Table des matières

1 Fonctions invariantes **3**
1.1 Définitions 3

2 Mécanique Lagrangienne **5**
2.1 Lagrangien 5
2.1.1 Equations d'Euler-Lagrange 5
2.1.2 Points critiques de l'action 5
2.1.3 Equation de Newton 7
2.1.4 Sous-variétés 7
2.1.5 Exemple et exercice : surface de révolution 7
2.2 Symétries 8
2.2.1 Exemple et exercice, mouvement dans un champ central . 9
2.2.2 Exemple et exercice, mouvement dans un champ central (bis) 10
2.2.3 Problème à deux corps 11

3 Mécanique Hamiltonienne **12**
3.1 Equations d'Hamilton 12
3.1.1 Transformation de Legendre 12
3.2 Crochet de Poisson 14
3.2.1 Algèbre des intégrales premières 14
3.2.2 Evolution des observables 15
3.3 Coordonnées symplectiques 16
3.3.1 Variables en involution 16
3.3.2 Variables conjuguées 16
3.3.3 Coordonnées symplectiques 18
3.4 Systèmes complétement intégrables 19
3.4.1 Systèmes complètement intégrables 19
3.4.2 Méthode d'Hamilton-Jacobi 19
3.4.3 Au delà des systèmes complètement intégrables 23
3.5 Toupie de Lagrange
Mouvement du solide 24

4 Mesures Invariantes **28**
4.1 Existence de mesures invariantes 30
4.2 Théorème de récurrence de Poincaré 31
4.3 Théorèmes ergodiques 34
4.3.1 Théorèmes ergodiques 34
4.4 Ergodicité du décalage de Bernouilli 37
4.4.1 Quelques remarques sur la preuve 39
4.4.2 Du décalage aux systèmes hyperboliques 39
4.4.3 Feuilletages absolument continus 41

5 Codage et entropie **42**
5.1 Codage 42
5.1.1 Codage, partition génératrice 42
5.2 Entropie 43
5.2.1 Entropie d'une mesure sur un ensemble fini 43
5.2.2 Entropie d'une mesure sur $F^{\mathbb{N}}$ 44
5.2.3 Entropie d'une transformation 44
5.3 Calcul de l'entropie 45
5.4 L'entropie comme invariant 47

Chapitre 1

Fonctions invariantes

Pour étudier un système dynamique, il est utile de chrecher 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.

Noux 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.

1.1 Définitions

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 et exercices

(i) 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 $df(X) = 0$.

(ii) 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

$$\left\{ \begin{array}{lcl} \frac{d}{dt}\dot{q} & = & -m \ \frac{\partial}{\partial q}U(q) \\ \frac{d}{dt}q & = & \dot{q}. \end{array} \right.$$

Montrez alors que l'énergie totale

$$E = \frac{1}{2}m\dot{q}^2 + U(q),$$

est une intégrale première du mouvement.

(iii) On a déjà vu que l'Hamiltonien H est une intégrale première du mouvement pour les équations de Hamilton.

$$\left\{ \begin{array}{lcl} \frac{d}{dt}p & = & -\frac{\partial}{\partial q}H, \\ \frac{d}{dt}q & = & \frac{\partial}{\partial p}H. \end{array} \right.$$

(iv) Soit x_0 un point fixe du flot ϕ^t. On suppose que le bassin d'attraction de x_0

$$B = \{y, \lim_{t\to\infty} \phi_t(y) = x_0\},$$

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 ?

(v) Plus généralement, soit γ une orbite bornée de du flot ϕ_t ; On suppose que le bassin d'attraction de γ

$$B = \{y, \lim_{t\to\infty} d(\phi_t(y), \gamma) = 0\},$$

est ouvert. Montrez alors que les seules fonctions continues invariantes définies au voisinage de γ sont constantes

Chapitre 2

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.

2.1 Lagrangien

2.1.1 Equations d'Euler-Lagrange

Soit U un ouvert d'un espace vectoriel E muni des coordonnées $q = (q_i)_{i=1\ldots n}$. 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 $q = (q_i)_{i=1,\ldots,n}$ comme précédemment. Le deuxième facteur correspond aux *vitesses* et les coordonnées correspondantes sont notées cette fois-ci $\dot{q} = (\dot{q}_i)_{i=1,\ldots,n}$. 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

$$\left\{ \begin{array}{rcl} \frac{d}{dt}\frac{\partial}{\partial \dot{q}} L(q,\dot{q}) & = & \frac{\partial}{\partial q} L(q,\dot{q}) \\ \frac{d}{dt} q & = & \dot{q}. \end{array} \right.$$

2.1.2 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 $c :]a, b[\mapsto U$ 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 $c^s : [a,b] \to U$ 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

$$\frac{d}{ds}|_{s=0} S(c^s).$$

On montre alors :

Théorème 2.1.2.1 *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 $c^s = (c_1^s, \ldots, c_n^s)$ 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

$$\begin{aligned}
\frac{d}{ds}S(c_s) &= \int_a^b \frac{\partial}{\partial s} L\big(c^s, \frac{d}{dt}c^s\big)dt \\
&= \sum_i \Big(\int_a^b \frac{\partial}{\partial q_i} L\big(c^s, \frac{d}{dt}c^s\big) \frac{\partial}{\partial s} c_i^s dt \\
&+ \int \frac{\partial}{\partial \dot{q}_i} L\big(c^s, \frac{d}{dt}c^s\big) \frac{\partial^2}{\partial s \partial t} c_i^s dt\Big).
\end{aligned}$$

En intégrant par partie le dernier terme, on obtient

$$\begin{aligned}
\frac{d}{ds}S(c^s) &= \sum_i \Big(\int_a^b \frac{\partial}{\partial q_i} L\big(c^s, \frac{d}{dt}c^s\big) \frac{\partial}{\partial s} c_i^s dt \\
&- \int_a^b \frac{d}{dt}\frac{\partial}{\partial \dot{q}_i} L\big(c^s, \frac{d}{dt}c^s\big) \frac{\partial}{\partial s} c_i^s dt\Big) \\
&+ \sum_i [\frac{\partial}{\partial \dot{q}_i} L\big(c^s, \frac{d}{dt}c^s\big) \frac{\partial}{\partial s} c_i^s]_a^b.
\end{aligned}$$

Si nous posons $w(t) = \frac{\partial}{\partial s}|_{s=0} c^s(t)$, et si nous considérons une variation à extrémités fixées, nous obtenons

$$\frac{d}{ds}|_{s=0} S(c_s) = \sum_i \int_a^b \Big(\frac{\partial}{\partial q_i} L\big(c, \frac{d}{dt}c\big) - \frac{d}{dt}\frac{\partial}{\partial \dot{q}_i} L\big(c, \frac{d}{dt}c\big)\Big) w_i dt.$$

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 $\int_a^b fgdt = 0$ pour toute fonction g telle que $g(a) = g(b) = 0$ alors $f = 0$.

$\diamond$

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.

2.1.3 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

$$L = \frac{1}{2}m\|\dot{q}\|^2 - U(q).$$

2.1.4 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.

2.1.5 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.

(i) 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.

(ii) On restreint le Lagrangien $L = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$ à S. Autreement dit on considre 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) ?

(iii) Ecrire les équations d'Euler-Lagrange dans ces coordonnées. Que se passe-t-il dans le cas du cylindre ($r = constante$) ?

2.2 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 2.2.0.1 (E. Noether) *Soit φ^s un groupe à un paramètre s de symétries de L. Alors, la fonction*

$$I(q, \dot{q}) = \sum_i \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}) \frac{d}{ds}|_{s=0} \varphi_i^s(q),$$

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

$$\begin{aligned} \frac{d}{dt} I(q, \dot{q}) &= \frac{d}{dt} \sum_i \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}) \frac{d}{ds} \varphi_i^s(q), \\ &= \sum_i \big(\frac{d}{dt} (\frac{\partial}{\partial \dot{q}_i} L(q, \dot{q})) \frac{d}{ds} \varphi_i^s(q) + \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}) \frac{d}{dt} \frac{d}{ds} \varphi_i^s(q) \big). \end{aligned}$$

On utilise les équations d'Euler-Lagrange dans le premier terme de droite pour obtenir

$$\frac{d}{dt} I(q, \dot{q}) = \sum_i \big(\frac{\partial}{\partial q_i} L(q, \dot{q}) \frac{d}{ds} \varphi_i^s(q) + \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}) \frac{d}{ds} \frac{d}{dt} \varphi_i^s(q) \big)$$

On reconnait 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 $\diamond$.

2.2.1 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

$$L = \frac{1}{2}\|\dot{q}\|^2 - U(q).$$

(i) *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 $q(t_0)$ $\not\emptyset$, 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 ?

(ii) 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.

(iii) On introduit la fonction *énergie potentielle effective*

$$V = U + \frac{M^2}{2r^2}.$$

L'introduction de cette fonction va permettre de réecrire le système comme un système en dimension 1. Plus précisemment, montrez alors que l'énergie totale s'écrit

$$E = \frac{\dot{r}^2}{2} + V(r),$$

et que les équations du mouvement sont

$$\frac{d^2}{dt^2} r = -\frac{\partial}{\partial r} V.$$

(iv) *Remarque.* La présence d'une seule intégrale première nous a permis ici de réduire la dimension de l'espace des phases de 4 a 2, alors que l'on s'attendrait à un gain d'un seul paramètre. Ceci est dû à des symétries "cachées" dont l'intreprétation est plus claire dans le cadre de la géométrie symplectique, cadre sous-jacent à la mécanique.

(v) *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.

(*vi*) *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.

2.2.2 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'hypothèses 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.

(*i*) Montrez que l'angle entre un apocentre et péricentre successif s'écrit

$$\Theta = \int_{r_{min}}^{r_{max}} \frac{M dr}{r^2\sqrt{2(E - V(r))}}.$$

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 dégénerescence 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 supposerons par la suite que Θ est constant.

(*ii*) Montrez qu'il existe des orbites circulaires pour n'importe quel rayon r.

(*iii*) 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 $r = M/x$.

(*iv*) 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 $\alpha = 0$)

(*v*) On suppose que $\lim_{r\to+\infty} U(r) = +\infty$ Montrez alors que

$$\lim_{E\to+\infty} \Theta(E, M) = \frac{\pi}{2}.$$

Indication On utilisera le changement de variables $y = r_{min}/r$ pour écrire

$$\Theta = \int_{y_{min}}^{1} \frac{dy}{\sqrt{2(Z(1) - Z(y))}}, \; Z(y) = \frac{y^2}{2} + \frac{M^2}{r_{min}^2} U\big(\frac{r_{min}^2}{My^2}\big).$$

2.2.3 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

$$L = \frac{M\|\dot{Q}\|^2}{2} + \frac{m\|\dot{q}\|^2}{2} + \frac{k(m+M)}{\|q - Q\|}.$$

(i) 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.

(ii) On suppose que cette vitesse est nulle, et on utilise donc des coordonnées dans lesquelles l'origine est le centre de gravité $mq + MQ$. 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.

(iii) 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 3

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.

3.1 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 coordonées de $\mathbb{R}^{2n}$ sont (q_i, p_i), elles s'écrivent

$$\left\{ \begin{array}{lcl} \frac{d}{dt}p & = & -\frac{\partial}{\partial q}H \\ \frac{d}{dt}q & = & \frac{\partial}{\partial p}H. \end{array} \right.$$

On introduit le champ de vecteur associé à cette équation différentielle appelé aussi *champ hamiltonien*, ou *gradient hamiltonien* ou encore *gradient symplectique*

$$X_H = (\frac{\partial}{\partial p}H, -\frac{\partial}{\partial q}H).$$

3.1.1 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.

(i) Montrez tout d'abord que

$$\frac{\partial H}{\partial p}(q, \frac{\partial L}{\partial \dot{q}}) \quad = \quad \dot{q},$$

(ii) Montrez ensuite

$$\frac{\partial H}{\partial q}(q, \frac{\partial L}{\partial \dot{q}}) \quad = \quad -\frac{\partial L}{\partial q}(q, \dot{q}).$$

(iii) 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.

(iv) 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 2.2.3.

$$L = \frac{a\dot{r}^2}{2} + \frac{ar^2\dot{\theta}^2}{2} + \frac{b}{r}.$$

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.

3.2 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 3.2.0.1 *Nous avons*

(i) *Antisymétrie* $\{f, g\} = -\{g, f\}$,

(ii) *Dérivation* $\{fh, g\} = f\{h, g\} + h\{f, g\}$,

(iii) *Identité de Jacobi* $\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0$,

(iv) $\{f, g\} = -df(X_g) = dg(X_f)$

(v) f *est une intégrale première du système associé à l'hamiltonien* H *si et seulement si* $\{f, H\} = 0$.

(vi) $X_{\{f,g\}} = [X_f, X_g]$. *En particulier, si* $\{f, g\}$ *est une fonction constante alors les flots de* X_f *et* X_g *commuttent.*

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. $\diamond$

3.2.1 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 ce 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'algébriser 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 polynome en les variables p et q. Physiquement, les fonctions sont des polynomes ... ou au pire des séries formelles

$$f = \sum_{i_1,\ldots,i_l,j_1,\ldots,j_k} a_{i_1,\ldots,i_l,j_1,\ldots,j_k} p_{i_1} \ldots p_{i_l} q_{j_1} \ldots q_{j_k}.$$

Le problème de recherche d'intégrale première devient alors algébrique : la recherche de solutions d'équations linéaires en les $a_{i_1,\ldots,i_l,j_1,\ldots,j_k}$. 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.

3.2.2 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 écrire 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 satistait 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.

3.3 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 précisemment le crochet de Poisson, s'expriment simplement. Ce seront les changements de variables adaptés à la mécanique Hamiltonienne.

3.3.1 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 3.3.1.1 *Soit Q_i des variables en involution. Soit H une fonction telle que $H = h(Q_1, \ldots, Q_l)$. Alors $\{H, Q_i\} = 0$.*

Démonstration Ceci découle de manipulations sur le crochet de Poisson

$$\begin{aligned}
\{H, Q_j\} &= \sum_i \frac{\partial H}{\partial p_i}\frac{\partial Q_j}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial Q_j}{\partial p_i} \\
&= \sum_{i,k} \frac{\partial h}{\partial Q_k}\frac{\partial Q_k}{\partial p_i}\frac{\partial Q_j}{\partial q_i} - \frac{\partial h}{\partial Q_k}\frac{\partial Q_k}{\partial q_i}\frac{\partial Q_j}{\partial p_i} \\
&= \sum_k \frac{\partial h}{\partial Q_k}\{Q_j, Q_k\} \\
&= 0.
\end{aligned}$$

$\diamond$

3.3.2 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 3.3.2.1 *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 supposerons 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

$$\{Q_i, Q_j\} = \{P_i, P_j\} = 0, \ \{P_i, Q_j\} = \delta_{ij}.$$

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 commuttent (cf (vi) de 3.2.0.1). 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,

$$T_y W \oplus V_y = \mathbb{R}^{2n}. \tag{3.1}$$

Considérons alors l'application Ψ définie de la manière suivante.

$$\left\{ \begin{array}{rcl} W \times \mathbb{R}^{n+k} & \to & \mathbb{R}^{2n} \\ (w, s_i, \dots, s_{n+k}) & \mapsto & \varphi_{s_1}^1 \circ \dots \circ \varphi_{s_{n+k}}^k(w). \end{array} \right.$$

L'application Ψ est un difféomorphisme au voisinage de x. Grâce à 3.1, on montre en effet que $D_x\Psi$ est surjective et donc inversible.

Notons enfin π, l'application

$$\left\{ \begin{array}{rcl} W \times \mathbb{R}^{n+k} & \to & \mathbb{R} \\ (w, s_i, \dots, s_{n+k}) & \mapsto & s_{k+1}. \end{array} \right.$$

Posons

$$P_{k+1} = \pi \circ \Psi^{-1}.$$

Les flots des champs de vecteurs que nous avons définis commuttent tous et nous pouvons en déduire que la fonction P_{k+1} vérifie

$$P_{k+1} \circ \varphi_s^i \quad = \quad s\delta_{ik} + P_{k+1}$$

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 $dQ_1, \ldots, dQ_n, dP_1, \ldots, dQ_k$. Pour cela il suffit de remarquer que

$$dP_{k+1}(X_{n+k+1}) = 1,$$

alors que, si $i \leq k$ et $j \leq n$,

$$dP_i(X_{n+k+1}) = dQ_j(X_{n+k+1}) = 0.$$

$\diamond$

Exercice

Adaptez la preuve ci-dessus pour montrer que l'on peut toujours compléter une familles de variables en involution non-dégénérées.

3.3.3 Coordonnées symplectiques

Nous dirons que la famille de fonctions $(Q_1, \ldots, Q_n, P_1, \ldots P_n)$ forment des *coordonnées symplectiques* au voisinage d'un point x de $\mathbb{R}^{2n}$ si nous avons les relations

$$\{P_i, P_j\} = \{Q_i, Q_j\} = 0, \ \{P_i, Q_j\} = \delta_{ij}.$$

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 3.3.3.1 *Soit $(Q_1, \ldots, Q_n, P_1, \ldots, P_n)$ des coordonnées symplectiques, alors*

(i) l'application $G = (Q_1, \ldots, Q_n, P_1, \ldots, P_n)$. est un difféomorphisme local,

(ii) De plus, si $H = h(Q) = h(Q_1, \ldots, Q_n)$ alors

$$\left\{ \begin{array}{lcl} \{Q_i, H\} & = & 0, \\ \{P_i, H\} & = & \frac{\partial h}{\partial Q_i}(Q). \end{array} \right.$$

Ce théorème a une conséquence utile sur l'intégration des équations d'Hamilton.

Corollaire 3.3.3.2 *Soit H un Hamiltonien défini sur $\mathbb{R}^{2n}$. Soit ϕ_t son flot. Notons $f^t = f \circ \phi_t$. Soit $(Q_1, \ldots, Q_n, P_1, \ldots P_n)$ des coordonnées symplectiques. Supposons qu'il existe une fonction h telle que $H = h \circ (Q_1, \ldots, Q_n)$. Alors*

$$\left\{ \begin{array}{lcl} P_i^t & = & P_i - t\frac{\partial h}{\partial Q_i}(Q), \\ Q_i^t & = & Q_i. \end{array} \right.$$

Démonstration du théorème. Nous avons déjà montré la moitié de [2], l'autre partie se montre par la même méthode. $\diamond$

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. $\diamond$

3.4 Systèmes complétement intégrables

Nous allons montrer comment la connaissance de suffisamment d'intégrales premières permet de décrire le système en choisissant des coordonnées symplectiques adaptées.

3.4.1 Systèmes complètement intégrables

Nous allons examiner le cas le plus optimiste, celui ou 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 définie sur U et telles que

(i) les formes df_i sont linéairement indépendantes,

(ii) 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 complétement 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.

3.4.2 Méthode d'Hamilton-Jacobi

La méthode d'Hamilton-Jacobi cherche à rendre un système complètement integrable. 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 $(q_1, \ldots, q_n, p_1, \ldots, p_n)$.

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'equation d'Hamilton-Jacobi.

Théorème 3.4.2.1 *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 theorème 3.3.2.1. 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 extrémement simple et renvoyons au livre d'Arnol'd [1] 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 thèorème 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

$$\sum_i (\frac{\partial P_i}{\partial q_l}\frac{\partial Q_i}{\partial p_m} - \frac{\partial P_i}{\partial p_m}\frac{\partial Q_i}{\partial q_l}) = \delta_{lm}$$

Des manipulations analogues donnent

$$\begin{aligned}\sum_i (\frac{\partial P_i}{\partial p_l}\frac{\partial Q_i}{\partial p_m} - \frac{\partial P_i}{\partial p_m}\frac{\partial Q_i}{\partial p_l}) &= 0 \\ \sum_i (\frac{\partial P_i}{\partial q_l}\frac{\partial Q_i}{\partial q_m} - \frac{\partial P_i}{\partial q_m}\frac{\partial Q_i}{\partial q_l}) &= 0.\end{aligned}$$

Traduisons ces trois dernières équations de manière algébrique. Introduisons les matrices

$$A = \begin{pmatrix} \frac{\partial P}{\partial p} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial Q}{\partial q} \end{pmatrix}, \ J = \begin{pmatrix} 0 & -Id \\ Id & 0 \end{pmatrix}.$$

Notons B^* la transposée de la matrice B. Les trois équations que nous venons d'obtenir s'écrivent

$$(JA)(JA)^* = Id.$$

Nous en déduisons

$$(JA^*)(JA^*)^* = Id.$$

Les équations correspondantes nous donnent

$$\begin{aligned}\sum_i (\frac{\partial Q_l}{\partial q_i}\frac{\partial P_m}{\partial p_i} - \frac{\partial Q_l}{\partial p_i}\frac{\partial P_m}{\partial q_i}) &= \delta_{lm} \\ \sum_i (\frac{\partial P_m}{\partial p_i}\frac{\partial P_l}{\partial q_i} - \frac{\partial P_m}{\partial q_i}\frac{\partial P_l}{\partial p_i}) &= 0 \\ \sum_i (\frac{\partial Q_m}{\partial p_i}\frac{\partial Q_l}{\partial q_i} - \frac{\partial Q_m}{\partial q_i}\frac{\partial Q_l}{\partial p_i}) &= 0\end{aligned}$$

Ces équations nous donnent les crochets de Poisson que nous recherchons.$\diamond$

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 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

$$(\frac{\partial S}{\partial r})^2 \frac{1}{2a} + (\frac{\partial S}{\partial \theta})^2 \frac{1}{2ar^2} + \frac{b}{r} = Q_1.$$

Equation que nous écrivons.

$$(\frac{\partial S}{\partial \theta})^2 = 2aQ_1 r^2 - 2abr - (\frac{\partial S}{\partial r})^2 r^2.$$

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

$$\begin{aligned} \frac{\partial S}{\partial \theta} &= \frac{\partial N}{\partial \theta} = Q_2 \\ \frac{\partial S}{\partial r} &= \frac{\partial M}{\partial r} = \frac{1}{r}\sqrt{Q_2^2 + 2abr - 2aQ_1 r^2} \end{aligned}$$

Il faut maintenant expliciter les nouvelles coordonnées symplectiques. Nous avons choisi

$$Q_1 = (\frac{\partial S}{\partial r})^2 \frac{1}{2a} + (\frac{\partial S}{\partial \theta})^2 \frac{1}{2ar^2} + \frac{b}{r} = H.$$

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

$$P_1 = \int \frac{a}{\sqrt{Q_2^2 + 2abr - 2aQ_1r^2}} dr.$$

De même, nous obtenons

$$\begin{aligned} \frac{\partial P_2}{\partial r} &= -\frac{\partial^2 S}{\partial r \partial Q_2} = -\frac{Q_2}{r\sqrt{Q_2^2 + 2abr - 2aQ_1r^2}} \\ \frac{\partial P_2}{\partial \theta} &= -\frac{\partial^2 S}{\partial \theta \partial Q_1} = -1. \end{aligned}$$

Et donc

$$\begin{aligned} P_2 &= -\theta - \int \frac{Q_2}{r\sqrt{Q_2^2 + 2abr - 2aQ_1r^2}} dr \\ &= -\theta + \int \frac{Q_2}{\sqrt{\frac{Q_2^2}{r^2} + \frac{2ab}{r} - 2aQ_1}} d(\frac{1}{r}) \\ &= -\theta + Argsh(\frac{\alpha}{r} + \beta) \text{ ou } -\theta + Arcsin(\frac{\alpha}{r} + \beta) \end{aligned}$$

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 3.3.3.2, 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 illustrait 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 prmière 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

$$\sinh(\theta - \theta_0) = \frac{\alpha}{r} + \beta$$

3.4.3 Au delà des systèmes complètement intégrables

En complèment, nous indiquons brièvement l'importance de l'étude des sytèmes complètement intégrables.

Enonçons tout d'abord un résultat important sur les systèmes totalement intégrables. Le théorème d'Arnol'd-Liouville.

Théorème 3.4.3.1 *Soit H un Hamiltonien complètement intégrable, et $Q_1, \ldots, Q_n$ 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*

$$\left\{ \begin{array}{rcl} \mathbb{R}^n & \to & M_c \\ (s_1, \ldots, s_n) & \mapsto & \varphi^1_{s_1} \circ \ldots \circ \varphi^n_{s_n}(x_0), \end{array} \right.$$

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 *tores de Liouville*.

Cette situation est bien sûr très agréable lorsque l'on cherche explicitement les solutions. Malheureusement, un système n'est qu'exceptionnellement complétement 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 un pertubation 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 sytèmes complètement intégrables, le processus de *quantification* qui permet de remonter du système classique au système quantique sous-jacent est mathématiquement justifiable. Une methode de quantification reste également disponible pour les systèmes proches des systèmes complètement intégrables.

3.5 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,

$$SO(3) = \{A \in M_3(\mathbb{R}) \ / \ AA^* = Id\}.$$

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

$$F = \mathcal{S}O(3) = \{A \in M_3(\mathbb{R}) \ / \ A = -A^*\}$$

est *l'algèbre de Lie* de $SO(3)$. L'identification est construite de la manière suivante : pour tout q de $SO(3)$, $T_qSO(3)$ est un sous-espace vectoriel de $M(3, \mathbb{R})$. L'identification est donnée par $(q, A) \in SO(3) \times F$ on associe $qA \in T_qSO(3)$. Dans cette identication, A est appelé *vitesse angulaire*.

Dans cette identification l'application tangente à L_g - la mutiplication à gauche par g - est donnée par $TL_g(w, A) = (gw, A)$. De même, l'application tangente à R_g - la mutiplication 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

(i) Montrez que l'espace des configurations du solide s'identifie au groupe $SO(3)$.

(ii) 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 $\overrightarrow{OM}$.

(iii) On veut maintenant décrire le Lagrangien associé à l'énergie cinétique. Physiquement, si on a un ensemble S - continu - de particule affectés 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 $\mathbb{R}^3$.

(iv) 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 $u_0 = \overrightarrow{OG}$.

$$U(q) = Mg \ \langle q(u_0), e_z \rangle = \lambda \langle q e_z, e_z \rangle.$$

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 $SO(3)$.

(v) On considèrera à partir demaintenant 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

$$\mu_1(q, u) = \langle e_z, q\mathcal{I}u\rangle$$

(vi) 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 canonnique (e_x, e_y, e_z) et nous considérons les rotations

$$\begin{aligned} R^1_\phi &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) \\ 0 & -\sin(\phi) & \cos(\phi) \end{pmatrix} \\ R^3_\theta &= \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned}$$

Nous considérons alors l'application

$$\Psi : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \to & SO(3) \\ (\theta, \phi, \psi) & \mapsto & R^3_\psi \circ R^1_\phi \circ R^3_\theta. \end{array} \right.$$

Nous admettrons que Ψ est un difféomorphisme d'un ouvert de $\mathbb{R}^3$ dans un ouvert dense de $SO(3)$.

(vii) Montrez maintenant que

$$\begin{aligned} & T_{\theta,\phi,\psi}\Psi(\dot\theta, \dot\phi, \dot\psi) = (\Psi(\theta, \phi, \psi), \dot\theta e_z + \dot\phi R^3_{-\theta}(e_x) + \dot\psi R^3_{-\theta} R^1_{-\phi}(e_z)) \\ = \; & (\Psi(\theta, \phi, \psi), \dot\phi\cos\theta + \dot\psi\sin\phi\sin\theta, \dot\phi\sin\theta - \dot\psi\sin\phi\cos\theta, \dot\theta + \dot\psi\cos\phi). \end{aligned}$$

($viii$) 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

(ix) 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

$$\left\{ \begin{array}{lcl} p_\theta &=& 2b\dot{\theta} + 2b\dot{\psi}\cos\phi \\ p_\phi &=& 2a\dot{\phi} \\ p_\psi &=& 2\dot{\psi}(a\sin^2\phi + b\cos^2\phi) + 2b\dot{\theta}\cos\phi \end{array} \right.$$

(x) Montrez que l'inverse de la transformation de Legendre s'écrit

$$\left\{ \begin{array}{lcl} \dot{\theta} &=& p_\theta \frac{1}{2b}(1 + \frac{b\cos^2\phi}{a\sin^2\phi}) - p_\psi \frac{\cos\phi}{2a\sin^2\phi} \\ \dot{\phi} &=& \frac{1}{2a}p_\phi \\ \dot{\psi} &=& \frac{1}{2a\sin^2\phi}(p_\psi - p_\theta\cos\phi) \end{array} \right.$$

Hamiltonien

(x) 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

$$H = \frac{1}{2a}p_\phi^2 - \frac{\cos\phi}{2a\sin^2\phi}p_\psi p_\theta + \frac{1}{4a\sin^2\phi}p_\psi^2 + \frac{1}{4b}(1 + \frac{b\cos^2\phi}{a\sin^2\phi})p_\theta^2 + \lambda\cos\phi.$$

(xi) 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 ?

(xii) Montrez enfin que l'on peut exprimez 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.

Chapitre 4

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'un grande richesse.

Nous utiliserons librement le théorème de représentation de Riesz que nous pouvons concevoir comme une définition de la notion de mesure.

Théorème 4.0.0.2 *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

(i) 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.

(ii) Soit $\gamma : [0, l] \to U$ 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.

(iii) 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

$$\frac{d}{dt} \log \det(A) = trace(\frac{dA}{dt} A^{-1}).$$

(iv) Montrez que le flot d'un gradient hamiltonien préserve la mesure de Lebesgue.

(v) *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 $\mathcal{A} = \{A_{-n}, \ldots, A_n\}$ 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.

(v) 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.

4.1 Existence de mesures invariantes

L'existence de mesures invariantes est garantie par le théorème de Kakutani-Markov.

Théorème 4.1.0.3 *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 4.1.0.4 *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,

$$\alpha_n(u) \to \alpha(u).$$

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 4.1.0.5 *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

$$\|\mu\| = \sup_{f/\forall x \in K, |f(x)| \leq 1} \{|\int f d\mu|\} \leq 1.$$

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

$$\mu_{n_p} \overset{p \to \infty}{\rightharpoonup} \mu_\infty.$$

Nous allons montrer que μ_∞ est invariante. Remarquons que

$$\|T_* \mu_n - \mu_n\| = \frac{1}{n} \|\mu - T_*^n \mu\| \leq \frac{2}{n}.$$

Ainsi

$$T_* \mu_n - \mu_n \rightharpoonup 0.$$

Par continuité de T_*, nous en déduisons

$$T_* \mu_\infty = \mu_\infty. \diamond$$

4.2 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 recurrence de Poincaré. Introduisons une définition. Le *support* d'un 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 4.2.0.6 *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*

$$\|y - x\| \leq \epsilon, \ \|\phi_{nT}(y) - x\| \leq \epsilon.$$

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 $U = Supp(\mu) \cap B(x, \epsilon)$. 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

$$\sum_{i=1}^{i=n} \mu(U_n) = n\mu(U) \stackrel{n\to\infty}{\longrightarrow} \infty.$$

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 parait 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 probalibilité d'un état initial. C'est le principe de base de la physqiue 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

(i) Soit x_0 un point fixe du flot ϕ^t. On suppose que le bassin d'attraction de x_0

$$B = \{y, \lim_{t\to\infty} \phi_t(y) = x_0\},$$

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.

(ii) 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.

(iii) *Approximations simultanées de réels.* On considère le *tore*

$$\mathbb{T}^n = \underbrace{S^1 \times \ldots \times S^1}_{n} \subset \mathbb{C}^n.$$

(a) Montrez que $\mathbb{T}^n$ est une sous-variété de $\mathbb{C}^n$.

(b) On considère les n-champs champs de vecteurs, définis en complexe par

$$\partial_{\theta_k}(z_1, \ldots, z_n) = (z_1, \ldots, iz_k, \ldots, z_n).$$

Soit φ_t^k le flot de ∂_{θ_k}. Montrez que

$$\varphi_t^k(z_1, \ldots, z_n) = (z_1, \ldots, e^{it} z_k, \ldots, z_n).$$

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

$$\left\{ \begin{array}{ccc} [0, 2\pi]^n & \to & \mathbb{T}^n \\ (t_1, \ldots, t_n) & \mapsto & \varphi_{t_1}^1 \circ \ldots \circ \varphi_{t_n}^n(x). \end{array} \right.$$

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 $(\lambda_1, \ldots, \lambda_n)$ n nombres réels. Montrez à l'aide du théorème de récurrence de Poincaré que

$$\forall \epsilon > 0,\ \forall k_0,\ \exists k \geq k_0,\ \exists p_i \in \mathbb{Z}\ /\ \ \sup_i \|\lambda_i - \frac{p_i}{k}\| \leq \frac{\epsilon}{k}.$$

Indication. Utilisez le flot

$$\phi_t(x) = \psi_x(2\pi t\lambda_1, \ldots, 2\pi t\lambda_n).$$

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.

4.3 Théorèmes ergodiques

Nous dirons que la mesure μ est *ergodique* pour la tranformation 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.

Premiers exemples

(i) 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 reéls strictement positifs, μ_i des mesures invariantes non proportionnelles.

(ii) *Rotations.* On considère S^1 muni de sa meusure invariante par rotation. Montrez que les rotations d'angle irrationnel sont ergodiques. Qu'en est-il des rotations d'angle rationnel ?

(iii) *Décalage de Bernouilli.* Nous verrons par la suite que le décalage de Bernouilli est ergodique pour les mesures produits.

4.3.1 Théorèmes ergodiques

Le théorème ergodique de Birkhoff est d'une grande importance. Il s'énonce "physiquement" de la manière suivante : *les moyennes temporelles et spatiale 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 4.3.1.1 *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,

$$\lim_{n\to\infty}\frac{1}{n}\big(\sum_{i=1}^{i=n}\varphi(T^n(x)\big) = \lim_{n\to\infty}\frac{1}{n}\big(\sum_{i=1}^{i=n}\varphi(T^{-n}(x)\big) = \bar{\varphi}(x).$$

De plus,

$$\int \varphi d\mu = \int \bar{\varphi} d\mu.$$

Remarques

(i) Le cas souvent utilisé est celui des transformations ergodiques pour une mesure de probabilité ; dans ce cas $\bar{\varphi}$ est une constante qui vaut

$$\bar{\varphi} = \int \varphi d\mu.$$

(ii) 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 4.3.1.2 *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*

$$\lim_{n\to\pm\infty}\|\frac{1}{|n|}\big(\sum_{i=0}^{i=n}\varphi\circ T^n\big) - \bar{\varphi}\|_{L^2} = 0.$$

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 4.3.1.3 *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*

$$\lim_{n\to\infty}\|\frac{1}{n}\big(\sum_{i=1}^{i=n}U^n(u)) - \pi(u)\big)\| = 0.$$

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

$$\|\frac{1}{n}\big(\sum_{i=0}^{i=n}U^n(v)\big)\| \to 0.$$

Or $V = Ker(1-U)$ et donc $V^\perp = Im(1-U^*)$, où U^* est l'adjoint de U. Ainsi $v = (1-U^*)(w)$. Nous avons alors

$$\|\frac{1}{n}\big(\sum_{i=1}^{i=n} U^n(v)\big)\| = \|\frac{1}{n}\big(\sum_{i=1}^{i=n} U^n(w) - \sum_{i=1}^{i=n} U^n U^*(w)\big)\|$$

Comme U est unitaire on a $UU^* = 1$. Ainsi

$$\|\frac{1}{n}\big(\sum_{i=1}^{i=n} U^n(v)\big)\| = \|\frac{1}{n}\big(w - U^n(w)\big)\| \leq \frac{2\|w\|}{n} \to 0.\diamond$$

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 4.3.1.4 *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*

$$\lim_{n\to\infty} \frac{1}{n}\big(\sum_{i=1}^{i=n} \varphi(T^n(x)) = \int \varphi d\mu.$$

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.$\diamond$

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.

4.4 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 $\{0, 1, \ldots, M\}$. 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 appelerons 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

$$\begin{aligned} f \in \mathcal{F}_g^- \quad &\Longrightarrow \quad \lim_{n\to-\infty} d(\sigma^n(f), \sigma^n(g)) = 0 \\ f \in \mathcal{F}_g^+ \quad &\Longrightarrow \quad \lim_{n\to+\infty} d(\sigma^n(f), \sigma^n(g)) = 0. \end{aligned}$$

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 appelerons *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

$$g \in G_f \implies \in E \cap \mathcal{F}_f^+ \cap \mathcal{F}_g^- \neq \emptyset.$$

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

$$\lim_{n\to+\infty} \frac{1}{n}\big(\sum_{i=1}^{i=n} \varphi \circ \sigma^n(x)\big) = \lim_{n\to+\infty} \frac{1}{n}\big(\sum_{i=1}^{i=n} \varphi \circ \sigma^{-n}(x)\big) = \bar{\varphi}(x).$$

Remarquons maintenant que φ est absolument continue et bornée par une constante K. Ainsi pour tout ϵ strictement positif, il existe α tel que

$$d(f,g) \leq \alpha \implies |\varphi(f) - \varphi(g)| \leq \epsilon.$$

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

$$\begin{aligned} |\frac{1}{n}\big(\sum_{i=1}^{i=n} \varphi\sigma^n(f)\big) - \frac{1}{n}\big(\sum_{i=1}^{i=n} \varphi\sigma^n(g)\big)| \quad &\leq \quad \frac{1}{n}\sum_{i=n}^{i=n} |\varphi\sigma^n(f) - \varphi\sigma^n(f)| \\ &\leq \quad \frac{1}{n}\big(2Kn_0 + \alpha(n-n_0)\big). \end{aligned}$$

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 $\mathcal{F}_f^- \cup \mathcal{F}_g^+$. $\diamond$.

4.4.1 Quelques remarques sur la preuve

(i) 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.

4.4.2 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 appelerons 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 appelerons *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

$$(\pi^+, \pi^-) \ : \ U \to (\mathcal{F}_{x_0}^+ \cap U) \times (\mathcal{F}_{x_0}^- \cap U)$$

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

$$\begin{array}{lll} \forall y \in \mathcal{F}_x^+, \quad d(T^n(x), T^n(y)) & \overset{n \to +\infty}{\longrightarrow} & 0 \\ \forall y \in \mathcal{F}_x^-, \quad d(T^n(x), T^n(y)) & \overset{n \to -\infty}{\longrightarrow} & 0 \end{array}$$

Exemples

(i) 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.

(ii) 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 locale et que A est hyperbolique

(iii) 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_xU, tels que

$$\begin{array}{rcl} T_xU & = & E_x^+ \oplus E_x^- \\ \forall u \in E_x^\pm, \forall n \geq 0 & , & \|DT^{\pm n}(u)\| \leq \|u\| C e^{\mp An}. \end{array}$$

Quand E est réduit à un point, on reconnait 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 $E_y^\pm$.

4.4.3 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, telle que l'application $\pi = (\pi^+, \pi^-)$,

$$\pi = (\pi^+, \pi^-) \ : \ U \to (\mathcal{F}_{x_0}^+ \cap U) \times (\mathcal{F}_{x_0}^- \cap U)$$

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 5

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.

5.1 Codage

Commençons par une remarque. Si T est une transformation d'un espace X, nous en déduisons une application $\mathcal{T}$

$$\left\{ \begin{array}{lcl} X & \to & X^{\mathbb{Z}} \\ x & \mapsto & (\ldots, T^{-n}(x), \ldots, T^{-1}(x), x, T(x), \ldots T^{n}(x), \ldots). \end{array} \right.$$

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 $X^{\mathbb{Z}}$.

5.1.1 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'élement de F, tel que

$$x \in \xi_{\dot{\xi}(x)}.$$

Posons, pour tout sous-ensemble A de $\mathbb{Z}$,

$$\hat{\xi}^A : \left\{ \begin{array}{ccc} X & \to & F^A \\ x & \mapsto & \big(\dots, \dot{\xi}(T^n(x)), \dots\big)_{n \in A} \end{array} \right.$$

Nous poserons enfin $\hat{\xi} = \hat{\xi}^{\mathbb{Z}}$ (resp. $\hat{\xi}^+ = \hat{\xi}^{\mathbb{Z}}$) 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

(i) Donnez une partition génératrice pour le décalage de Bernouilli.
(ii) 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.
(iii) Les difféomorphismes d'Anosov admettent des partitions génératrices (cf 4.4.2).

5.2 Entropie

Nous allons progressivment constuire l'entropie pour une mesure sur un ensemble fini F, puis sur $F^{\mathbb{Z}}$, puis enfin pour tout système dynamique.

5.2.1 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 élementaires de l'entropie. Elles se démontrent en utilisant la convexité de la fonction S

Proposition 5.2.1.1 (i) *$h(\mu) \geq 0$, l'égalité étant atteinte exactement pour les mesures les moins bien distribuées pour lesquelle l'un des μ_f vaut 1 et tous les autres 0.*

(*ii*) $h(\mu) \leq \log |F|$, *l'égalité étant atteinte exactement pour la mesure d'équilibre pour lesquelles tous les* μ_f *valent* $1/|F|$.

(*iii*) *Si* p *est la projection de* $F \times G$ *sur* F, *pour* F *et* G *des ensembles finis. Alors*

$$h(p_*\mu) \leq h(\mu) \leq h(p_*\mu) + \log(|G|).$$

(*iv*) 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 $\xi = \{\xi_1, \ldots, \xi_n\}$ une partition d'un ensemble X equipé d'une mesure μ. L'*entropie* de cette partition est donnée par

$$h(\xi) = \sum_{i=1}^{i=n} S(\mu(\xi)).$$

En d'autre termes, $h(\xi) = h(\dot{\xi}_*\mu)$.

5.2.2 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

$$h(\mu) = \limsup_{n \to \infty} \frac{h(p^n_*(\mu))}{n}$$

Par exemple, si μ_0 est une mesure sur F l'entropie de la mesure produit μ est celle de μ_0.

5.2.3 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,

$$\begin{aligned} h(T, \xi, \mu, n) &= h(\hat{\xi}_*^{[0,n]}\mu). \\ h(T, \xi, \mu) &= h(\hat{\xi}_*^{+}\mu) = \limsup_{n \to \infty} \frac{h(T, \xi, \mu, n)}{n}, . \end{aligned}$$

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 $\xi, T^{-1}(\xi), \ldots, T^{-n}\xi$.

Nous appelerons 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 sytèmes 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.

5.3 Calcul de l'entropie

Tous les résultats de cette section sont sans démonstration. Un bonne référence est le livre de Rudolph [2]. Tout d'abord une première proposition précise la définition de $h(T, \xi, \mu)$.

Proposition 5.3.0.1 *Soit T une transformation d'un espace X préservant une mesure μ. Alors pour toute partition finie μ on a*

$$h(T, \xi, \mu) = \lim_{n\to\infty} \frac{h(T, \xi, \mu, n)}{n},$$

où de manière équivalente si μ est une mesure sur $F^{\mathbb{N}}$ invariante par décalage alors

$$h(\mu) = \lim_{n\to\infty} \frac{h(p_*^n \mu)}{n}.$$

Le théorème suivant permet de calculer $h(T, \mu)$ à l'aide d'une seule partition.

Théorème 5.3.0.2 [Kolmogorov-Sinaï] *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 5.3.0.3 [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 $\xi_{i_0} \cap T^{-1}\xi_{i_1} \ldots \cap T^{-n}\xi_{i_n}$ qui contienne x.*

Alors pour μ-presque tout x, on a

$$h(T, \xi, \mu) = \lim_{n\to\infty} \frac{-\log(\mu(A_n(x))}{n}.$$

Ces résultats, et particulièrement celui de Kolmogorov-Sinaï vont nous permettre de calculer l'entropie sur une série d'exemples simples, pour lesquelles une partition génératrice est connue.

Quelques Calculs.

(i) 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 équidistribuée sur F.

(ii) 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 subsidivision au plus 2 intervalles sont coupés en 2. On en déduit donc que $h(R_\theta, \xi_0, \lambda, n) \leq \log(2n)$. Ainsi

$$h(R_\theta, \mu) \leq \lim_{n\to\infty} \frac{\log(2n)}{n} = 0.$$

(iii) Plus généralement, on peut montrer que l'entropie d'un système hamiltonien complètement intégrable est nulle.

(iv) 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 4.4.2. Nous appelerons *rectangles de Markov* la projection injective d'un rectangle dont les côtés sont parrallèles aux directions propres de la matrice. Nous parlerons d'un côté *positif* d'un rectangle s'il est parrallèle à la direction contractante, d'un côté négatif s'il est parrallèle 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 traîter 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

$$h(T, \mu) = \log(\lambda) = \log(\frac{3+\sqrt{5}}{2}).$$

Remarques

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 dûe 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.

5.4 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 équidistribuée. 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 léatoire 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 ou 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 [2] qui permet de montrer que l'entropie est le seul invariant de l'équivalence métrique des décalages de Bernouilli.

Théorème 5.4.0.4 [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 Bernouilli.

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] V. Arnol'd *Méthodes Mathématiques de la Mécanique Classique* Editions MIR

[2] D. Rudolph*Fundamentals of Measurable Dynamics* Oxford University Press