphi_{s_{n+k}}^k(w). \end{cases}$$

L'application Ψ est un difféomorphisme au voisinage de x . Grâce à 14.1, on montre en effet que $D_x \Psi$ est surjective et donc inversible.

Notons enfin π , l'application

$$\begin{cases} W \times \mathbb{R}^{n+k} & \rightarrow \mathbb{R} \\ (w, s_i, \dots, s_{n+k}) & \mapsto s_{k+1}. \end{cases}$$

Posons

$$P_{k+1} = \pi \circ \Psi^{-1}.$$

Les flots des champs de vecteurs que nous avons définis commutent tous et nous pouvons en déduire que la fonction P_{k+1} vérifie

$$P_{k+1} \circ \varphi_s^i = s\delta_{ik} + P_{k+1}$$

En dérivant ces dernières relations, nous allons bien obtenir les relations de commutation souhaitées.

Il nous reste à montrer que dP_{k+1} n'appartient pas à l'espace vectoriel engendré par $dQ_1, \dots, dQ_n, dP_1, \dots, dP_k$. Pour cela il suffit de remarquer que

$$dP_{k+1}(X_{n+k+1}) = 1,$$

alors que, si $i \leq k$ et $j \leq n$,

$$dP_i(X_{n+k+1}) = dQ_j(X_{n+k+1}) = 0.$$

□

Exercice

Adaptez la preuve ci-dessus pour montrer que l'on peut toujours compléter une familles de variables en involution non-dégénérées.

14.3.3 Coordonnées symplectiques

Nous dirons que la famille de fonctions $(Q_1, \dots, Q_n, P_1, \dots, P_n)$ forment des *coordonnées symplectiques* au voisinage d'un point x de $\mathbb{R}^{2n}$ si nous avons les relations

$$\{P_i, P_j\} = \{Q_i, Q_j\} = 0, \quad \{P_i, Q_j\} = \delta_{ij}.$$

D'après ce que nous venons de voir au sujet des variables conjuguées, il suffit de trouver n fonctions en involution non dégénérées pour construire des coordonnées symplectiques.

Le résultat suivant justifie la terminologie "coordonnées" et permet de comprendre l'utilité de ces coordonnées

Théorème 14.3.3.1 *Soit $(Q_1, \dots, Q_n, P_1, \dots, P_n)$ des coordonnées symplectiques, alors*

- l'application $G = (Q_1, \dots, Q_n, P_1, \dots, P_n)$ est un difféomorphisme local,
- De plus, si $H = h(Q) = h(Q_1, \dots, Q_n)$ alors

$$\begin{cases} \{Q_i, H\} &= 0, \\ \{P_i, H\} &= \frac{\partial h}{\partial Q_i}(Q). \end{cases}$$

Ce théorème a une conséquence utile sur l'intégration des équations d'Hamilton.

Corollaire 14.3.3.2 *Soit H un Hamiltonien défini sur $\mathbb{R}^{2n}$. Soit ϕ_t son flot. Notons $f^t = f \circ \phi_t$. Soit $(Q_1, \dots, Q_n, P_1, \dots, P_n)$ des coordonnées symplectiques. Supposons qu'il existe une fonction h telle que $H = h \circ (Q_1, \dots, Q_n)$. Alors*

$$\begin{cases} P_i^t &= P_i - t \frac{\partial h}{\partial Q_i}(Q), \\ Q_i^t &= Q_i. \end{cases}$$

Démonstration du théorème. Nous avons déjà montré la moitié de [2], l'autre partie se montre par la même méthode. $\square$

Démonstration du corollaire. D'après le théorème, nous avons

$$\begin{aligned} \dot{Q}_i = \{Q_i, H\} &= 0, \\ \dot{P}_i = \{P_i, H\} &= -\frac{\partial h}{\partial Q_i}(Q). \end{aligned}$$

Les Q_i sont donc des intégrales premières. Les fonctions $\frac{\partial h}{\partial Q_i}(Q)$ sont elles aussi des intégrales premières. Nous en déduisons $\ddot{P}_i = 0$. Ceci nous permet de conclure. $\square$

14.4 Systèmes complètement intégrables

Nous allons montrer comment la connaissance de suffisamment d'intégrales premières permet de décrire le système en choisissant des coordonnées symplectiques adaptées.

14.4.1 Systèmes complètement intégrables

Nous allons examiner le cas le plus optimiste, celui où le système hamiltonien possède suffisamment d'intégrales premières.

Nous dirons que le hamiltonien H défini sur $\mathbb{R}^{2n}$ est *complètement intégrable* (sur un ouvert U) s'il possède n intégrales premières non dégénérées en involution. Autrement dit, s'il existe n fonctions f_i définies sur U et telles que

1. les formes df_i sont linéairement indépendantes,
2. les crochets de Poisson $\{f_i, f_j\}$ et $\{f_i, H\}$ sont nuls.

Attention : cette définition n'est pas stable dans la littérature. Certains auteurs autorisant un ensemble (petit) où (i) cesse d'être vrai. D'autres parlent de systèmes complètement intégrables lorsqu'ils sont complètement intégrable, dans le sens ci-dessus, au voisinage de tout point. Nous avons choisi la définition la plus contraignante.

Exercice

Montrez que le problème à deux corps est complètement intégrable.

14.4.2 Méthode d'Hamilton-Jacobi

La méthode d'Hamilton-Jacobi cherche à rendre un système complètement intégrable. C'est un procédé qui permet, dans les bons cas, de trouver le changement de variables - les coordonnées symplectiques - le plus adapté.

On se donne donc un hamiltonien H défini sur un ouvert de $\mathbb{R}^n$ muni de ses coordonnées symplectiques naturelles $(q_1, \dots, q_n, p_1, \dots, p_n)$.

On dit que la fonction $S = S(q, Q)$ définie sur $\mathbb{R}^{2n}$, muni cette fois ci des coordonnées (q_i, Q_i) est une solution de l'équation d'Hamilton-Jacobi associée à H s'il existe une fonction h telle que

$$H(q_i, \frac{\partial S}{\partial q_i}(q_i, Q_i)) = h(Q_i).$$

Il faut considérer ici S comme une fonction des q_i et de paramètres supplémentaires Q_i .

On dira aussi que S est *non dégénérée* en (q_0, Q_0) si

$$\det(\frac{\partial^2 S}{\partial q_i \partial Q_j})|_{q_0, Q_0} \neq 0.$$

Le théorème d'Hamilton-Jacobi montre l'intérêt des solutions de l'équation d'Hamilton-Jacobi.

Théorème 14.4.2.1 *Soit S une solution non dégénérée de l'équation d'Hamilton-Jacobi pour H . Il existe alors des coordonnées symplectiques Q_i, P_i définies au voisinage de*

$$(q_0, \frac{\partial}{\partial q_i} S|_{(q_0, Q_0)})$$

telles que

$$\begin{cases} p_i &= \frac{\partial}{\partial q_i} S(q_i, Q_i), \\ P_i &= -\frac{\partial}{\partial Q_i} S(q_i, Q_i). \end{cases}$$

En particulier, H est complètement intégrable.

Signalons que la fonction S s'appelle *fonction génératrice* des coordonnées symplectiques (Q, P) . Ce théorème nous permet donc de construire des coordonnées symplectiques beaucoup plus sympathique et de manière plus directe que la preuve du théorème 14.3.2.1. Par contre, comme il faut s'y attendre, trouver des solutions de l'équation d'Hamilton-Jacobi est difficile. On y arrive dans certains cas. Nous verrons un exemple extrêmement simple et renvoyons au livre d'Arnol'd [3] pour des exemples plus intéressants, en particulier celui du *dipôle gravitique* qui modélise la trajectoire d'une particule massive dans deux puits de potentiels gravitationnels.

DÉMONSTRATION : Montrons tout d'abord qu'il existe des fonctions Q_i telles que

$$p_i = \frac{\partial S}{\partial q_i}(q_i, Q_i).$$

Ceci est une conséquence du théorème des fonctions implicites et de la condition de non dégénérescence pour S . L'existence des fonctions P_i découle alors de leur définition. Il nous reste à calculer les crochets de Poisson.

Considérons la fonction $W = S(q, Q)$ vu comme fonction de q et p . Nous obtenons les relations suivantes

$$\begin{aligned}\sum_i P_i \frac{\partial Q_i}{\partial p_m} &= \frac{\partial W}{\partial p_m} \\ \sum_i P_i \frac{\partial Q_i}{\partial q_l} &= \frac{\partial W}{\partial q_l} + p_l.\end{aligned}$$

en dérivant en haut par rapport à q_l , en bas par rapport à p_m et en utilisant l'identité de Schwarz, il vient

$$\sum_i \left(\frac{\partial P_i}{\partial q_l} \frac{\partial Q_i}{\partial p_m} - \frac{\partial P_i}{\partial p_m} \frac{\partial Q_i}{\partial q_l} \right) = \delta_{lm}$$

Des manipulations analogues donnent

$$\begin{aligned}\sum_i \left(\frac{\partial P_i}{\partial p_l} \frac{\partial Q_i}{\partial p_m} - \frac{\partial P_i}{\partial p_m} \frac{\partial Q_i}{\partial p_l} \right) &= 0 \\ \sum_i \left(\frac{\partial P_i}{\partial q_l} \frac{\partial Q_i}{\partial q_m} - \frac{\partial P_i}{\partial q_m} \frac{\partial Q_i}{\partial q_l} \right) &= 0.\end{aligned}$$

Traduisons ces trois dernières équations de manière algébrique. Introduisons les matrices

$$A = \begin{pmatrix} \frac{\partial P}{\partial p} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial Q}{\partial q} \end{pmatrix}, \quad J = \begin{pmatrix} 0 & -\text{Id} \\ \text{Id} & 0 \end{pmatrix}.$$

Notons B^* la transposée de la matrice B . Les trois équations que nous venons d'obtenir s'écrivent

$$(JA)(JA)^* = \text{Id}.$$

Nous en déduisons

$$(JA^*)(JA^*)^* = \text{Id}.$$

Les équations correspondantes nous donnent

$$\begin{aligned}\sum_i \left(\frac{\partial Q_l}{\partial q_i} \frac{\partial P_m}{\partial p_i} - \frac{\partial Q_l}{\partial p_i} \frac{\partial P_m}{\partial q_i} \right) &= \delta_{lm} \\ \sum_i \left(\frac{\partial P_m}{\partial p_i} \frac{\partial P_l}{\partial q_i} - \frac{\partial P_m}{\partial q_i} \frac{\partial P_l}{\partial p_i} \right) &= 0 \\ \sum_i \left(\frac{\partial Q_m}{\partial p_i} \frac{\partial Q_l}{\partial q_i} - \frac{\partial Q_m}{\partial q_i} \frac{\partial Q_l}{\partial p_i} \right) &= 0\end{aligned}$$

Ces équations nous donnent les crochets de Poisson que nous recherchons. $\square$

La méthode d'Hamilton-Jacobi nous permet d'une part de montrer qu'un Hamiltonien est complètement intégrable, mais aussi de donner des coordonnées symplectiques adaptées au problème. Cette méthode est efficace quand les variables ont des séparables. Nous allons la voir fonctionner sur un exemple.

Le problème à deux corps

Nous considérons l'Hamiltonien du problème à deux corps,

$$H = \frac{p_r^2}{2a} + \frac{p_\theta^2}{2ar^2} + \frac{b}{r}.$$

L'équation d'Hamilton-Jacobi correspondante est

$$\left(\frac{\partial S}{\partial r}\right)^2 \frac{1}{2a} + \left(\frac{\partial S}{\partial \theta}\right)^2 \frac{1}{2ar^2} + \frac{b}{r} = Q_1.$$

Equation que nous écrivons.

$$\left(\frac{\partial S}{\partial \theta}\right)^2 = 2aQ_1r^2 - 2abr - \left(\frac{\partial S}{\partial r}\right)^2r^2.$$

Nous avons séparé les variables r et θ - c'est-à-dire écrit l'équation sous la forme $A(r) + B(\theta) = 0$ - preuve que ces variables étaient *séparables* dans l'équation de départ. Nous cherchons alors S sous la forme $S = M(r) + N(\theta)$. Nous obtenons alors

$$\begin{aligned} \frac{\partial S}{\partial \theta} &= \frac{\partial N}{\partial \theta} = Q_2 \\ \frac{\partial S}{\partial r} &= \frac{\partial M}{\partial r} = \frac{1}{r} \sqrt{Q_2^2 + 2abr - 2aQ_1r^2} \end{aligned}$$

Il faut maintenant expliciter les nouvelles coordonnées symplectiques. Nous avons choisi

$$Q_1 = \left(\frac{\partial S}{\partial r}\right)^2 \frac{1}{2a} + \left(\frac{\partial S}{\partial \theta}\right)^2 \frac{1}{2ar^2} + \frac{b}{r} = H.$$

De même

$$Q_2 = \frac{\partial S}{\partial \theta} = p_\theta.$$

Nous retrouvons sans surprise nos deux intégrales premières habituelles.

Intéressons nous aux variables conjuguées

$$\begin{aligned} P_1 &= -\frac{\partial S}{\partial Q_1} \\ P_2 &= -\frac{\partial S}{\partial Q_2}. \end{aligned}$$

Nous avons

$$\begin{aligned}\frac{\partial P_1}{\partial r} &= -\frac{\partial^2 S}{\partial r \partial Q_1} = \frac{a}{\sqrt{Q_2^2 + 2abr - 2aQ_1 r^2}} \\ \frac{\partial P_1}{\partial \theta} &= -\frac{\partial^2 S}{\partial \theta \partial Q_1} = 0.\end{aligned}$$

Ainsi

$$P_1 = \int \frac{a}{\sqrt{Q_2^2 + 2abr - 2aQ_1 r^2}} dr.$$

De même, nous obtenons

$$\begin{aligned}\frac{\partial P_2}{\partial r} &= -\frac{\partial^2 S}{\partial r \partial Q_2} = -\frac{Q_2}{r\sqrt{Q_2^2 + 2abr - 2aQ_1 r^2}} \\ \frac{\partial P_2}{\partial \theta} &= -\frac{\partial^2 S}{\partial \theta \partial Q_1} = -1.\end{aligned}$$

Et donc

$$\begin{aligned}P_2 &= -\theta - \int \frac{Q_2}{r\sqrt{Q_2^2 + 2abr - 2aQ_1 r^2}} dr \\ &= -\theta + \int \frac{Q_2}{\sqrt{\frac{Q_2^2}{r^2} + \frac{2ab}{r} - 2aQ_1}} d\left(\frac{1}{r}\right) \\ &= -\theta + \operatorname{argsh}\left(\frac{\alpha}{r} + \beta\right) \text{ ou } -\theta + \arcsin\left(\frac{\alpha}{r} + \beta\right)\end{aligned}$$

Les fonctions P_1 et P_2 sont donc des fonctions “classiques”, leurs valeurs exactes dépendant des constantes de départ. La possession de ces variables conjugués et du corollaire 14.3.3.2, nous permet de calculer numériquement l’évolution du système. Bien sûr, nous n’avons pas choisi ici le chemin le plus direct, mais celui qui illustre la méthode.

Remarquons déjà que dans nos nouvelles coordonnées $H = Q_1$, en particulier P_2 est une intégrale première du mouvement. La connaissance de cette nouvelle intégrale première nous donne alors l’équation des trajectoires en coordonnées polaires.

En effet, on obtient suivant les cas les relations suivantes

$$\sin(\theta - \theta_0) = \frac{\alpha}{r} + \beta$$

ou

$$\sinh(\theta - \theta_0) = \frac{\alpha}{r} + \beta$$

Demi-plan supérieur

On considère le *demi-plan supérieur*

$$U = \{(x, y)/y > 0\}.$$

On considère le Lagrangien défini sur $U \times \mathbb{R}^2$ par

$$L(x, y, \dot{x}, \dot{y}) = \frac{1}{2y^2}(\dot{x}^2 + \dot{y}^2).$$

On s'intéresse tout d'abord à la transformée de Legendre. Elle est donnée par

$$\mathcal{L}(\dot{x}, \dot{y}) = (p_x, p_y) = \left(\frac{\dot{x}}{y^2}, \frac{\dot{y}}{y^2}\right).$$

Le Hamiltonien dans ces variables est donc

$$H = p_x \dot{x} + p_y \dot{y} - L(x, y, \dot{x}, \dot{y}) = \frac{y^2}{2}(p_x^2 + p_y^2).$$

En particulier, p_x est une intégrale première.

On s'intéresse maintenant à la solution $S = S(x, y, a, b)$ de l'équation d'Hamilton-Jacobi associée à H . Celle-ci est

$$y^2 \left(\left(\frac{\partial S}{\partial x} \right)^2 + \left(\frac{\partial S}{\partial y} \right)^2 \right) = a.$$

On peut séparer les variables dans cette équation. Elle est équivalente à

$$\left(\frac{\partial S}{\partial x} \right)^2 + \left(\left(\frac{\partial S}{\partial y} \right)^2 - \frac{a}{y^2} \right) = 0.$$

Equation que l'on peut résoudre en posant $S = A(y) + B(x)$

$$\begin{aligned} \frac{\partial B}{\partial x} &= b \iff B(x) = bx \\ \frac{\partial A}{\partial y} &= \pm \sqrt{\frac{a}{y^2} - b^2}. \end{aligned}$$

On se souviendra que $a \geq 0$ et que $\frac{a}{y^2} \geq b^2$. D'après le théorème d'Hamilton-Jacobi, les variables a et b en fonction de x, y, p_x et p_y sont solutions des équations

$$\frac{\partial S}{\partial x}(x, y, a, b) = p_x \quad \frac{\partial S}{\partial y}(x, y, a, b) = p_y.$$

On a donc bien $b = p_x$. Par ailleurs, par construction on a $H = 2a$.

On s'intéresse maintenant aux variables conjuguées p_a, p_b des variables a et b , données par le théorème d'Hamilton-Jacobi. Les variables conjuguées sont données par

$$\begin{aligned} p_a &= -\frac{\partial S}{\partial a} \\ p_b &= -\frac{\partial S}{\partial b}. \end{aligned}$$

On dérive

$$\begin{aligned}\frac{\partial p_a}{\partial y} &= -\frac{\partial^2 S}{\partial a \partial y} = \mp \frac{1}{2y\sqrt{a-b^2y^2}} \\ \frac{\partial p_a}{\partial x} &= 0 \\ \frac{\partial p_b}{\partial x} &= -1 \\ \frac{\partial p_b}{\partial y} &= \pm \frac{by}{\sqrt{a-b^2y^2}}.\end{aligned}$$

On obtient donc

$$p_b = -x + f(a, b) \mp \sqrt{\frac{a}{b^2} - y^2}.$$

Dans les nouvelles coordonnées symplectiques, on a par construction $H = 2a$. Donc $\frac{\partial H}{\partial b} = 0$; Les équations de Hamilton donnent

$$\frac{dp_b}{dt} = -\frac{\partial H}{\partial b} = 0.$$

Ainsi p_b est une intégrale première du mouvement. On en déduit que pour toute courbe (x, y) solution des équations d'Euler-Lagrange, on a

$$-x \mp \sqrt{\frac{a}{b^2} - y^2} = K,$$

et donc

$$(x - K)^2 + y^2 = \frac{a}{b^2},$$

qui est bien l'équation d'un cercle dont le centre est sur l'axe réel.

14.4.3 Au delà des systèmes complètement intégrables

En complément, nous indiquons brièvement l'importance de l'étude des systèmes complètement intégrables.

Enonçons tout d'abord un résultat important sur les systèmes totalement intégrables. Le théorème d'Arnol'd-Liouville.

Théorème 14.4.3.1 *Soit H un Hamiltonien complètement intégrable, et $Q_1, \dots, Q_n$ ses intégrales premières en involution. On suppose que $M_c = Q^{-1}\{c\}$ est compact. Alors M_c est un tore. De plus si φ^i sont les flots des gradients hamiltoniens des Q_i , si x_0 appartient à M_c , alors l'application*

$$\begin{cases} \mathbb{R}^n & \rightarrow & M_c \\ (s_1, \dots, s_n) & \mapsto & \varphi_{s_1}^1 \circ \dots \circ \varphi_{s_n}^n(x_0), \end{cases}$$

est un revêtement. Dans ces coordonnées sur M_c , le mouvement est linéaire.

Dans ces conditions, on parle de mouvement *quasi-périodique*. Les tores de niveaux invariants sont les *tors de Liouville*.

Cette situation est bien sûr très agréable lorsque l'on cherche explicitement les solutions. Malheureusement, un système n'est qu'exceptionnellement complètement intégrable. Poincaré a ainsi montré que le problème à trois corps sous l'influence mutuelle de la gravitation n'est pas complètement intégrable.

Malgré tout, être proche d'un système complètement intégrable est une condition importante : le célèbre théorème KAM - pour Kolmogorov-Arnol'd-Moser - assure que pour des systèmes hamiltoniens proches des systèmes complètement intégrables, certains des tores invariants persistent et de plus le mouvement y est quasi-périodique. Par exemple, pour le problème à trois corps Terre-Lune-Soleil - que l'on considère comme une perturbation du système Terre-Soleil - certaines des orbites restent sur des tores, et en particulier la Terre et le Soleil ne s'éloignent pas indéfiniment.

Signalons enfin que pour les systèmes complètement intégrables, le processus de *quantification* qui permet de remonter du système classique au système quantique sous-jacent est mathématiquement justifiable. Une méthode de quantification reste également disponible pour les systèmes proches des systèmes complètement intégrables.

14.5 Toupie de Lagrange Mouvement du solide

Nous traiterons sous la forme d'un exercice les mouvements du solide.

Commençons par rappeler quelques notations et résultats sur $SO(3)$. Le *groupe de Lie* est l'ensemble des matrices orthogonales 3×3 , c'est-à-dire en notant A^* la transposée de la matrice A ,

$$SO(3) = \{A \in M_3(\mathbb{R}) / AA^* = \text{Id}\}.$$

Le groupe $SO(3)$ est une sous-variété de l'espace des matrices 3×3 , noté $M_3(\mathbb{R})^3 = \mathbb{R}^9$. Son espace tangent $TSO(3)$ s'identifie naturellement à $SO(3) \times F$ où

$$F = SO(3) = \{A \in M_3(\mathbb{R}) / A = -A^*\}$$

est l'*algèbre de Lie* de $SO(3)$. L'identification est construite de la manière suivante : pour tout q de $SO(3)$, $T_q SO(3)$ est un sous-espace vectoriel de $M(3, \mathbb{R})$. L'identification est donnée par $(q, A) \in SO(3) \times F$ on associe $qA \in T_q SO(3)$. Dans cette identification, A est appelé *vitesse angulaire*.

Dans cette identification l'application tangente à L_g - la multiplication à gauche par g - est donnée par $TL_g(w, A) = (gw, A)$. De même, l'application tangente à R_g - la multiplication à droite par g - est donnée par $TR_g(w, A) = (wg, g^{-1}Ag)$.

Enfin, F s'identifie à $\mathbb{R}^3$ par l'identification qui à un vecteur u associe l'endomorphisme antisymétrique $v \mapsto u \wedge v$

On considère un solide ayant un point fixe à l'origine de $\mathbb{R}^3$ (pensez à une toupie, le point fixe étant le point de contact avec le sol). On va étudier son mouvement dans un champ de gravité d'intensité g . On étudiera en particulier le cas, *toupie de Lagrange*, où cette toupie a un axe de symétrie vertical dans la position de référence. On prendra comme référence une position du solide dans lequel le centre de gravité est sur l'axe vertical en un point G , O sera le point par lequel la toupie est fixée. On note enfin e_z le champ de vecteur unitaire vertical.

Description Lagrangienne du problème

1. Montrez que l'espace des configurations du solide s'identifie au groupe $SO(3)$.
2. Soit donc q une position du solide et u un vecteur de $\mathbb{R}^3$ vue comme un élément de $T_q SO(3)$. Ce vecteur représente une "vitesse du solide". Exprimez alors la vitesse d'un point M du solide en fonction de u et de $\overrightarrow{OM}$.
3. On veut maintenant décrire le Lagrangien associé à l'énergie cinétique. Physiquement, si on a un ensemble S - continu - de particules affectées d'une distribution de masse dm et d'une distribution de vitesse v , l'énergie cinétique du système est alors

$$T = \int_S \frac{v^2 dm}{2}.$$

Dans le cas d'un solide, montrez que le lagrangien T associé est alors donné par

$$T(q, u) = \langle u, \mathcal{I}u \rangle$$

où $\mathcal{I}$ est une matrice symétrique définie positive - la *matrice d'inertie* du solide - et $\langle, \rangle$ désigne le produit scalaire usuel de $\mathbb{R}^3$.

4. On rappelle que l'énergie potentielle de gravitation d'un solide est $U = Mgh$ où M est la masse du solide et h l'altitude du centre de gravité. Montrez que ce potentiel U s'écrit en fonction de la position $q \in SO(3)$ et du vecteur $u_0 = \overrightarrow{OG}$.

$$U(q) = Mg \langle q(u_0), e_z \rangle = \lambda \langle qe_z, e_z \rangle.$$

Recherche d'intégrales premières

Remarque : dans ces deux paragraphes on utilise le théorème de Noether pour les sous-variétés. Pour ceux qui ne sentent pas à l'aise avec ce point de vue, nous verrons dans le paragraphe suivant comment retrouver les intégrales premières en utilisant une paramétrisation de $SO(3)$.

1. On considérera à partir maintenant le lagrangien $L = T - U$ qui est celui associé au mouvement de la toupie dans un champ de pesanteur

dirigée dans la direction du champ de vecteur vertical v . Montrer alors que le système possède une invariance par rotation, donnée par l'action - à gauche - d'un sous-groupe S^1 sur $SO(3)$. Montrez à l'aide du théorème de Noether que la fonction suivante est une intégrale première

$$\mu_1(q, u) = \langle e_z, q\mathcal{I}u \rangle$$

2. Dans le cas où la toupie est homogène et possède un axe de symétrie passant par le point d'ancrage (toupie de Lagrange). Montrer que T est invariant sous l'action à droite d'un sous-groupe S^1 de $SO(3)$. En déduire l'allure de la matrice d'inertie

$$\mathcal{I} = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix}$$

Montrez à l'aide du théorème de Noether que la fonction suivante est une intégrale première

$$\mu_2(q, u) = \langle e_z, \mathcal{I}u \rangle,$$

Coordonnées

Pour mieux comprendre ce système, nous aimerions en connaître la version hamiltonienne. Comme il nous manque la version intrinsèque de la mécanique hamiltonienne, nous allons utiliser des coordonnées, appelés *angles d'Euler*. Nous avons fixé les coordonnées usuelles de $\mathbb{R}^3$ associées à sa base canonique (e_x, e_y, e_z) et nous considérons les rotations

$$\begin{aligned} R_\phi^1 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) \\ 0 & -\sin(\phi) & \cos(\phi) \end{pmatrix} \\ R_\theta^3 &= \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned}$$

Nous considérons alors l'application

$$\Psi : \begin{cases} \mathbb{R}^3 & \rightarrow SO(3) \\ (\theta, \phi, \psi) & \mapsto R_\psi^3 \circ R_\phi^1 \circ R_\theta^3. \end{cases}$$

Nous admettons que Ψ est un difféomorphisme d'un ouvert de $\mathbb{R}^3$ dans un ouvert dense de $SO(3)$.

1. Montrez maintenant que

$$\begin{aligned} T_{\theta, \phi, \psi} \Psi(\dot{\theta}, \dot{\phi}, \dot{\psi}) &= (\Psi(\theta, \phi, \psi), \dot{\theta}e_z + \dot{\phi}R_{-\theta}^3(e_x) + \dot{\psi}R_{-\theta}^3R_{-\phi}^1(e_z)) \\ &= (\Psi(\theta, \phi, \psi), \dot{\phi} \cos \theta + \dot{\psi} \sin \phi \sin \theta, \dot{\phi} \sin \theta - \dot{\psi} \sin \phi \cos \theta, \dot{\theta} + \dot{\psi} \cos \phi). \end{aligned}$$

2. Montrez alors que le Lagrangien de notre système s'écrit

$$\begin{aligned} L &= a\dot{\phi}^2 + b\dot{\theta}^2 + \dot{\psi}^2(a\sin^2\phi + b\cos^2\phi) \\ &+ 2b\dot{\psi}\dot{\theta}\cos\phi - \lambda\cos\phi. \end{aligned}$$

Montrez ensuite, soit en utilisant les résultats du paragraphe précédent, soit par l'application directe du théorème de Noether, que

$$\begin{aligned} \mu_1 &= b(\dot{\theta} + \dot{\psi}\cos\phi) \\ \mu_2 &= b\dot{\theta}\cos\phi + \dot{\psi}(a\sin^2\phi + b\cos^2\phi). \end{aligned}$$

Transformée de Legendre

1. on s'intéresse à la transformation de Legendre

$$\mathcal{L} : (\dot{\theta}, \dot{\phi}, \dot{\psi}) \mapsto (p_\theta = \frac{\partial L}{\partial \dot{\theta}}, p_\phi = \frac{\partial L}{\partial \dot{\phi}}, p_\psi = \frac{\partial L}{\partial \dot{\psi}}).$$

Montrez alors que

$$\begin{cases} p_\theta &= 2b\dot{\theta} + 2b\dot{\psi}\cos\phi \\ p_\phi &= 2a\dot{\phi} \\ p_\psi &= 2\dot{\psi}(a\sin^2\phi + b\cos^2\phi) + 2b\dot{\theta}\cos\phi \end{cases}$$

2. Montrez que l'inverse de la transformation de Legendre s'écrit

$$\begin{cases} \dot{\theta} &= p_\theta \frac{1}{2b} (1 + \frac{b\cos^2\phi}{a\sin^2\phi}) - p_\psi \frac{\cos\phi}{2a\sin^2\phi} \\ \dot{\phi} &= \frac{1}{2a} p_\phi \\ \dot{\psi} &= \frac{1}{2a\sin^2\phi} (p_\psi - p_\theta \cos\phi) \end{cases}$$

Hamiltonien

1. A l'aide de la transformation de Legendre montrez que p_ψ et p_θ sont des intégrales premières de l'équation de Hamilton. Vérifiez (ou admettez) que le Hamiltonien s'écrit

$$H = \frac{1}{2a}p_\phi^2 - \frac{\cos\phi}{2a\sin^2\phi}p_\psi p_\theta + \frac{1}{4a\sin^2\phi}p_\psi^2 + \frac{1}{4b}(1 + \frac{b\cos^2\phi}{a\sin^2\phi})p_\theta^2 + \lambda\cos\phi.$$

2. En utilisant les crochets de Poisson, montrez à nouveau que p_ψ et p_θ sont des intégrales premières. Le système est-il totalement intégrable?
3. Montrez enfin que l'on peut exprimer p_ϕ en fonctions des constantes H , p_ψ , p_θ et de ϕ puis que l'on peut résoudre explicitement les équations de Hamilton.

Troisième partie

Introduction à la théorie ergodique

Chapitre 15

Mesures Invariantes

Si en général T est une application d'un espace mesuré (M, μ) vers un espace N muni d'une tribu, telle que l'image inverse d'un mesurable est mesurable, on définit la *mesure image* de μ par T , notée $T_*\mu$, par

$$T_*\mu(U) = \mu(f^{-1}(U)).$$

Si T envoie l'espace M dans lui-même, on dira que T *préserve la mesure* μ , ou que la mesure μ est *invariante* par T , si $T_*\mu = \mu$.

Nous nous intéresserons principalement dans ce texte aux flots de champs de vecteurs.

L'étude des transformations sous cet angle, celui de la théorie de la mesure, s'appelle *théorie ergodique*. Elle examine les transformations sous l'angle le plus grossier, c'est-à-dire en oubliant la plupart des structures que peut porter l'espace M . Pourtant cette "simple" étude est d'une grande richesse.

Nous utiliserons librement le théorème de représentation de Riesz que nous pouvons concevoir comme une définition de la notion de mesure.

Théorème 15.0.0.2 *Soit K un ensemble compact. Soit $C^0(K)$ l'ensemble des fonctions continues sur K . Soit m une forme linéaire sur $C^0(K)$, telle que*

$$f \geq 0 \implies m(f) \geq 0.$$

Alors, il existe une et une seule mesure borelienne finie μ sur K telle que

$$m(f) = \int f d\mu.$$

Une mesure est donc caractérisée par sa valeur contre les fonctions continues.

Exemples et exercice

1. Soit m un point fixe d'un champ de vecteur, montrez alors que la masse de Dirac δ_m porté par m est invariante par le flot.

2. Soit $\gamma : [0, l] \rightarrow U$ l'orbite périodique d'un champ de vecteur. Montrez que la mesure μ_γ définie par

$$\int f d\mu_\gamma = \int_0^l f(\gamma_t) dt,$$

est invariante.

3. Soit X un champ de vecteur sur $\mathbb{R}^n$. Montrez que X préserve la mesure de Lebesgue si et seulement si

$$\sum_i \frac{\partial X}{\partial x_i} = 0;$$

Indication. Utilisez la formule de changement de variable

$$\int_A |J(\phi)| f \circ \phi d\lambda = \int_{\phi(A)} f,$$

et la formule, pour une famille de matrices A dépendant de t

$$\frac{d}{dt} \log \det(A) = \text{trace} \left(\frac{dA}{dt} A^{-1} \right).$$

4. Montrez que le flot d'un gradient hamiltonien préserve la mesure de Lebesgue.
5. *Décalage de Bernouilli.* Voici un exemple qui provient de la théorie de l'information. On considère un ensemble fini F , par exemple $F = \{0, 1\}$. On s'intéresse à l'ensemble $F^{\mathbb{Z}}$ considéré comme l'ensemble des applications de $\mathbb{Z}$ dans F , où comme l'ensemble des tirages bi-infinis dans l'ensemble F . Soit $\mathcal{A} = \{A_{-n}, \dots, A_n\}$ une famille finie de sous-ensembles de F . Le cylindre $C(\mathcal{A})$ est défini par

$$C(\mathcal{A}) = \{f / f(i) \in A_i\}.$$

On munit $F^{\mathbb{Z}}$ de la tribu engendrée par les cylindres. On fixe une mesure μ_0 de probabilité sur F et on considère la mesure produit μ sur $F^{\mathbb{Z}}$. On rappelle que cette mesure est caractérisée par la propriété suivante.

$$\mu(C(\mathcal{A})) = \prod_i \mu_0(A_i).$$

Montrez alors que le *décalage de Bernouilli - shift* en anglais - donné par

$$\sigma(f)(x) = f(x+1),$$

préserve ces mesures produits.

6. On considère le cercle

$$S^1 = \{z \in \mathbb{C} / |z| = 1\}.$$

On considère l'application $\psi : t \mapsto \exp(2i\pi t)$ qui paramètre $S^1 \setminus \{1\}$ par $]0, 2\pi[$. On note R_θ la rotation $u \mapsto e^{i\theta}u$. On dit qu'elle est d'*angle irrationnel* si $\frac{\theta}{2\pi} \notin \mathbb{Q}$. Montrez que $\lambda_c = \psi_*\lambda$ est la seule mesure invariante par toutes les rotations. Montrez qu'une rotation d'angle irrationnel possède une seule mesure invariante.

15.1 Existence de mesures invariantes

L'existence de mesures invariantes est garantie par le théorème de Kakutani-Markov.

Théorème 15.1.0.3 *Soit K un ensemble compact muni de la tribu borelienne. Toute transformation continue de K préserve une mesure finie invariante. De même tout groupe à un paramètre de transformations continues de K préserve une mesure finie.*

Dans le cas des champs de vecteurs, on obtient

Proposition 15.1.0.4 *Soit K un ensemble compact invariant par le flot φ_t d'un champ de vecteur défini sur un ouvert U , alors il existe une mesure finie invariante par φ_t et dont le support est inclus dans K , c'est-à-dire telle que $\mu(U \setminus K) = 0$.*

Indication de la démonstration. Nous allons donner simplement une esquisse de la démonstration. Celle-ci, quoique simple, utilise des arguments d'analyse fonctionnelle, en particulier la compacité de la boule unité du dual d'un Banach pour la topologie faible. Nous allons devoir admettre ces résultats.

Rappelons quelques notations.

Soit E un espace de Banach. Soit E^* son dual. On le munit de la norme

$$\|\alpha\| = \sup\{\alpha(u)/|u| \leq 1\}.$$

On dira que la suite $\{\alpha_n\}_{n \in \mathbb{N}}$ converge $*$ -faiblement vers α et on notera

$$\alpha_n \rightharpoonup \alpha,$$

si, pour tout u de E ,

$$\alpha_n(u) \rightarrow \alpha(u).$$

Par définition, pour tout u , l'application de E^* dans $\mathbb{R}$, définie par

$$\alpha \mapsto \alpha(u),$$

est continue pour la topologie $*$ -faible.

Nous admettrons le résultat suivant.

Théorème 15.1.0.5 *La boule unité du dual d'un espace de Banach est compacte pour la topologie $*$ -faible.*

Revenons aux mesures. Si K est un compact, nous avons vu les mesures de probabilité par le théorème de représentation de Riesz comme des formes linéaires sur l'espace des fonctions continues $C^0(K)$. Ce dernier espace est un espace de Banach quand on le munit de la norme

$$\|f\| = \sup_{x \in K} |f(x)|.$$

On remarque alors que les mesures sont des formes continues, puisque

$$\|\mu\| = \sup_{f/\forall x \in K, |f(x)| \leq 1} \left\{ \left| \int f d\mu \right| \right\} \leq 1.$$

De plus, l'ensemble $\mathcal{M}$ de mesures de probabilités est un fermé pour la topologie *-faible. On peut en effet écrire, d'après le théorème de représentation de Riesz.

$$\mathcal{M} = \{\alpha \in E^* / \alpha(1) = 1\} \bigcap_{f \geq 0} \{\alpha \in E^* / \alpha(f) \geq 0\},$$

et dans le terme de droite tous les ensembles sont des fermés pour la topologie *-faible, puisqu'image réciproque de fermé par les applications continues :

$$\alpha \mapsto \alpha(f).$$

Ainsi donc, $\mathcal{M}$ est un fermé pour la topologie faible, qui d'ailleurs dans ce cas particulier s'appelle *topologie vague*. En utilisant le théorème que nous venons de citer nous obtenons que $\mathcal{M}$ est compact pour la topologie vague.

Soit T une transformation continue de K . T_* est alors une transformation continue de $\mathcal{M}$ muni de la topologie vague. Soit μ une mesure quelconque et posons

$$\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} T_*^i \mu.$$

La famille μ_n est une famille de mesure et nous pouvons donc extraire une sous-suite qui converge pour la topologie vague

$$\mu_{n_p} \xrightarrow{p \rightarrow \infty} \mu_\infty.$$

Nous allons montrer que μ_∞ est invariante. Remarquons que

$$\|T_* \mu_n - \mu_n\| = \frac{1}{n} \|\mu - T_* \mu\| \leq \frac{2}{n}.$$

Ainsi

$$T_* \mu_n - \mu_n \rightarrow 0.$$

Par continuité de T_* , nous en déduisons

$$T_* \mu_\infty = \mu_\infty. \diamond$$

15.2 Théorème de récurrence de Poincaré

L'existence d'une mesure invariante à des conséquences sur la dynamique du système. L'un des premières est celle dite de récurrence de Poincaré. Introduisons une définition. Le *support* d'une mesure borelienne μ définie sur un espace topologique M est l'intersection du complémentaire des ouverts de mesure nulle

$$Supp(\mu) = \bigcap_{U \text{ ouvert } / \mu(U)=0} (M \setminus U).$$

Autrement dit $Supp(\mu)$ est l'ensemble fermé caractérisé par le fait que pour tout ouvert U

$$\mu(U) = 0 \iff U \cap Supp(\mu) = \emptyset.$$

Théorème 15.2.0.6 *Soit φ_t le flot d'un champ de vecteur préservant une mesure finie μ . Alors pour tout point x du support de μ , pour tout réel T positif, pour tout ε , il existe un point y du support de μ un entier n positif, tels que*

$$\|y - x\| \leq \varepsilon, \quad \|\phi_{nT}(y) - x\| \leq \varepsilon.$$

Autrement dit tout point du support de la mesure, possède un point arbitrairement proche de lui qui va revenir arbitrairement proche de lui au bout d'un temps arbitrairement long.

DÉMONSTRATION : . Soit x un point de $Supp(\mu)$ et $U = Supp(\mu) \cap B(x, \varepsilon)$. Nous voulons montrer que pour tout T , il existe $n \geq 0$ tel que $U \cap \varphi_{nT}(U) \neq \emptyset$.

Considérons donc $U_n = \phi_{2nT}(U)$. Comme $\mu(U) > 0$ et que ϕ_t préserve μ , nous avons

$$\sum_{i=1}^{i=n} \mu(U_n) = n\mu(U) \xrightarrow{n \rightarrow \infty} \infty.$$

Comme μ est finie, les U_n ne peuvent être tous disjoints. Nous en déduisons qu'il existe $i > j$ tel que $U_i \cap U_j \neq \emptyset$. Ainsi $\phi_{2iT}(U) \cap \phi_{2jT}(U) \neq \emptyset$ et donc

$$U \cap \phi_{2(i-j)T}(U) \neq \emptyset. \diamond$$

Paradoxe de Zermelo

Attachons nous un moment à commenter un paradoxe classique et célèbre qui met en contradiction le théorème de récurrence et le principe de Clausius de croissance de l'entropie.

Imaginons deux chambres communicantes. Dans leur état initial, l'une des chambres est remplie de gaz, l'autre est vide. On ouvre la porte, d'après le principe de Clausius, le gaz va se répartir équitablement entre les deux chambres. Selon le théorème de récurrence de Poincaré, au bout d'un certain temps aussi grand que l'on veut, on retrouvera une situation où presque tout le gaz se retrouve dans l'une des chambres.

Si l'on veut rendre le paradoxe plus spectaculaire, en imaginant cette fois-ci que l'univers est fini et que l'évolution préserve une mesure - après tout, dans une vision purement déterministe du monde, le mouvement des atomes est régi par un hamiltonien - vous allez vous retrouver au bout d'un temps arbitraire par exemple plus grand que la durée de vie du système solaire à relire à nouveau ces lignes, après avoir mangé à peu près le même petit déjeuner.

La manière habituelle de résoudre ce paradoxe est d'évaluer le temps de récurrence, c'est-à-dire le temps de retour à un état presque initial. On s'aperçoit alors que s'il y a très peu de molécules de gaz - par exemple deux - la récurrence de Poincaré se fait assez vite, ce qui paraît bien plausible. Par contre, dans le

cas d'un grand nombre de molécules - c'est-à-dire le cas normal - les calculs font apparaître un temps colossal bien plus grand que la durée de vie du système solaire. La discussion s'arrête en général là.

De manière un peu plus satisfaisante, il faut se rendre compte que le principe de Clausius est un principe statistique. On ne cherche pas à comprendre l'évolution d'une particule, mais l'évolution d'une statistique, c'est-à-dire d'une mesure. Nous ne connaissons pas l'état initial, seulement la probabilité d'un état initial. C'est le principe de base de la physique statistique, renforcé bien sûr par le principe d'incertitude d'Heisenberg. Or l'évolution d'une probabilité n'a pas de raison d'être récurrente, au contraire notre connaissance du système peut devenir de plus en plus diffuse. C'est ce qui se passe pour les transformations chaotiques, où cette probabilité peut tendre à être invariante. Ce que nous pouvons effectivement prévoir - et c'est la seule chose raisonnable à faire - c'est que le gaz va être en moyenne répartie équitablement entre les deux chambres. Tout autre prévision étant déraisonnable.

Exercices

1. Soit x_0 un point fixe du flot ϕ^t . On suppose que le bassin d'attraction de x_0

$$B = \{y, \lim_{t \rightarrow \infty} \phi_t(y) = x_0\},$$

est ouvert. On se donne une mesure μ invariante dont le support est inclus dans B . Montrez que cette mesure est la masse de Dirac en x_0 .

2. Montrez que si K est un compact de $\mathbb{R}^2$ invariant par le flot d'un champ de vecteur, alors K contient un point fixe ou une orbite périodique. *Indication.* Utilisez le théorème de Kakutani-Markov, le théorème de récurrence de Poincaré et le théorème de Poincaré-Bendixson.
3. *Approximations simultanées de réels.* On considère le tore

$$\mathbb{T}^n = \underbrace{S^1 \times \dots \times S^1}_n \subset \mathbb{C}^n.$$

- (a) Montrez que $\mathbb{T}^n$ est une sous-variété de $\mathbb{C}^n$.
- (b) On considère les n -champs de vecteurs, définis en complexe par

$$\partial_{\theta_k}(z_1, \dots, z_n) = (z_1, \dots, iz_k, \dots, z_n).$$

Soit φ_t^k le flot de ∂_{θ_k} . Montrez que

$$\varphi_t^k(z_1, \dots, z_n) = (z_1, \dots, e^{it}z_k, \dots, z_n).$$

En déduire que $\mathbb{T}^n$ est invariant par ces champs de vecteurs.

- (c) Pour tout x de $\mathbb{T}^n$, on considère ψ_x

$$\begin{cases} [0, 2\pi]^n & \rightarrow \mathbb{T}^n \\ (t_1, \dots, t_n) & \mapsto \varphi_{t_1}^1 \circ \dots \circ \varphi_{t_n}^n(x). \end{cases}$$

Montrez que ψ_x est un difféomorphisme sur un ouvert dense de $\mathbb{T}^n$, puis que la mesure $\mu_x = (\psi_x)_*\lambda$ ne dépend pas de x , où λ est la mesure de Lebesgue, et enfin que cette mesure $\mu = \mu_x$ est invariante par les différents flots.

- (d) Soit $(\lambda_1, \dots, \lambda_n)$ n nombres réels. Montrez à l'aide du théorème de récurrence de Poincaré que

$$\forall \varepsilon > 0, \forall k_0, \exists k \geq k_0, \exists p_i \in \mathbb{Z} / \sup_i \left\| \lambda_i - \frac{p_i}{k} \right\| \leq \frac{\varepsilon}{k}.$$

Indication. Utilisez le flot

$$\phi_t(x) = \psi_x(2\pi t\lambda_1, \dots, 2\pi t\lambda_n).$$

Remarque. Le cas $n = 1$ est bien connu et se montre directement à l'aide des fractions continues par exemple, par contre pour n plus grand il est plus difficile de construire une autre preuve que celle mentionnée dans cet exercice.

15.3 Théorèmes ergodiques

Nous dirons que la mesure μ est *ergodique* pour la transformation T - ou que la transformation T est ergodique - si les seules fonctions invariantes sont constantes μ presque partout.

De manière équivalente, une transformation est ergodique si les seuls ensembles invariants sont de mesure nulle ou pleine - c'est-à-dire que leur complémentaire est de mesure nulle.

EXEMPLE :

1. Montrez qu'une mesure invariante μ n'est pas ergodique, si et seulement si on peut écrire $\mu = a\mu_1 + b\mu_2$, où b et a sont des réels strictement positifs, μ_i des mesures invariantes non proportionnelles.
2. *Rotations.* On considère S^1 muni de sa mesure invariante par rotation. Montrez que les rotations d'angle irrationnel sont ergodiques. Qu'en est-il des rotations d'angle rationnel ?
3. *Décalage de Bernouilli.* Nous verrons par la suite que le décalage de Bernouilli est ergodique pour les mesures produits.

15.3.1 Théorèmes ergodiques

Le théorème ergodique de Birkhoff est d'une grande importance. Il s'énonce "physiquement" de la manière suivante : *les moyennes temporelles et spatiales sont égales pour une transformation ergodique.* Nous l'énoncerons dans le cas d'une transformation unique, le cas des groupes à un paramètre s'énonçant de manière identique.

Théorème 15.3.1.1 *Soit T une transformation de X préservant une mesure finie μ . Soit φ une fonction dans $L^1(\mu)$. Il existe alors une fonction mesurable $\bar{\varphi}$ invariante telle que pour μ -presque tout x ,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left(\sum_{i=1}^{i=n} \varphi(T^i(x)) \right) = \lim_{n \rightarrow \infty} \frac{1}{n} \left(\sum_{i=1}^{i=n} \varphi(T^{-i}(x)) \right) = \bar{\varphi}(x).$$

De plus,

$$\int \varphi d\mu = \int \bar{\varphi} d\mu.$$

Remarques

1. Le cas souvent utilisé est celui des transformations ergodiques pour une mesure de probabilité; dans ce cas $\bar{\varphi}$ est une constante qui vaut

$$\bar{\varphi} = \int \varphi d\mu.$$

2. Dans le cas du décalage de Bernouilli, le théorème ergodique de Birkhoff est la justification mathématique de la loi “expérimentale” des grands nombres.

Nous ne démontrerons pas ce théorème dans le cas général. Nous démontrerons le cas particulier où φ est une fonction dans L^2 et nous montrerons ce que l’on appelle le théorème ergodique de Von Neuman.

Théorème 15.3.1.2 *Soit T une transformation préservant une mesure μ . Soit φ un élément de $L^2(\mu)$. Alors il existe une fonction $\bar{\varphi}$, T -invariante telle que*

$$\lim_{n \rightarrow \pm\infty} \left\| \frac{1}{n} \left(\sum_{i=0}^{i=n} \varphi \circ T^i \right) - \bar{\varphi} \right\|_{L^2} = 0.$$

Commençons par remarquer qu’une transformation T préservant une mesure μ définit une transformation unitaire de $L^2(\mu)$ donnée par $T^* : \varphi \mapsto \varphi \circ T$. Le théorème suit alors du résultat suivant.

Théorème 15.3.1.3 *Soit U un opérateur unitaire d’un espace de Hilbert H . Soit V le sous-espace fermé constitué des vecteurs invariants et π la projection sur V . Alors, pour tout u de H*

$$\lim_{n \rightarrow \infty} \left\| \frac{1}{n} \left(\sum_{i=1}^{i=n} U^i(u) \right) - \pi(u) \right\| = 0.$$

DÉMONSTRATION : En posant $v = u - \pi(u)$, on se ramène au cas où $\pi(v) = 0$ - c’est-à-dire $v \in V^\perp$ - et nous devons donc montrer

$$\left\| \frac{1}{n} \left(\sum_{i=0}^{i=n} U^i(v) \right) \right\| \rightarrow 0.$$

Or $V = \text{Ker}(1 - U)$ et donc $V^\perp = \text{Im}(1 - U^*)$, où U^* est l'adjoint de U . Ainsi $v = (1 - U^*)(w)$. Nous avons alors

$$\left\| \frac{1}{n} \left(\sum_{i=1}^{i=n} U^n(v) \right) \right\| = \left\| \frac{1}{n} \left(\sum_{i=1}^{i=n} U^n(w) - \sum_{i=1}^{i=n} U^n U^*(w) \right) \right\|$$

Comme U est unitaire on a $UU^* = 1$. Ainsi

$$\left\| \frac{1}{n} \left(\sum_{i=1}^{i=n} U^n(v) \right) \right\| = \left\| \frac{1}{n} (w - U^n(w)) \right\| \leq \frac{2\|w\|}{n} \rightarrow 0. \diamond$$

Un critère d'ergodicité

Le lemme suivant qui est une sorte de réciproque du théorème d'ergodicité nous donne un moyen de démontrer qu'une transformation est ergodique.

Lemme 15.3.1.4 *Soit T une transformation continue d'un espace compact X préservant une mesure de probabilité μ . Supposons que pour toute fonction continue φ et pour μ -presque tout x , on a*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left(\sum_{i=1}^{i=n} \varphi(T^n(x)) \right) = \int \varphi d\mu.$$

Alors T est ergodique

Démonstration du lemme. Supposons que la mesure est non ergodique. On sait alors que l'on peut écrire $\mu = t\mu_1 + (1 - t)\mu_2$ où $t \in]0, 1[$, et où les mesures μ , μ_1 et μ_2 sont des mesures de probabilités distinctes. Remarquons que

$$\mu(F) = 1 \implies \mu_1(F) = \mu_2(F) = 1.$$

Par le théorème de représentation de Riesz, il existe une fonction φ telle que $\int \varphi d\mu_1 \neq \int \varphi d\mu_2$. Posons

$$\Sigma_n = \frac{1}{|n|} \sum_{i=1}^{i=n} \varphi \circ T^n.$$

Soit W l'ensemble des points x tels que la suite $\{\Sigma_n(x)\}$ converge quand n tend vers $+\infty$ et vers $-\infty$ vers une limite commune que nous noterons $\bar{\varphi}$. Nous savons alors que W est de mesure pleine pour toute mesure invariante par le théorème de Birkhoff et que

$$\begin{aligned} \int \bar{\varphi} d\mu_1 &= \int \varphi d\mu_1 \\ \int \bar{\varphi} d\mu_2 &= \int \varphi d\mu_2. \end{aligned}$$

Ainsi $\int \bar{\varphi} d\mu_1 \neq \int \bar{\varphi} d\mu_2$. Nous en déduisons que pour tout ensemble E de μ mesure pleine il existe $x \in E \cap W$, $y \in E \cap W$, tel que $\bar{\varphi}(x) \neq \bar{\varphi}(y)$. C'est ce que nous voulions démontrer. $\square$

Remarque. Nous aurions pu adapter la preuve de ce lemme en de façon à n'utiliser que n'utilisant que le théorème ergodique de Von Neuman, ainsi que la propriété suivante de la convergence dans L^2 : si la suite $\{\varphi\}_n$ converge vers φ dans L^2 , alors une sous-suite converge simplement sur un ensemble de mesure pleine vers φ .

15.4 Ergodicité du décalage de Bernouilli

Nous allons démontrer l'ergodicité du décalage de Bernouilli. Plusieurs raisons justifient ce choix. Tout d'abord nous avons vu plus haut que le le théorème de Birkhoff justifie dans ce cas la loi des grands nombres. Ensuite, la démonstration de l'ergodicité se généralise aux cas des transformations dites *hyperboliques* et nous essaierons de justifier cette assertion. Enfin, le décalage joue un rôle fondamental dans les systèmes dynamiques, c'est en quelque sorte l'archétype d'une tranformation quelconque, comme nous l'évoquerons au paragraphe suivant consacré au codage et à l'entropie.

Description du décalage

Nous allons décrire les propriétés du décalage - noté σ - qui nous seront utiles par la suite.

Identifions F à l'ensemble $\{0, 1, \dots, M\}$. Commençons par munir $X = F^{\mathbb{Z}}$ de la distance d définie par

$$d(f, g) = \sum_{i=-\infty}^{i=\infty} \frac{1}{2^{-|i|}} |f(i) - g(i)|.$$

Pour cette distance X est un espace compact.

Notons ensuite pour tout f de X

$$\mathcal{F}_f^+ = \{g / \forall m \geq 1, f(n) = g(n)\}$$

et

$$\mathcal{F}_f^- = \{g / \forall m \leq 0, f(n) = g(n)\}$$

Remarquons rapidement que ces ensembles forment une partition de X

$$\begin{aligned} g \in \mathcal{F}_g^\pm &\iff f \in \mathcal{F}_f^\pm \\ \mathcal{F}_f^\pm \cap \mathcal{F}_g^\pm &\neq \emptyset \iff \mathcal{F}_f^\pm = \mathcal{F}_g^\pm \\ X &= \bigcup_f \mathcal{F}_f^+ = \bigcup_f \mathcal{F}_f^-. \end{aligned}$$

Nous appellerons les ensembles $\mathcal{F}_f^+$ *feuilles stables* et les ensembles $\mathcal{F}_f^-$ *feuilles instables* et il est utile de penser de manière abusive à ces objets comme définissant deux familles de feuilletages de X .

Les propriétés suivantes des feuilles donnent au décalage de Bernouilli la structure d'un *système dynamique hyperbolique*. Ce sont les seules que nous utiliserons.

On montre facilement que

$$\begin{aligned} f \in \mathcal{F}_g^- &\implies \lim_{n \rightarrow -\infty} d(\sigma^n(f), \sigma^n(g)) = 0 \\ f \in \mathcal{F}_g^+ &\implies \lim_{n \rightarrow +\infty} d(\sigma^n(f), \sigma^n(g)) = 0. \end{aligned}$$

Bien sûr ces deux propriétés font penser aux variétés stables et instables au voisinage des points fixes hyperboliques.

Enfin, nous avons une propriété - que nous appellerons *absolue continuité des feuilletages* - reliant les feuilles et la mesure :

Tout ensemble E de mesure pleine possède un sous-ensemble G de mesure pleine tel que pour tout f de G il existe un sous-ensemble G_f de E de mesure pleine tel que

$$g \in G_f \implies g \in E \cap \mathcal{F}_f^+ \cap \mathcal{F}_g^- \neq \emptyset.$$

A l'aide de cette propriété dont nous repoussons la démonstration, montrons l'ergodicité du décalage.

Soit φ une fonction continue de X dans $\mathbb{R}$. D'après le théorème de Birkhoff, il existe un ensemble E de mesure pleine tel que pour tout x de E

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \left(\sum_{i=1}^{i=n} \varphi \circ \sigma^n(x) \right) = \lim_{n \rightarrow +\infty} \frac{1}{n} \left(\sum_{i=1}^{i=n} \varphi \circ \sigma^{-n}(x) \right) = \bar{\varphi}(x).$$

Remarquons maintenant que φ est absolument continue et bornée par une constante K . Ainsi pour tout ε strictement positif, il existe α tel que

$$d(f, g) \leq \alpha \implies |\varphi(f) - \varphi(g)| \leq \varepsilon.$$

Ceci nous permet de montrer que si $f \in \mathcal{F}_g^+$ et que si f et g appartiennent toutes deux à E alors $\bar{\varphi}(f) = \bar{\varphi}(g)$. En effet, on sait que pour $n \geq n_0$, $d(\sigma^n(f), \sigma^n(g)) \leq \alpha$ et donc

$$\begin{aligned} \left| \frac{1}{n} \left(\sum_{i=1}^{i=n} \varphi \sigma^n(f) \right) - \frac{1}{n} \left(\sum_{i=1}^{i=n} \varphi \sigma^n(g) \right) \right| &\leq \frac{1}{n} \sum_{i=n}^{i=n} |\varphi \sigma^n(f) - \varphi \sigma^n(g)| \\ &\leq \frac{1}{n} (2Kn_0 + \alpha(n - n_0)). \end{aligned}$$

A la limite, on obtient bien $\bar{\varphi}(f) = \bar{\varphi}(g)$. Le même raisonnement montre que si $f \in \mathcal{F}_g^-$ alors $\bar{\varphi}(f) = \bar{\varphi}(g)$.

Nous allons conclure maintenant. Soit G l'ensemble de mesure pleine obtenu par l'absolue continuité des feuilletages. Nous en déduisons que pour tout f

élément de G et tout g élément de G_f , nous avons $\bar{\varphi}(f) = \bar{\varphi}(g)$. En effet, il existe h appartenant à E , tel que $h \in \mathcal{F}_f^+ \cap \mathcal{F}^-g$. Ainsi, $\bar{\varphi}(f) = \bar{\varphi}(h) = \bar{\varphi}(g)$ et donc $\bar{\varphi}$ est constante sur l'ensemble de mesure pleine G_f .

La fonction $\bar{\varphi}$ est donc constante presque partout. Elle vaut donc $\int \varphi d\mu$.

Il ne nous reste plus qu'à démontrer l'absolue continuité des feuilletages.

Pour cela posons $\mathbb{Z}^- = \{z \in \mathbb{Z} / z < 0\}$. Nous avons alors $\mathbb{Z} = \mathbb{Z}^- \sqcup \mathbb{N}$ et donc

$$F^{\mathbb{Z}} = F^{\mathbb{Z}^-} \times F^{\mathbb{N}}$$

où F^A désigne l'ensemble des applications de A dans F . Notons π_- (resp. π_+) la projection de $F^{\mathbb{Z}}$ sur $F^{\mathbb{Z}^-}$ (resp. $F^{\mathbb{N}}$) qui se déduisent de cette identification. Ces projections sont en fait des restrictions. Nous poserons $f^\pm = \pi_\pm(f)$.

Par construction on a $\pi_\pm^{-1}\pi_\pm(f) = \mathcal{F}_f^\pm$.

Les ensembles $F^{\mathbb{Z}^-}$ et $F^{\mathbb{N}}$ sont chacun munis d'une mesure produit notée μ_+ et μ_- et nous avons alors $\mu = \mu_+ \otimes \mu_-$ et $\mu_\pm = p_\pm^*(\mu)$.

Soit E un ensemble de mesure pleine, le théorème de Fubini nous dit alors que pour μ -presque tout f , le sous-ensemble de $F^{\mathbb{N}}$ défini par

$$G_f^+ = \{h / (f^-, h) \in E\},$$

est de μ_+ -mesure pleine. Ainsi,

$$G_f = E \cap \pi_+^{-1}(G_f^+),$$

est de mesure pleine. Par construction enfin, si g est un élément de G_f , g_+ appartient à G_f^+ et donc (f^-, g^+) appartient à E . Pour conclure, il suffit de remarquer que (f^-, g^+) appartient à $\mathcal{F}_f^- \cup \mathcal{F}_g^+$. $\square$

REMARQUES :

1. On pourrait adapter la preuve de façon à n'utiliser que le théorème ergodique de Von Neumann, et la propriété suivante de la convergence dans L^2 : si la suite $\{\varphi\}_n$ converge vers φ dans L^2 , alors une sous-suite converge simplement sur un ensemble de mesure pleine vers φ .

15.4.1 Du décalage aux systèmes hyperboliques

Dans la preuve de l'ergodicité du décalage nous avons isolé certaines propriétés que d'autres systèmes pourraient avoir.

Nous allons donner une définition d'un système hyperbolique en faisant ressortir ces propriétés. La définition donnée ici n'est pas tout à fait compatible avec les définitions de la littérature. Nous appellerons cette version simplifiée $*$ -hyperbolicité.

Commençons par quelques définitions préliminaires.

Produit local

Soit X est un espace topologique muni d'une distance d . Si $\mathcal{F}$

$$X = \sqcup_l \mathcal{F}_l,$$

est une partition de X nous appellerons *feuille passant par x* et nous noterons $\mathcal{F}_x$ l'ensemble de la partition à laquelle appartient x .

Nous dirons que deux partitions

$$X = \sqcup_l \mathcal{F}_l^+, X = \sqcup_l \mathcal{F}_l^-,$$

définissent une *structure de produit local*, si tout x_0 de X possède un voisinage ouvert U , tel que pour tout y de U , l'intersection de $\mathcal{F}_y^+ \cap U$ (resp. $\mathcal{F}_y^- \cap U$) avec $\mathcal{F}_{x_0}^- \cap U$ (resp. $\mathcal{F}_{x_0}^+ \cap U$) est réduite à un seul élément noté $\pi^-(y)$ (resp. $\pi^+(y)$) et si de plus l'application ainsi définie

$$(\pi^+, \pi^-) : U \rightarrow (\mathcal{F}_{x_0}^+ \cap U) \times (\mathcal{F}_{x_0}^- \cap U)$$

est un homéomorphisme

Homéomorphisme *-hyperbolique

Nous dirons alors qu'un homéomorphisme T d'un espace topologique X muni d'une métrique d est **-hyperbolique* si X possède deux partitions $(\mathcal{F}^+, \mathcal{F}^-)$ dites *feuilletage stables* et *feuilletages instables* définissant une structure de produit local et telles que

$$\begin{aligned} \forall y \in \mathcal{F}_x^+, \quad d(T^n(x), T^n(y)) &\xrightarrow{n \rightarrow +\infty} 0 \\ \forall y \in \mathcal{F}_x^-, \quad d(T^n(x), T^n(y)) &\xrightarrow{n \rightarrow -\infty} 0 \end{aligned}$$

EXEMPLES

1. Le cas du décalage est notre premier exemple. Dans cet exemple, la structure de produit local est en fait globale, puisque l'ouvert U est l'espace X tout entier.
2. Un deuxième exemple est celui défini plus tôt dans le cours et associé à la matrice

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Cette matrice à coefficients entiers agit par homéomorphisme sur le tore $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. Notons p la projection de $\mathbb{R}^2$ sur $\mathbb{T}^2$. Soit Λ^+ (resp. Λ^-) le sous-espace propre associé à la valeur propre plus petite que 1, (resp. plus grande). On obtient alors deux partitions vérifiant

$$\mathcal{F}_{p(x)}^\pm = p(\Lambda^\pm + x).$$

Il n'est alors pas difficile de vérifier que ces deux partitions définissent une structure de produit local et que A est hyperbolique

3. De manière plus générale, soit T un difféomorphisme d'un ouvert (ou d'une sous-variété) U de $\mathbb{R}^n$ dans lui-même. Nous dirons qu'un ensemble E fermé et invariant par T est *hyperbolique* s'il existe des constantes C et A positives, si pour tout x de E , il existe deux sous-espaces vectoriels E_x^+ et E_x^- de $T_x U$, tels que

$$\begin{aligned} T_x U &= E_x^+ \oplus E_x^- \\ \forall u \in E_x^\pm, \forall n \geq 0 \quad , \quad \|DT^{\pm n}(u)\| &\leq \|u\| C e^{\mp A n}. \end{aligned}$$

Quand E est réduit à un point, on reconnaît la définition d'un point fixe hyperbolique. Si $E = U$, on dira que T est un *difféomorphisme d'Anosov*. On peut alors montrer que T en tant que qu'application de E dans lui-même est $*$ -hyperbolique. Pour cela on montre, par des arguments analogues à ceux du théorème de la variété stable de la première partie du cours, que les feuilletages stables et instables existent ; dans cette situation, les feuilles $\mathcal{F}_x^\pm$ sont des sous-variétés et leur espace tangent en y est justement $E_y^\pm$.

15.4.2 Feuilletages absolument continus

Maintenant, il nous reste à définir les mesures invariantes adaptées à la situation. Nous dirons tout d'abord que deux mesures μ_1 et μ_2 sont *absolument continues* l'une par rapport à l'autre si

$$\mu_1(E) = 0 \iff \mu_2(E) = 0.$$

Ensuite, plaçons nous dans la situation de deux partitions $\mathcal{F}^+$ et $\mathcal{F}^-$ de X définissant un produit local. Rappelons que pour tout x_0 de X , il existe alors un voisinage ouvert U de x_0 , telle que l'application $\pi = (\pi^+, \pi^-)$,

$$\pi = (\pi^+, \pi^-) : U \rightarrow (\mathcal{F}_{x_0}^+ \cap U) \times (\mathcal{F}_{x_0}^- \cap U)$$

est un homéomorphisme. Nous dirons alors que ces deux partitions sont *absolument continues* par rapport à la mesure μ si, en notant $\nu = \mu|_U$, la mesure $\pi_*^{-1}(\pi_*^+ \nu \otimes \pi_*^- \nu)$ est absolument continue par rapport à ν .

Dans le cas du décalage de Bernouilli, la mesure produit à cette propriété ; dans le cas du tore, la mesure de Lebesgue a également cette propriété.

Dans cette situation, on montre alors en suivant la même preuve que pour le décalage, le résultat suivant.

Si les feuilletages stables et instables d'une transformation $$ -hyperbolique T sont absolument continus par rapport à la mesure invariante μ , celle-ci est ergodique pour T .*

Chapitre 16

Codage et entropie

Le décalage vont nous permettre de décrire les systèmes dynamiques du point de vue de la théorie de la mesure, et en particulier d'associer un invariant, l'*entropie* de Kolmogorov-Sinaï, à toute mesure invariante. Cet invariant mesure le désordre du système, la vitesse à laquelle il devient purement aléatoire. Il permet enfin de décrire complètement de nombreux systèmes du point de vue de la théorie de la mesure.

L'idée fondamentale est celle du codage : décrire le systèmes, à l'aide d'une partition finie de l'espace des phases, comme une suite de symboles.

Dans ce chapitre, la plupart des théorèmes seront énoncés sans démonstration. Ce sont à chaque fois des résultats dont la démonstration est plus délicate que ce que nous avons vu jusque là.

16.1 Codage

Commençons par une remarque. Si T est une transformation d'un espace X , nous en déduisons une application $\mathcal{T}$

$$\begin{cases} X & \rightarrow X^{\mathbb{Z}} \\ x & \mapsto (\dots, T^{-n}(x), \dots, T^{-1}(x), x, T(x), \dots, T^n(x), \dots). \end{cases}$$

Par construction, nous avons alors

$$\mathcal{T} \circ T = \sigma \circ \mathcal{T},$$

où σ est le décalage de Bernouilli sur l'espace - énorme - $X^{\mathbb{Z}}$. Ainsi, tout système dynamique est un sous-système du décalage. Cette situation n'est pourtant pas très intéressante : on ne voit pas en quoi on a simplifié la situation en remplaçant X par l'ensemble $X^{\mathbb{Z}}$.

16.1.1 Codage, partition génératrice

L'idée de codage, analogue à celle-ci, est plus efficace. Considérons une partition *finie* $\xi = \{\xi_i\}_{i \in F}$ de X . Nous en déduisons une application $\hat{\xi}$ de X dans

$F^{\mathbb{Z}}$. Pour cela, notons $\dot{\xi}(x)$ l'élément de F , tel que

$$x \in \xi_{\dot{\xi}(x)}.$$

Posons, pour tout sous-ensemble A de $\mathbb{Z}$,

$$\hat{\xi}^A : \begin{cases} X & \rightarrow & F^A \\ x & \mapsto & (\dots, \dot{\xi}(T^n(x)), \dots)_{n \in A} \end{cases}$$

Nous poserons enfin $\hat{\xi} = \hat{\xi}^{\mathbb{Z}}$ (resp. $\hat{\xi}^+ = \hat{\xi}^{\mathbb{N}}$) à valeurs dans $F^{\mathbb{Z}}$ (resp. $F^{\mathbb{N}}$).

L'idée physique sous-jacente est claire : en fait l'espace des phases est partitionné de manière finie - les états de même mesure physique - le système dynamique est alors représenté par la succession des états dans lequel l'appareil de mesure trouve la particule.

L'application $\hat{\xi}$ n'est en général ni une injection, ni une surjection. Nous dirons que la partition ξ est *génératrice* si l'application $\hat{\xi}$ est une injection pour un ensemble de mesure pleine.

Exemples et exercices

1. Donnez une partition génératrice pour le décalage de Bernouilli.
2. Soit R_θ une rotation d'angle θ irrationnel, montrez que la partition $S^1 = e^{i[0, \pi]} \cup e^{i[\pi, 2\pi[}$ est génératrice.
3. Les difféomorphismes d'Anosov admettent des partitions génératrices (cf 2).

16.2 Entropie

Nous allons progressivement construire l'entropie pour une mesure sur un ensemble fini F , puis sur $F^{\mathbb{Z}}$, puis enfin pour tout système dynamique.

16.2.1 Entropie d'une mesure sur un ensemble fini

Posons $S(x) = -x \log(x)$. Le premier cas que nous considérerons est celui d'une mesure de probabilité μ sur un ensemble fini F . Une telle mesure n'est rien d'autre qu'une collection de nombre positifs $\mu_f = \mu\{f\}$ pour tout f de F . Par définition, l'*entropie* de la mesure μ est le nombre

$$h(\mu) = \sum_{f \in F} S(\mu_f).$$

La proposition suivante résume quelques propriétés élémentaires de l'entropie. Elles se démontrent en utilisant la convexité de la fonction S

Proposition 16.2.1.1 1. $h(\mu) \geq 0$, l'égalité étant atteinte exactement pour les mesures les moins bien distribuées pour laquelle l'un des μ_f vaut 1 et tous les autres 0.

2. $h(\mu) \leq \log |F|$, l'égalité étant atteinte exactement pour la mesure d'équilibre pour lesquelles tous les μ_f valent $1/|F|$.
3. Si p est la projection de $F \times G$ sur F , pour F et G des ensembles finis. Alors

$$h(p_*\mu) \leq h(\mu) \leq h(p_*\mu) + \log(|G|).$$

4. Additivité de l'entropie. Si μ et ν sont deux mesures sur des ensembles finis alors $S(\mu \otimes \nu) = S(\mu) + S(\nu)$

Entropie d'une partition

On se donne $\xi = \{\xi_1, \dots, \xi_n\}$ une partition d'un ensemble X équipé d'une mesure μ . L'entropie de cette partition est donnée par

$$h(\xi) = \sum_{i=1}^n S(\mu(\xi_i)).$$

En d'autres termes, $h(\xi) = h(\dot{\xi}_*\mu)$.

16.2.2 Entropie d'une mesure sur $F^{\mathbb{N}}$

La dernière proposition suggère une définition de l'entropie d'une mesure sur $F^{\mathbb{N}}$ (ou sur $F^{\mathbb{Z}}$). Notons pour cela p^n la projection de $F^{\mathbb{N}}$ sur $F^{[0,n]}$. Soit μ une mesure de probabilité sur $F^{\mathbb{N}}$, nous posons alors l'entropie de μ comme le nombre

$$h(\mu) = \limsup_{n \rightarrow \infty} \frac{h(p_*^n(\mu))}{n}$$

Par exemple, si μ_0 est une mesure sur F l'entropie de la mesure produit μ est celle de μ_0 .

16.2.3 Entropie d'une transformation

Chaque partition ξ de l'espace X , nous donne une application de codage $\hat{\xi}^+$ à valeurs dans F^n .

Nous pouvons poser alors pour μ une mesure sur X ,

$$\begin{aligned} h(T, \xi, \mu, n) &= h(\hat{\xi}_*^{[0,n]}\mu). \\ h(T, \xi, \mu) &= h(\hat{\xi}_*^+\mu) = \limsup_{n \rightarrow \infty} \frac{h(T, \xi, \mu, n)}{n}, . \end{aligned}$$

En termes de partitions, on vérifie que $h(T, \xi, \mu, n)$ est l'entropie de la partition ξ_n dont les éléments sont les intersections des éléments des partitions $\xi, T^{-1}(\xi), \dots, T^{-n}(\xi)$.

Nous appellerons enfin *entropie (de Kolmogorov-Sinaï) de la transformation T* le nombre

$$h(T, \mu) = \sup_{\xi} (h(T, \xi, \mu)).$$

Nous avons donc défini l'entropie pour les mesures sur les ensembles $F^{\mathbb{N}}$ et étendu cette notion à tous les systèmes dynamiques en utilisant l'idée de codage.

Pour le moment, nous ne savons pas grand chose sur ce nombre, il pourrait aussi bien valoir toujours $+\infty$. Nous allons dans la suite expliquer comment peut se calculer l'entropie, dans les rares cas où son calcul est possible explicitement. Nous tenterons de rendre compte de sa signification comme celle de la *mesure du désordre du système* et enfin expliquer en quoi ce nombre est un invariant fondamental des systèmes dynamiques.

16.3 Calcul de l'entropie

Tous les résultats de cette section sont sans démonstration. Une bonne référence est le livre de Rudolph [4]. Tout d'abord une première proposition précise la définition de $h(T, \xi, \mu)$.

Proposition 16.3.0.1 *Soit T une transformation d'un espace X préservant une mesure μ . Alors pour toute partition finie μ on a*

$$h(T, \xi, \mu) = \lim_{n \rightarrow \infty} \frac{h(T, \xi, \mu, n)}{n},$$

où de manière équivalente si μ est une mesure sur $F^{\mathbb{N}}$ invariante par décalage alors

$$h(\mu) = \lim_{n \rightarrow \infty} \frac{h(p_*^n \mu)}{n}.$$

Le théorème suivant permet de calculer $h(T, \mu)$ à l'aide d'une seule partition.

Théorème 16.3.0.2 [KOLMOGOROV-SINAI] *Soit ξ une partition génératrice pour une transformation inversible, alors $h(T, \xi, \mu) = h(T, \mu)$.*

Un dernier théorème permet d'accéder au calcul de l'entropie en se consacrant au voisinage d'un point.

Théorème 16.3.0.3 [SHANNON-MCMILLAN-BREIMAN] *Soit T une transformation préservant une mesure ergodique μ et ξ une partition finie. Notons $A_n(x)$ l'unique sous-ensemble de la forme $\xi_{i_0} \cap T^{-1}\xi_{i_1} \dots \cap T^{-n}\xi_{i_n}$ qui contienne x .*

Alors pour μ -presque tout x , on a

$$h(T, \xi, \mu) = \lim_{n \rightarrow \infty} \frac{-\log(\mu(A_n(x)))}{n}.$$

Ces résultats, et particulièrement celui de Kolmogorov-Sinai vont nous permettre de calculer l'entropie sur une série d'exemples simples, pour lesquelles une partition génératrice est connue.

Quelques Calculs.

1. Pour le décalage sur $F^{\mathbb{Z}}$ muni d'une mesure produit μ on a donc $h(\sigma, \mu) = h(\mu)$. Si cette mesure est une mesure produit obtenue à partir d'une mesure μ_0 sur F , on a $h(\sigma, \mu) = h(\mu_0)$. En général, on a $h(\sigma, \mu) \leq \log(|F|) = h(\sigma, \nu)$, où ν est la mesure produit obtenue à partir de la mesure équilibrée sur F .
2. Si on considère une rotation R_θ d'angle irrationnel sur S^1 , nous avons vu que la partition ξ_0 en deux demi-cercles est génératrice. Dans cette situation, la partition ξ^n est constitué d'au plus $2n$ intervalles - à chaque étape de la subdivision au plus 2 intervalles sont coupés en 2. On en déduit donc que $h(R_\theta, \xi_0, \lambda, n) \leq \log(2n)$. Ainsi

$$h(R_\theta, \mu) \leq \lim_{n \rightarrow \infty} \frac{\log(2n)}{n} = 0.$$

3. Plus généralement, on peut montrer que l'entropie d'un système hamiltonien complètement intégrable est nulle.
4. Nous allons tenter de justifier le calcul de l'entropie du difféomorphisme du tore $\mathbb{T}^2$ associé à la matrice A à coefficient entiers vu en 2. Nous appellerons *rectangles de Markov* la projection injective d'un rectangle dont les côtés sont parallèles aux directions propres de la matrice. Nous parlerons d'un côté *positif* d'un rectangle s'il est parallèle à la direction contractante, d'un côté négatif s'il est parallèle à la direction dilatante. On peut alors montrer qu'il existe une *partition de Markov* ξ , c'est-à-dire une partition génératrice en rectangles de Markov, telles que les images (resp. préimages) des côtés positifs (resp. négatifs) sont inclus dans les côtés positifs (resp. négatifs). L'existence d'une partition de Markov est une propriété générale des systèmes hyperboliques, mais le cas que nous évoquons peut se traiter directement.

Notons alors l (resp. L) la longueur du plus petit (resp. plus grand) côté positif. Soit de même c , C pour les côtés négatifs. Soit λ la plus grande valeur propre de la matrice A . La partition itérée ξ_n est elle aussi formée de rectangles de Markov dont la longueur des côtés positifs est comprise entre l/λ^n et L/λ^n , alors que la longueur des côtés négatifs - qui sont eux dilatés - reste comprise entre c et C . Nous en déduisons que pour tout rectangle ξ_i de ξ_n nous avons

$$\frac{lc}{\lambda^{-n}} \leq \mu(\xi_i) \leq \frac{LC}{\lambda^{-n}}.$$

En particulier,

$$\log(lc) - n \log(\lambda) \leq \log(\mu(\xi_i)) \leq \log(LC) - n \log(\lambda).$$

Le théorème de Shannon-McMillan-Breiman nous donne alors que

$$h(T, \mu) = \log(\lambda) = \log\left(\frac{3 + \sqrt{5}}{2}\right).$$

REMARQUE :

L'entropie d'un système dynamique est très difficile à calculer explicitement dans la plupart des cas. On arrive cependant à démontrer des inégalités sur l'entropie. On montre ainsi que l'entropie des difféomorphismes d'Anosov est toujours strictement positive. Une inégalité importante due à Y. Pesin relie l'entropie d'un difféomorphisme F à l'asymptotique des valeurs propres de $DF.DF^*$. L'entropie d'un système dynamique permet aussi de choisir une mesure invariante, celle d'entropie maximale qui est unique pour la plupart des systèmes dynamiques.

16.4 L'entropie comme invariant

Nous pouvons maintenant avec ces quelques exemples essayer de justifier l'entropie comme une *mesure du désordre d'un système*.

Nous avons vu que pour le décalage sur $F^{\mathbb{Z}}$ l'entropie est majoré par l'entropie de la mesure produit de la mesure équilibrée. Cette mesure est celle qui rend le système le plus aléatoire : tout d'abord pour une mesure produit, chaque "pas" est totalement indépendant des précédents, de manière probabiliste les applications de $F^{\mathbb{Z}}$ dans F qui associe à une suite son $n^{\text{ième}}$ élément sont des variables indépendantes. De plus chaque pas est équiprobable, il s'agit donc du système le plus aléatoire possible, celui d'une succession de tirage au dé. Au contraire, les mesures produits d'entropie nulle sur le décalage, proviennent d'une mesure supporté pas un seul élément ; il s'agit là du système le moins aléatoire possible, celui où l'on tire avec probabilité 1 toujours le même élément.

De manière plus générale, c'est-à-dire pour les mesures qui ne sont pas des produits, des théorèmes montrent que les événements dans le lointain passé et le futur éloigné vont tendre à être indépendants pour une mesure d'entropie positive. Ainsi, la mesure peut être considéré comme une mesure produit.

Signalons enfin que l'entropie est un invariant. Nous dirons que deux transformations T_1 et T_2 des espaces X_1 et X_2 respectivement munis des mesures μ_1 et μ_2 , sont *métriquement équivalents*, s'il existe deux ensembles invariants de mesure pleine E_1 et E_2 , un isomorphisme φ entre E_1 et E_2 préservant les ensembles mesurables et tel que $\varphi \circ T_1 = T_2 \circ \varphi$. Il s'agit de l'une des notions d'équivalence les plus faibles que l'on puisse imaginer sur les systèmes dynamiques.

Il est clair d'après la définition que l'entropie est un invariant pour l'équivalence métrique. Ceci a permis de distinguer - pour la première fois - différents décalages de Bernoulli.

Signalons enfin un résultat important [4] qui permet de montrer que l'entropie est le seul invariant de l'équivalence métrique des décalages de Bernoulli.

Théorème 16.4.0.4 [ORSTEIN-WEISS] *Soit deux décalages de Bernoulli - construits avec des ensembles finis éventuellement différents - pour des mesures produits. Si ces deux décalages ont la même entropie, alors ils sont métriquement équivalents.*

Les constructions utilisées dans ce théorème donnent plus : elles permettent de reconnaître si un système est équivalent ou non à un décalage de Bernoulli.

L'entropie est donc l'outil fondamental qui permet de décrire les systèmes dynamiques du point de vue de la théorie de la mesure.

Bibliographie

- [1] ??Rappels d'algèbres linéaires
- [2] ??Espaces métriques complets
- [3] V. Arnol'd *Méthodes Mathématiques de la Mécanique Classique* Editions MIR
- [4] D. Rudolph *Fundamentals of Measurable Dynamics* Oxford University Press

Univ. Paris-Sud,
Laboratoire de Mathématiques, Orsay F-91405 Cedex
CNRS, Orsay cedex, F-91405