

---

# **Symétries et Moment**

---

**Patrick Iglesias**



**À Alexandre, Lya et Nelly**



---

# Introduction

---

Ce livre est le texte d'un cours donné entre 1996 et 1997 à des étudiants de second et troisième cycles allemands dans le cadre des écoles d'été du *Studenstiftung*. Je profite de l'occasion qu'il m'est faite ici pour remercier le professeur Alan Huckleberry qui m'a invité à y participer et qui a permis ainsi la rédaction de ces notes.

Il existe déjà beaucoup de traités de mécanique analytique/symplectique et l'on peut légitimement se demander pourquoi en rajouter ? La question est pertinente. . . Parmi tous ces livres de mécanique, j'en vois deux qui peuvent faire office de références : *Structure des systèmes dynamiques* de Jean-Marie Souriau [Sou2] et *Méthodes mathématiques de la mécanique classique* de Vladimir I. Arnold [Arn1]. Sans oublier un livre plus ancien, qu'il faut impérativement inclure dans cette courte liste : la *Mécanique Analytique* de Joseph-Louis Lagrange [Lag8], édition de 1811/1815. Même si cela peut paraître désuet ; la mécanique symplectique moderne en est, il n'y a pas de doute, l'héritière directe. La lecture de ce chef d'œuvre est une source de compréhension profonde pour celui qui veut dépasser le simple stade du savoir-faire technique. Ces livres sont excellents et peuvent satisfaire la curiosité de chacun comme ils ont satisfait la mienne. Ils sont tous, à leur manière, complets ; cependant, ils laissent dans l'ombre quelques constructions qui méritaient, selon moi, d'être développées davantage. C'est ce qui a motivé la rédaction de ce livre.

Dans son ouvrage, Souriau consacre un tout petit chapitre à définir la structure symplectique de l'espace des mouvements d'un système dynamique à partir d'un lagrangien. Il préfère l'introduire directement à partir du système des

forces en présence — à la Newton, j'ai envie de dire. Pourtant, contrairement à ce que l'on trouve dans la littérature, sa méthode — à partir d'un lagrangien homogène — est ingénieuse et élégante. Elle évite de nombreuses constructions fastidieuses que l'on trouve habituellement dans les traités de mécanique symplectique. Elle introduit directement le meilleur cadre pour l'étude des géodésiques d'une variété riemannienne, ou de systèmes analogues. En même temps cette méthode ne se résume pas à la construction hamiltonienne ordinaire. Elle fait naître des questions (auxquelles je n'ai pas répondu dans ce livre) sur les situations critiques de lagrangiens hors du cadre régulier des hypothèses. L'approche « lagrangienne » des systèmes de la mécanique est suffisamment répandue en physique théorique pour justifier un développement particulier de cette méthode : c'est ce que j'ai essayé de faire dans la deuxième partie de ce cours. J'espère qu'en développant cette idée je ne l'ai pas étouffée, ce qui est souvent le risque de ce genre d'entreprise.

Mais au fond pourquoi faire de la mécanique symplectique ? C'est aussi une question que l'on peut se poser, surtout après en avoir fait sa préoccupation quasi-quotidienne pendant de longues années. Une façon de se rassurer, de se convaincre que l'on n'a pas perdu son temps, est de se retourner vers l'histoire et l'interroger : qui a introduit la géométrie symplectique, pourquoi, quelles questions a-t-elle permis de résoudre ? C'est cette réponse, historique, que j'ai essayé de traduire dans le premier chapitre.

C'est Lagrange qui a introduit les premiers éléments de calcul symplectique entre 1808 et 1811. Cette structure s'est imposée d'elle-même lorsqu'il a voulu étendre un résultat de Laplace sur la stabilité du grand axe des planètes. C'est donc vers Lagrange qu'il faut se tourner pour comprendre le sens de cette théorie, dont le développement moderne, actuel, continue d'apporter de nouvelles réponses à d'anciennes questions.

La réduction des systèmes dynamiques possédant des symétries particulières est un de ces problèmes que la géométrie symplectique moderne a permis de résoudre d'une façon tout à fait nouvelle. À cette question est attaché aujourd'hui le nom d'Emmy Noether et le théorème célèbre qui porte son nom. Mais la version moderne de son théorème est la construction de l'*application moment*, associée

naturellement à l'action d'un groupe de Lie préservant une 2-forme fermée. La nature de cette application moment est assez claire lorsqu'il s'agit du cadre originel du théorème de Noether : lorsque la 2-forme est issue d'un lagrangien invariant. Elle est plus obscure dans le cas général de la mécanique symplectique. Cette (relative) obscurité se retrouve dans les diverses propriétés de variance de cette application moment et des différentes cohomologies qui lui sont associées. Il suffirait d'« intégrer » la 2-forme en question pour que tout s'éclaire et que la situation particulière du théorème de Noether devienne la situation générale, et c'est bien ce que l'on peut et doit faire. Mais bien sûr tout a un prix, et cette simplification nous demandera d'introduire des espaces et groupes bizarres, quotients irrationnels de la droite réelle, de construire avec ces groupes des fibrés au-dessus de nos variétés, espaces qui ne sont plus eux-mêmes tout à fait des variétés. Mais qu'importe, à la fin l'application moment se révèle dans toute sa simplicité ; c'est en tous les cas mon opinion, même si la simplicité est une notion bien subjective et changeante.

Voilà, ce sont quelques leçons, plutôt destinées à aiguïser votre curiosité sur le sujet qu'à prétendre la satisfaire.



---

# Lagrange et les origines du calcul symplectique

---

Entre 1808 et 1811, Lagrange développe une *théorie de la variations des constantes* appliquée aux problèmes de la mécanique [Lag5, Lag6, Lag7], une application particulière de la *théorie générale de la variation des constantes* qu'il a introduite en 1775 [Lag2]. C'est l'acte de naissance du calcul *symplectique*<sup>1</sup>, terme qui ne sera inventé qu'en 1946 par Hermann Weyl<sup>2</sup> [Wey]. Le but que poursuit alors Lagrange est la généralisation d'un théorème de Laplace sur la stabilité séculaire du grand axe de l'orbite elliptique d'une planète, perturbée par l'attraction d'autres corps célestes.

Depuis Kepler on sait *résoudre* explicitement le problème des éphémérides des planètes, c'est-à-dire, calculer avec une précision aussi grande que l'on veut la position de la terre (ou de toute autre planète) connaissant sa position et sa vitesse à un instant donné, à condition toutefois de seulement considérer l'attraction du Soleil et de négliger l'influence des autres planètes. Mais bien que ce savoir soit important, il est largement insuffisant pour ce qui est du mouvement réel des planètes. L'influence des autres planètes sur la Terre est-elle vraiment négligeable, et ne va-t-elle pas à terme déstabiliser sa trajectoire et l'expulser aux confins de l'espace ?

---

<sup>1</sup>Voir l'annexe C pour une introduction à la géométrie symplectique.

<sup>2</sup>Voir la note historique à la fin de ce chapitre.

Il faut donc traiter le problème dans sa globalité : calculer la position d'une planète quelconque, connaissant les positions et vitesses de toutes les planètes, et en ne négligeant l'influence d'aucune d'entre elles. La difficulté de cette question donne le vertige, et on ne sait y répondre, encore actuellement, ni analytiquement ni même numériquement.

On pourrait croire, en effet, qu'avec l'avènement de l'ordinateur cette question soit devenue académique : pourquoi ne pas intégrer naïvement les équations du mouvement par une méthode numérique quelconque ? Malheureusement, si les erreurs d'approximations, inévitables dans ce genre de calcul, sont négligeables sur un bref intervalle de temps, elles deviennent catastrophiques à long terme. Cette incertitude sur la position de la planète n'a rien à voir avec une éventuelle situation chaotique du système (le système à deux corps est d'ailleurs parfaitement *intégrable* dans tous les sens raisonnables que l'on veut bien donner à ce mot), elle est simplement la conséquence de l'accumulation des erreurs commises par l'ordinateur lors de l'intégration numérique des équations du mouvement. L'existence d'une méthode analytique d'intégration du mouvement est donc capitale pour résoudre convenablement cette question. Si cette remarque est vraie pour le problème à deux corps, elle l'est *a fortiori* pour le problème à  $n$  corps (*i.e.* un nombre quelconque de planètes en interactions). Or, comme nous l'avons déjà dit, nous ne connaissons toujours aucune méthode analytique satisfaisante susceptible de résoudre cette question. Lagrange a contourné cette difficulté en appliquant de façon astucieuse sa méthode de *la variation des constantes* aux problèmes de la mécanique analytique. Décrivons rapidement ce dont il s'agit.

**L'espace des mouvements keplériens :** considérons d'abord un corps matériel (une planète) attiré par un centre fixe (le Soleil) selon la loi de la gravitation universelle. Les équations différentielles qui décrivent son mouvement sont d'ordre deux dans l'espace à trois dimensions : il faudra donc six *constantes d'intégration*<sup>3</sup> pour le décrire. D'après Newton, nous savons que la trajectoire de ce corps est une

---

<sup>3</sup>À l'époque de Lagrange on utilisait ce terme de *constantes d'intégration* ; nous parlons aujourd'hui d'*espace de solutions*. Par exemple, l'équation différentielle ordinaire réelle  $dx/dt = x$  a toutes ses solutions de la forme  $x(t) = c \exp(t)$ , où  $c$  est une constante arbitraire — la fameuse constante d'intégration. Or  $c$  caractérise justement cette solution.

ellipse<sup>4</sup>, dont le foyer est le centre d'attraction<sup>5</sup>. Pour décrire complètement cette ellipse, il nous faut d'abord connaître le plan dans lequel elle s'inscrit (le plan de l'orbite) : on peut le repérer par le vecteur unitaire qui lui est orthogonal, ce qui fait deux paramètres. Pour définir l'ellipse dans son plan on peut choisir la position du deuxième foyer, ce qui donne deux nouveaux paramètres, et la longueur de l'ellipse, soit au total cinq paramètres pour situer et décrire la trajectoire du corps dans l'espace. Connaissant ces cinq paramètres nous pouvons tracer alors l'ellipse par la *méthode du jardinier* : plantons dans le plan de l'orbite deux piquets, un à chacun des foyers, entourons ces deux piquets d'une corde de la longueur donnée, puis à l'aide d'un bâton tendons cette corde et faisons lui parcourir (corde tendue) un tour complet ; la figure ainsi obtenue est l'ellipse en question.

Mais si ces cinq paramètres suffisent à définir complètement la trajectoire du corps céleste, ils ne suffisent pas à déterminer son *mouvement*. En effet, comment déterminer la position de la planète à chaque instant sur sa trajectoire si nous ne connaissons pas sa position à une origine des temps arbitraire ? Ou encore la date de son passage<sup>6</sup> à l'aphélie ? C'est le sixième paramètre qu'il faut introduire et que les astronomes appellent l'*époque*.

Nous aurions pu tout aussi bien choisir six autres paramètres : par exemple les position et vitesse initiales de la planète à l'origine des temps. Ils définissent aussi, de façon unique, le mouvement de la planète. Seul le caractère pratique de tel ou tel ensemble de paramètres peut influencer notre choix. Les astronomes appellent *éléments keplériens* de la planète un ensemble de six paramètres caractérisant son mouvement, cinq pour la figure de l'ellipse et un autre pour la loi horaire : l'époque.

L'ensemble des mouvements de la planète considérés indépendamment du choix des paramètres qui nous servent à les décrire<sup>7</sup> sera appelé *espace des mou-*

---

<sup>4</sup>Si Kepler a découvert le mouvement elliptique des planètes, c'est Newton qui l'a « déduit » de la loi de la gravitation universelle qui porte son nom. Pour une discussion plus approfondie sur ce sujet voir la thèse de François de Gandt [dGa].

<sup>5</sup>Les caractéristiques géométriques de cette ellipse étant, par ailleurs, liées aux position et vitesse initiales du corps.

<sup>6</sup>C'est un peu comme si nous connaissions le trajet d'un autobus sans connaître l'heure de ses passages aux arrêts, cela ne nous serait de peu d'utilité.

<sup>7</sup>Un ensemble muni de différentes familles de paramètres servant à repérer ses éléments est

*vements kepleriens*. Un point de cet espace, que nous noterons  $\mathcal{K}$  par la suite, représente donc un *mouvement keplerien* dans son intégralité : sa trajectoire et sa loi horaire. La façon de le repérer par six constantes est une question de commodité qui peut être dictée par les conditions particulières du problème que l'on étudie.

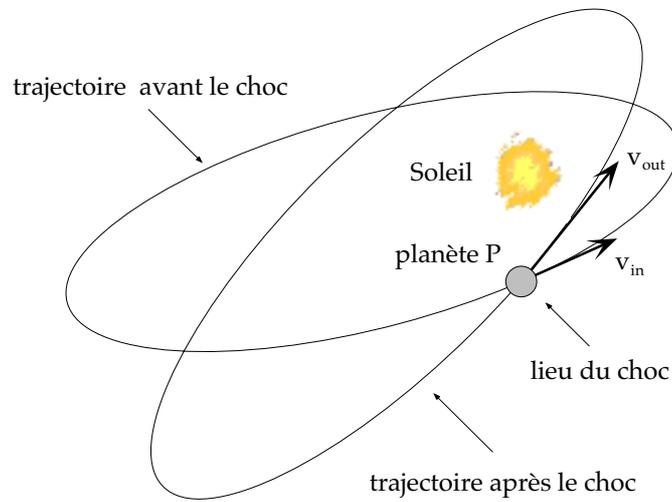


Figure II.1 : Méthode de la variation des constantes

**La méthode de la variation des constantes :** établissons maintenant la méthode de la variation des constantes telle qu'elle a été introduite par Lagrange. Lorsque la planète suit un mouvement keplerien, son état est complètement caractérisé par les 6 éléments kepleriens de son orbite qui définissent, nous l'avons dit, à la fois la figure de l'ellipse et sa loi horaire. Ce mouvement est un point  $m$  de l'espace  $\mathcal{K}$ . Supposons maintenant que la planète, qui suit le mouvement keplerien  $m$ , subisse un choc instantané dû à l'impact d'un astéroïde. Après le choc, elle suivra encore un mouvement keplerien  $m'$  différent du précédent. C'est appelée une *variété*. Par son utilisation des constantes d'intégration, pour repérer les différents mouvements de la planète, Lagrange traite de façon moderne l'ensemble des mouvements de la planète comme une variété.

une autre ellipse<sup>8</sup> parcourue selon une nouvelle loi horaire. Le mouvement (perturbé) de cette planète est donc décrit par son mouvement  $m$  avant le choc, son mouvement  $m'$  après le choc et l'instant  $t$  du choc. Supposons ensuite que la planète subisse une série de chocs de ce type. Le mouvement réel de la planète est donc décrit par une courbe dans l'espace des mouvements kepleriens, discontinue et constante par morceaux, chaque morceau de courbe décrivant le mouvement keplerien de la planète entre deux chocs successifs. En étendant ce raisonnement, Lagrange assimile l'interaction des autres planètes du système à une série infinie de chocs « infiniment petits et continus ». Il décrit ainsi le mouvement réel de la planète perturbée par une courbe, cette fois différentiable, tracée dans son espace des mouvements kepleriens. C'est ce que représente, de façon rudimentaire, la figure précédente (fig. II.1). C'est en précisant l'équation différentielle de cette courbe<sup>9</sup> qu'il a fait apparaître la *structure symplectique* de l'*espace des mouvements*. Il a donné l'expression des composantes de la forme symplectique de l'espace des mouvements kepleriens dans le système de coordonnées que sont les *éléments de la planète*. Il en a déduit entre autre la stabilité séculaire du grand axe des planètes.

J'ai essayé, dans ce chapitre, d'être le plus fidèle possible aux textes de Lagrange, désirant par là mettre en évidence le processus qui lui a permis, en voulant résoudre le problème du système des planètes, d'élaborer les premiers éléments de calcul symplectique.

NOTE. (PARENTHÈSES ET CROCHETS DE LAGRANGE) C'est le 22 août 1808 que Lagrange présente à l'Institut de France son *Mémoire sur la théorie des variations des éléments des planètes* [Lag5] où apparaissent pour la première fois ce que l'on nomme aujourd'hui les *parenthèses de Lagrange*, qui sont en termes modernes les composantes covariantes de la *forme symplectique* naturelle, de l'espace des mouvements d'une planète. Elles apparaissent comme les coefficients des variations des éléments de la planète dans l'expression de la force de perturbation, rapportée aux éléments kepleriens.

---

<sup>8</sup>Si le choc n'a pas été trop violent !

<sup>9</sup>Cette équation est connue aujourd'hui sous le nom d'*équation de Hamilton*, mais Sir W.R. Hamilton n'avait que six ans lorsque Lagrange la publia pour la première fois.

Ce mémoire a été suivi de celui *Sur la théorie générale de la variation des constantes arbitraires* [Lag6] présenté le 13 mars 1809, où Lagrange généralise sa méthode à tous les problèmes de mécanique. Ce mémoire contient en annexe une version notablement simplifiée, et définitive, de ses calculs. Le 19 février 1810, il publie un texte [Lag7] contenant les formules explicites d'inversion des *parenthèses* introduites dans son premier texte de 1808<sup>10</sup>. Ces coefficients apparaissent sous forme de crochets, les *crochets de Lagrange*, qui sont les composantes contravariantes de la forme symplectique naturelle de l'espace des éléments kepleriens de la planète. Ils apparaissent comme les coefficients de la force de perturbation, rapportée aux éléments de la planète, dans l'expression de la variation des éléments kepleriens.

C'est à partir de ces textes que Lagrange a écrit les chapitres relatifs à la dynamique dans la deuxième édition<sup>11</sup> de son traité de *Mécanique Analytique* [Lag8], publié en 1815, après sa mort<sup>12</sup>.

NOTE. (LE GROUPE SYMPLECTIQUE) Dans son ouvrage sur les groupes classiques [Wey], Hermann Weyl baptise *groupe symplectique* le groupe des transformations linéaires de  $\mathbf{R}^{2n}$  qui préservent la forme bilinéaire antisymétrique canonique<sup>13</sup>. Les relations étroites entre la structure définie par  $\omega$  et la structure complexe ( $\mathbf{R}^{2n} \sim$

<sup>10</sup>Il y a, à ce propos, une interaction entre Lagrange et Poisson. Elle est commentée dans l'annexe A.

<sup>11</sup>Dans la seconde partie, de la cinquième à la septième section

<sup>12</sup>Entre temps, sous l'influence d'un échange avec Poisson, les *parenthèses* sont devenues des *crochets* et vice-versa, ce qui rend la lecture comparée de ces textes malaisée.

<sup>13</sup>Soient  $X = (u, v)$  et  $X' = (u', v')$  deux vecteurs de  $\mathbf{R}^{2n} = \mathbf{R}^n \times \mathbf{R}^n$ , la forme symplectique canonique  $\omega$  est définie par

$$\omega(X, X') = u \cdot v' - u' \cdot v,$$

où le point désigne le produit scalaire ; autrement dit :

$$\omega(X, X') = \sum_{i=1}^n u_i v'_i - u'_i v_i$$

Le groupe symplectique est le groupe des transformation linéaires  $A$  de  $\mathbf{R}^{2n}$  tel que

$$\omega(AX, AX') = \omega(X, X')$$

pour tout  $X \in \mathbf{R}^{2n}$ . Pour une présentation plus générale de la géométrie symplectique voir l'annexe C.

$\mathbf{C}^n$ ) lui fait choisir le mot *symplectique* [du grec *sum-plektikos*], transposition de *complexe* [du latin *com-plexus*] pour désigner ce groupe, le mot *complexe* étant par ailleurs réservé.

symplectique :	sum	–	plecticos
	↓		↓
complexe :	com	–	plexus

Le suffixe *plekticos*  $\sim$  *plexus* signifiant *tenir, entrelacer*... l'idée de *complexe*, comme *symplectique*, sous-entend l'existence de plusieurs types d'objets (ici deux) maintenus ensemble dans une même structure. De façon rapide et en anticipant sur la suite, on peut dire que la *complexité* représente la dualité *réel-imaginaire*, et la *symplecticité* la dualité *position-vitesse*. Voilà ce qu'en dit lui même Weyl [Wey, p. 165] :

The name “complex group” formerly advocated by me in allusion to line complexes, as these are defined by the vanishing of antisymmetric bilinear forms, has become more and more embarrassing through collision with the word “complex” in the connotation of complex number. I therefore propose to replace it by the corresponding Greek adjective “symplectic”. Dickson calls the group the “Abelian linear group” in homage to Abel who first studied it.

En ce qui concerne la notion actuelle de *géométrie symplectique*, au sens de la théorie des variétés différentielles munies d'une forme symplectique, il semble que ce soit J.-M. Souriau qui l'ait introduite en 1953 dans son article *Géométrie symplectique différentielle, applications* [Sou1]. Dans un article plus récent *La structure symplectique de la mécanique décrite par Lagrange en 1811* [Sou5] le même auteur décrit un autre aspect des relations entre la géométrie symplectique et la mécanique de Lagrange.

## II.1 Espace des mouvements d'une planète

Pour comprendre et apprécier la méthode de la variation des constantes développée par Lagrange, il est nécessaire de bien connaître la résolution du problème à deux corps. Nous allons en donner un bref résumé dans ce qui suit.

Depuis Newton, on sait que les mouvements d'un point matériel (une planète) autour d'un centre fixe (le Soleil) sont décrits par l'équation différentielle<sup>14</sup> suivante :

$$\frac{d^2\mathbf{r}}{dt^2} = -\frac{\mathbf{r}}{r^3}, \quad (\text{II.1})$$

où  $\mathbf{r}$  désigne un vecteur non nul de l'espace  $\mathbf{R}^3$  et  $r$  son module. Même si cela tient plus à l'histoire des sciences qu'à la science elle-même, il est intéressant de savoir comment Newton est arrivé à cette équation. Le problème posé à l'époque de Newton consistait à prouver que le mouvement des planètes était dû à des forces centripètes qui *décroissent comme le carré de la distance*. A ce moment du développement des sciences, le terme de *force* était pris au sens d'*effort*, comme l'effort que déploie un cheval pour tirer une charrette. Pour résoudre cette question, Newton devait d'abord préciser le sens qu'il donnait au mot *force*. ce qu'il fit sous la forme de principes qu'il énonça d'abord dans son *De Motu* [New1, dGa] :

- Laissées à elles-mêmes les planètes suivraient un mouvement rectiligne uniforme (le principe de l'inertie).
- L'incurvation de leur trajectoire est due à une force extérieure qui les dirige vers le Soleil.
- Pour évaluer la force, il faut mesurer l'incurvation, c'est-à-dire la différence entre la trajectoire rectiligne virtuelle et la trajectoire incurvée réelle.

Ce troisième principe deviendra, dans ses *Principia Mathematica* [New2, New3], la deuxième loi de la mécanique sous la forme suivante :

- Le changement de mouvement est proportionnel à la force motrice imprimée.

Newton admet les trois lois de Kepler :

- Les orbites des planètes sont des ellipses avec le Soleil pour foyer.
- Les aires balayées par le rayon joignant le Soleil à la planète sont proportionnelles au temps de parcours.

---

<sup>14</sup>Il faudrait en toute rigueur multiplier  $\mathbf{r}$  par la constante d'attraction solaire, mais nous choisirons les unités de telle sorte qu'elle soit égale à 1.

- La période du mouvement de la planète autour du Soleil est proportionnelle à la racine carrée du cube du grand axe.

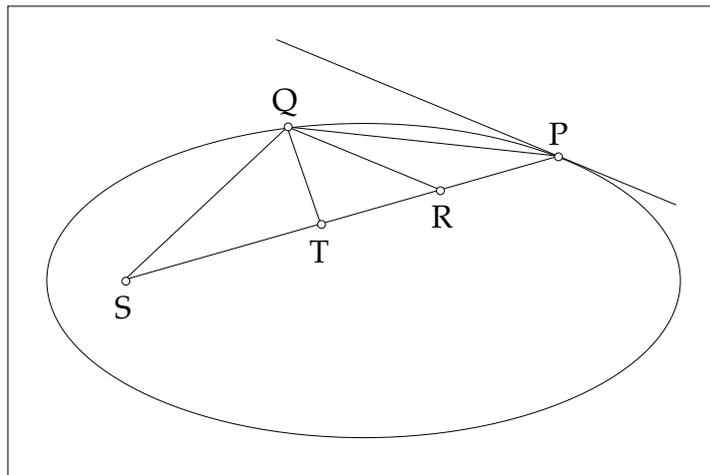


Figure II.2 : Le mouvement elliptique d'une planète

Supposons que la planète se trouve en un point  $P$  de l'ellipse à un instant donné et au point  $Q$  après un temps  $t$ . Si la planète avait été libre de toute interaction elle aurait suivi la tangente à l'ellipse passant par  $P$ , mais puisque sa trajectoire a été incurvée vers le Soleil, cela signifie qu'il lui a imprimé une *force centripète* : dirigée vers le centre. Cette force doit être mesurée par la déflexion de la planète, c'est-à-dire par la distance dont elle est *tombée* sur le Soleil, autrement dit par la longueur du segment  $PR$ , où  $R$  est la projection sur le segment  $SP$ , parallèlement à la tangente en  $P$ , du point  $Q$  (fig. II.2). Newton suppose alors que la planète tombe vraiment sur le Soleil, comme un corps pesant<sup>15</sup>. Il peut donc appliquer la loi de la chute des corps connue depuis Galilée : *la proportionnalité de la hauteur de chute avec le carré du temps de la chute*, c'est-à-dire :

$$PR = \frac{1}{2}gt^2, \quad (\text{II.2})$$

<sup>15</sup>Il n'est peut-être pas inutile d'insister sur cet aspect du raisonnement de Newton, puisque c'est là le caractère *universel* de sa loi de la gravitation : *comme les pommes, les planètes tombent*.

où  $g$  mesure la *pesanteur* au point  $P$ . C'est ce nombre qu'il faut calculer pour connaître l'intensité de la force d'attraction du Soleil. Le temps  $t$  est, d'après la deuxième loi de Kepler, proportionnel à l'aire du secteur balayé. Lorsque le temps  $t$  est petit, cette aire est à peu près égale à celle du triangle  $SPQ$  :

$$t \sim \text{aire}(SPQ) = c \frac{SP \times QT}{2}, \quad (\text{II.3})$$

où  $c$  est une constante de proportionnalité dépendant *a priori* de l'ellipse. La valeur de cette constante  $c$  s'obtient en appliquant la deuxième loi de Kepler à une orbite complète, et en utilisant la troisième loi de Kepler, *i.e.* « la période  $\tau$  de la planète est proportionnelle à la racine carrée du cube du grand axe ». Comme lors d'une révolution complète de la planète l'aire balayée est l'aire de l'ellipse toute entière, nous avons donc les deux égalités suivantes

$$\begin{cases} \tau = c \times \text{aire de l'ellipse} = c \times \pi ab, & (\text{deuxième loi de Kepler}) \\ \text{et} \\ \tau = k\sqrt{a^3} & (\text{troisième loi de Kepler}) \end{cases}$$

où  $a$  est le demi grand axe de l'ellipse et  $b$  l'autre demi grand axe,  $k$  étant une constante dépendant des planètes et des unités de mesure. L'aire des ellipses étant par ailleurs connue par la formule  $\text{aire} = \pi ab$ . Ces deux dernières expressions de  $\tau$  nous permettent d'obtenir une valeur approchée de la constante de gravitation cherchée, acceptable pour les temps petits :

$$g \sim k' \times \frac{b^2}{a} \times \frac{PR}{SP^2 \times QT^2}. \quad (\text{II.4})$$

où  $k'$  est une nouvelle constante intermédiaire. Mais cette approximation ne donne la valeur exacte de  $g$  qu'à la limite où les points  $Q$  et  $P$  se confondent, ce que nous notons  $Q \rightarrow P$ . Auparavant, il est nécessaire d'établir un résultat intermédiaire tiré de la géométrie des ellipses dont nous épargnons le calcul au lecteur :

$$\lim_{Q \rightarrow P} \frac{PR}{QT^2} = \frac{a}{2b^2}, \quad (\text{II.5})$$

qui permet de conclure :

$$g = \frac{K}{SP^2}, \quad (\text{II.6})$$

où  $K$  est une constante ne dépendant que des masses des planètes considérées.

Le coefficient de gravitation  $g$ , l'accélération de la planète, ne dépend donc pas des paramètres de l'ellipse, mais seulement du lieu géométrique de la planète et des masses en présence. Cette propriété remarquable mise en évidence par Newton peut s'exprimer ainsi :

« La force centripète exercée par le Soleil sur la planète est inversement proportionnelle au carré de la distance de la planète au Soleil ».

C'est à partir de ce résultat que Newton a précisé la notion de force comme variation de la quantité de mouvement et qu'il a établi ses équations générales de la mécanique que nous connaissons aujourd'hui<sup>16</sup> sous la forme condensée

$$F = m\gamma.$$

REMARQUE. Certains historiens des sciences (comme François de Gandt par exemple [dGa]) pensent que Newton n'a pas vraiment répondu au problème qui avait été posé à la communauté scientifique de l'époque, en particulier par Sir Christopher Wren, de la *Royal Society*. Ce problème, que l'on appelle *problème direct* peut être énoncé comme ceci : « Montrer que les seules trajectoires possibles d'un corps soumis à une attraction inversement proportionnelle au carré de la distance sont les sections coniques ayant le centre attracteur pour foyer ». Nous avons pu vérifier qu'en utilisant les trois lois de Kepler, Newton déduit effectivement que la planète est attiré par le Soleil selon une loi en inverse du carré de la distance qui les séparent. Ce qui est plutôt l'énoncé réciproque. Mais il ne faut pas oublier que ces mots : « force », « attraction »... n'ont pas de sens à l'époque, à peine un sens commun, vague, mais pas le sens précis, mathématique, opératoire que justement Newton leur a donné. En d'autres termes, c'est Newton lui-même qui pose la vraie

---

<sup>16</sup>Ces équations sont souvent présentées de façon obscure dans les manuels de l'enseignement secondaire. L'élève ne sait pas toujours ce qui est donné : l'expression de la force  $F$  comme une fonction  $F(t, \mathbf{r}, \mathbf{v})$  de certaines variables indépendantes, le temps  $t$  la position  $\mathbf{r}$  et la vitesse  $\mathbf{v}$ , et ce qu'il faut rechercher : la loi horaire, ou plutôt les lois horaires,  $t \mapsto \mathbf{r}(t)$  telles qu'à chaque instant  $t$  l'équation différentielle  $F(t, \mathbf{r}(t), \mathbf{v}(t)) = d^2\mathbf{r}(t)/dt^2$  soit satisfaite. Parfois même l'absence de réflexion sur le choix de la fonction  $F$  obscurcit encore davantage le discours.

question et qui y répond. C'est là sa contribution essentielle à la fondation de la mécanique moderne, d'avoir introduit les notions opératoires de force, d'attraction etc... sans lesquelles les questions évoquées plus haut sont infondées. Toujours d'après F. de Gandt [dGa], il semble que ce soit Jean Bernoulli qui, trente ans plus tard, ait réellement résolu le problème direct. Mais

Ces critiques et ces preuves nouvelles font cependant apparaître le mérite de Newton et l'apport irremplaçable des *Principia*. En effet si Jean Bernoulli corrige et améliore Newton, c'est en s'appuyant sur les résultats et les procédés des *Principia*.

c'est ce qu'écrit justement F. de Gandt dans sa thèse [dGa].

Démontrons le théorème de Newton par les méthodes modernes du calcul linéaire. Auparavant, fixons les quelques notations que nous utiliserons par la suite. Nous noterons par des lettres grasses  $\mathbf{r}, \mathbf{v}, \dots$  les vecteurs, c'est-à-dire les colonnes (ou parfois lignes lorsqu'il s'agit de co-vecteurs) de nombres réels :

$$\mathbf{r} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \dots$$

Les composantes des vecteurs seront notées normalement, par des lettres indexées  $r_i, v_i, \dots$ . Nous utiliserons la même lettre qui désigne un vecteur mais non grasse et sans indice pour désigner sa norme (par exemple  $r$  désigne la norme de  $\mathbf{r}$ ).

Transformons maintenant l'équation différentielle de Newton (II.1) en un système du premier ordre dans  $[\mathbf{R}^3 - \{0\}] \times \mathbf{R}^3$ ; les *mouvements* de la planète deviennent les solutions de

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}, \quad \frac{d\mathbf{v}}{dt} = -\frac{\mathbf{r}}{r^3}, \quad r = \|\mathbf{r}\|. \quad (\text{II.7})$$

Comme on le sait [lire aussi l'annexe B], l'énergie totale du système est conservée le long du mouvement<sup>17</sup>. Les astronomes appellent *constante des forces vives* le

<sup>17</sup>Comme on peut le lire dans l'annexe B, l'énergie totale d'un point matériel dans un champ de force est la somme de son énergie cinétique  $mv^2/2$ , où  $v^2$  est le carré de la norme du vecteur

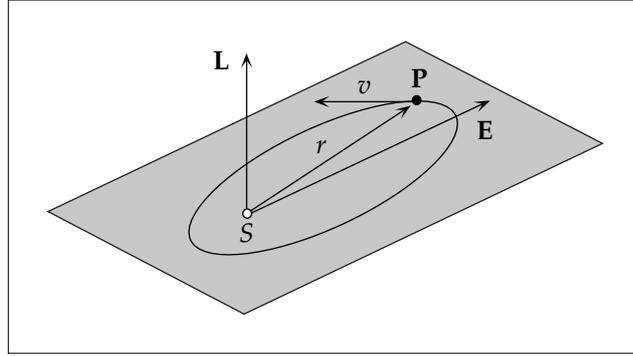


Figure II.3 : L'orbite de la planète P

double de l'énergie que l'on la notera<sup>18</sup>  $f$  :

$$f = v^2 - \frac{2}{r}. \quad (\text{II.8})$$

D'autre part, comme la force d'attraction gravitationnelle est centrale, le moment cinétique  $\mathbf{L}$  est lui aussi conservé<sup>19</sup> :

$$\mathbf{L} = \mathbf{r} \wedge \mathbf{v}. \quad (\text{II.9})$$

De cette invariance on déduit que le mouvement de la planète s'effectue dans le plan orthogonal à  $\mathbf{L}$ .

Un autre vecteur est miraculeusement conservé<sup>20</sup> le long du mouvement,

vitesse  $\mathbf{v}$  et  $m$  la masse du corps, et de l'énergie potentielle  $U$  dont dérive la force  $\mathbf{F}$  exercée sur le corps, c'est-à-dire  $F = (F_i)_{i=1}^3$  et  $F_i = \partial U / \partial r_i$ . L'énergie totale notée en général  $E$  est donc définie par  $E = mv^2/2 + U$ . La vertu essentielle de cette fonction du temps de la position et de la vitesse est d'être préservée le long du mouvement. On dit que c'est une *constante du mouvement*. Un des objectifs de ce livre est de montrer, entre autre, pourquoi la conservation de cette fonction, le long des mouvements, est la conséquence d'une symétrie : le décalage du temps.

<sup>18</sup>Suivant les notations de Lagrange.

<sup>19</sup>Le signe  $\wedge$  représente le produit vectoriel ordinaire de  $\mathbf{R}^3$ ,  $(\mathbf{r} \wedge \mathbf{v})_1 = r_2 v_3 - r_3 v_2 \dots$

<sup>20</sup>Contrairement aux autres constantes du mouvement ce vecteur n'est associé à aucune symétrie évidente (c'est-à-dire spatio-temporelle). Il est lié à une symétrie du groupe  $SO(4)$  qui n'apparaît que lors de la régularisation du problème à deux corps, voir par exemple [Sou3].

on peut le vérifier directement ; c'est *le vecteur de Laplace* :

$$\mathbf{E} = \mathbf{L} \wedge \mathbf{v} + \frac{\mathbf{r}}{r}. \quad (\text{II.10})$$

On déduit de cet invariant supplémentaire les trajectoires des planètes. En effet, on a immédiatement, en prenant la norme carrée du vecteur  $\mathbf{E}$  et le produit scalaire de  $\mathbf{E}$  par  $\mathbf{L}$  :

$$E^2 = 1 + fL^2 \quad \text{et} \quad \mathbf{E} \cdot \mathbf{L} = 0. \quad (\text{II.11})$$

Le vecteur  $\mathbf{E}$  est donc perpendiculaire au moment cinétique à tout instant, et se situe donc dans le plan du mouvement. En prenant le produit scalaire de  $\mathbf{E}$  par  $\mathbf{r}$  on obtient l'équation suivante, vérifiée le long du mouvement :

$$\mathbf{E} \cdot \mathbf{r} + L^2 = r. \quad (\text{II.12})$$

Soit  $\phi$  l'angle entre  $\mathbf{E}$  et  $\mathbf{r}$  (voir figure II.4), de sorte que :

$$\mathbf{E} \cdot \mathbf{r} = Er \cos \phi$$

on obtient alors :

$$Er \cos \phi + L^2 = r \quad \text{ou encore} \quad r = \frac{L^2}{1 - E \cos \phi}. \quad (\text{II.13})$$

On reconnaît ainsi l'équation, exprimée en coordonnées polaires  $(r, \phi)$ , d'une conique de paramètre  $L^2$ , d'excentricité  $E$  et d'axe la direction du vecteur  $\mathbf{E}$ . Les astronomes appellent l'angle  $\phi$  l'*anomalie vraie*<sup>21</sup>. Le vecteur  $\mathbf{E}$  pourrait s'appeler le *vecteur d'excentricité*.

Les trajectoires de la planète sont donc des sections coniques, avec le Soleil pour foyer. Leur nature dépend essentiellement du signe de l'énergie totale, comme le montre la formule<sup>22</sup> (II.11), les trois cas possibles sont :

<sup>21</sup>Dans ce contexte, le terme *anomalie* signifie simplement *paramètre*.

<sup>22</sup>Le lecteur peut vérifier que les trajectoires elliptiques correspondent à une excentricité  $E < 1$ , la norme de  $\mathbf{r}$  étant alors encadrée par  $L^2/(1 + E)$  et  $L^2/(1 - E)$ . Lorsque  $E = 1$ , le rayon  $\mathbf{r}$  est infini pour  $\phi = 0$  ou  $2\pi$ , mais la vitesse du point est alors nulle, comme on peut le vérifier par le calcul : l'orbite parabolique est en quelque sorte *tangente à l'infini*. Dans le troisième cas  $E > 1$ , l'infini est atteint pour  $\phi = \pm \arccos(1/E)$  et la vitesse du point à l'infini n'est pas nulle.

- $f < 0 \Rightarrow E < 1$ , l'excentricité est inférieure à 1 : l'orbite est elliptique.
- $f = 0 \Rightarrow E = 1$ , l'excentricité vaut 1 : l'orbite est parabolique.
- $f > 0 \Rightarrow E > 1$ , l'excentricité est supérieure à 1 : l'orbite est hyperbolique.

Dans le cas des orbites elliptiques, on trouve tout de suite la valeur du demi-grand axe, noté  $a$ . En effet, les points de l'ellipse les plus éloignés du foyer sont atteints pour  $\phi = 0$  et  $\phi = \pi$  ce qui donne les deux valeurs de  $r$  :  $r_0 = L^2/(1 + E)$  et  $r_\pi = L^2/(1 - E)$  ; la somme de ces deux valeurs est égale au grand axe, donc  $2a = L^2/(1 + E) + L^2/(1 - E)$ , d'où on déduit que  $a = L^2/(1 - E^2)$ , ce qui donne grâce à l'équation (II.11) :

$$a = -\frac{1}{f}. \quad (\text{II.14})$$

Nous pouvons décrire complètement la variété des mouvements keplériens elliptiques ( $f < 0$ ) si l'on exclut les chutes sur le centre, c'est-à-dire si on se restreint à  $\mathbf{L} \neq 0$ . Une trajectoire elliptique est bien définie par les deux vecteurs  $\mathbf{L}$  et  $\mathbf{E}$ , le vecteur  $\mathbf{E}$  donnant à la fois l'excentricité et l'axe de la conique, le plan étant défini comme l'orthogonal de  $\mathbf{L}$  et le paramètre de l'ellipse valant  $L^2$ . Autrement dit, l'espace des trajectoires keplériennes elliptiques est équivalent à l'ensemble des couples de vecteurs  $(\mathbf{E}, \mathbf{L}) \in \mathbf{R}^3 \times \mathbf{R}^3$  tels que :

$$E < 1, \quad L \neq 0 \quad \text{et} \quad \mathbf{E} \cdot \mathbf{L} = 0. \quad (\text{II.15})$$

C'est une sous-variété, de dimension 5, de  $\mathbf{R}^3 \times \mathbf{R}^3$ . Ce n'est pas encore l'espace des mouvements keplériens elliptiques : il nous faut pouvoir calculer la position de la planète à chaque instant. On pourrait, pour cela, choisir la position de la planète sur son orbite (c'est-à-dire l'anomalie vraie) à l'instant zéro. Mais ce choix donne lieu à des calculs pénibles qui ne seront pas faits ici. On considère plutôt le vecteur qui joint l'origine du cercle circonscrit à l'ellipse au point  $\mathbf{A}$  de ce cercle qui a la même projection orthogonale, sur l'axe dirigé par  $\mathbf{E}$ , que la planète  $P$  (voir figure II.4). Nous obtenons ainsi une description géométrique de l'espace des mouvements keplériens (ou plutôt d'un ouvert de l'espace des mouvements

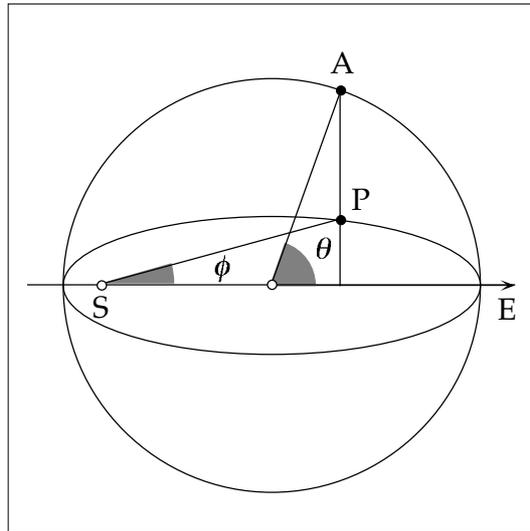


Figure II.4 : L'anomalie excentrique

keplériens, puisque nous avons supprimé l'ensemble des chutes sur le foyer<sup>23</sup> *i.e.* les mouvements correspondants aux conditions initiales  $\mathbf{L} = \mathbf{r} \wedge \mathbf{v} = 0$ ). Comme nous pouvons le constater directement, nous retrouvons les 6 paramètres nécessaires, ainsi que cela était précédemment indiqué.

Ce vecteur, ou plus précisément l'angle  $\theta$  qu'il fait avec l'axe de l'ellipse, est appelé *anomalie excentrique*<sup>24</sup>, et a été introduit par Kepler. L'intérêt de l'anomalie excentrique est l'équation différentielle suivante (*l'Équation de Kepler*), que ce paramètre vérifie le long du mouvement :

$$dt = \sqrt{a^3} [1 - E \cos \theta] d\theta, \quad (\text{II.16})$$

<sup>23</sup>Le problème des chutes sur le foyer donne lieu à la *régularisation* du problème de Kepler qui est un autre sujet très intéressant lié au problème à deux corps et qui sera peut-être inclus dans une version future de ce livre.

<sup>24</sup>Comme le montre la figure l'*anomalie excentrique* doit son nom à ce qu'il est le paramètre excentré de l'ellipse, le « vrai » centre, le centre d'attraction, étant bien entendu le foyer.

et qui donne, par intégration, une nouvelle constante du mouvement :

$$c = t - \sqrt{a^3} [\theta - E \sin \theta]. \quad (\text{II.17})$$

Cette constante est donc égale à la valeur de  $t$  pour  $\theta = 0$ , c'est-à-dire à la date du passage de la planète à l'aphélie. C'est ce paramètre que les astronomes appellent l'*époque* de la planète, et qu'ils choisissent à la place de l'anomalie excentrique à l'instant zéro<sup>25</sup>.

DÉMONSTRATION. (ÉQUATION DE KEPLER) *Plaçons nous dans le repère orthonormé d'origine le centre de symétrie de l'ellipse, et d'ordonnée l'axe du vecteur excentricité  $\mathbf{E}$ . L'équation d'un point courant de l'ellipse et sa vitesse sont données, en coordonnées polaires, par :*

$$P = a \begin{pmatrix} \cos \theta \\ \sqrt{1-E^2} \sin \theta \end{pmatrix} \quad \mathbf{v} = a \begin{pmatrix} -\sin \theta \\ \sqrt{1-E^2} \cos \theta \end{pmatrix} \frac{d\theta}{dt}. \quad (\text{II.18})$$

où  $a$  est le demi-grand axe et  $E < 1$  l'excentricité (la norme du vecteur  $\mathbf{E}$ ). Le carré de la vitesse vaut donc

$$v^2 = \|\mathbf{v}\|^2 = a^2(1 - E^2 \cos^2 \theta) \left( \frac{d\theta}{dt} \right)^2.$$

En comparant cette expression avec la constante des forces vives donnée par la formule (II.8) :  $v^2 = f + 2/r$ , en se rappelant que  $f = -1/a$ , et en utilisant la formule suivante donnant la distance  $r$  du foyer  $S$  au point courant  $P$

$$r = a\sqrt{1 + E \cos \theta},$$

on obtient l'équation de Kepler annoncée. ■

REMARQUE. Les mouvements keplériens sont donc définis par les valeurs de l'époque, du moment cinétique et du vecteur de Laplace. Mais il est évident,

<sup>25</sup>En réalité ce paramètre est mal défini puisque le mouvement de la planète est périodique. Il n'est vraiment défini que modulo  $\sqrt{a^3}$  (la période du mouvement). Il faudrait plutôt choisir  $C = \exp(2i\pi c/\sqrt{a^3})$ , ce qui est équivalent au choix de  $\mathbf{A}$  à l'instant zéro.

puisque tous les mouvements elliptiques sont périodiques, que cet espace des mouvements kepleriens est aussi l'ensemble des conditions initiales à l'instant  $t = 0$ , c'est-à-dire l'ouvert de  $\mathbf{R}^3 \times \mathbf{R}^3$  des couples  $(\mathbf{r}, \mathbf{v})$  vérifiant :

$$\mathbf{r} \wedge \mathbf{v} \neq 0 \quad \text{et} \quad v^2 - \frac{2}{r} < 0. \quad (\text{II.19})$$

La représentation d'un mouvement keplerien par ses conditions initiales ou par ses caractéristiques géométriques est *a priori* affaire de goût. Nous verrons toutefois que certaines représentations sont plus pratiques que d'autres. Dans sa *Mécanique*, Lagrange choisit les six *éléments kepleriens*  $(a, b, c, h, i, k)$ , où  $a$  est la valeur du demi-grand axe (l'inverse de la constante des forces vives au signe près),  $b$  le paramètre de l'ellipse (le carré du moment cinétique),  $c$  l'époque. Les éléments  $h$ ,  $i$  et  $k$  déterminent le plan de l'orbite et l'axe de l'ellipse dans ce plan :  $i$  est l'inclinaison du plan de l'orbite par rapport à un plan de référence,  $h$  est la *longitude des nœuds*, c'est-à-dire l'angle que fait la trace du plan de l'orbite sur le plan de référence (la *ligne des nœuds*), et  $k$  est la *longitude du périhélie*, c'est-à-dire l'angle que fait l'axe de l'ellipse avec la ligne des nœuds.

## II.2 La méthode de la variation des constantes

Maintenant que nous avons bien compris et résolu<sup>26</sup> le problème à deux corps (au moins en ce qui concerne les mouvements réguliers), il nous reste à traiter le problème à deux corps perturbé, et introduire ainsi les premiers calculs symplectiques comme l'a fait Lagrange. Nous nous bornerons, comme lui<sup>27</sup>, aux perturbations des orbites elliptiques ; les astronomes se sont naturellement toujours intéressés davantage aux perturbations des mouvements elliptiques puisqu'ils concernent, en particulier, les mouvements de notre planète. Évidemment, cela exclut les fortes perturbations qui auraient pour conséquence de transformer notre orbite elliptique en une orbite hyperbolique, cas pour lequel d'autres techniques sont à mettre en œuvre.

<sup>26</sup>C'est-à-dire que nous avons ramené le problème à un exercice de calcul numérique.

<sup>27</sup>Même s'il évoque explicitement la possibilité d'une modification de l'orbite de l'ellipse à la parabole, ou à l'hyperbole.

Nous avons déjà expliqué (dans le deuxième chapitre) la méthode de la variation des constantes (voir page 12). Rappelons succinctement de quoi il s'agit. Mais plutôt que de répéter ce que nous avons déjà écrit, laissons la plume à Lagrange [Lag8, tome II, p.58] :

« Un des premiers et des plus beaux résultats de la Théorie de Newton, sur le système du monde, consiste en ce que toutes les orbites des corps célestes sont de même nature, et ne diffèrent entre elles qu'à raison de la force de projection que ces corps peuvent être supposés avoir reçue dans l'origine des choses. Il suit de là que, si une planète ou une comète venait à recevoir une impulsion étrangère quelconque, son orbite en serait dérangée ; mais il n'y aurait que les éléments, qui sont les constantes arbitraires de l'équation, qui pourraient changer : c'est ainsi que l'orbite circulaire ou elliptique d'une planète pourrait devenir parabolique ou même hyperbolique, ce qui transformerait la planète en comète.

Il en est de même de tous les problèmes de Mécanique. Comme les constantes arbitraires introduites par les intégrations dépendent de l'état initial du système, qui peut être placé dans un instant quelconque, si l'on suppose que les corps viennent à recevoir pendant leur mouvement des impulsions quelconques, les vitesses produites par ces impulsions étant composées avec les vitesses déjà acquises par les corps, pourront être regardées comme des vitesses initiales, et ne feront que changer les valeurs des constantes.

Et si au lieu d'impulsions finies, qui n'agissent que dans un instant, on suppose les impulsions infiniment petites, mais dont l'action soit continue, les mêmes constantes deviendront tout à fait variables, et serviront à déterminer l'effet de ces sortes de forces, qu'il faudra regarder comme des forces perturbatrices. »

Ainsi, étant donnée une condition initiale du système, l'influence des forces perturbatrices se traduit sur l'espace des mouvements — l'espace des constantes d'intégration — par une courbe. Cette courbe, tracée pour le cas d'une planète

dans l'espace des mouvements keplériens<sup>28</sup> (non perturbés), exprime le mouvement réel de la planète. Il s'agit donc d'en exprimer l'équation, et éventuellement d'en extraire quelques renseignements, comme par exemple la stabilité du grand axe. Ce résultat a été découvert par Laplace en 1773<sup>29</sup>. Plutôt que de présenter cette méthode dans le seul cas des planètes, nous allons montrer maintenant comment Lagrange l'a inclus dans le cadre général de sa méthode de la variation des constantes pour tous les systèmes de la mécanique. Introduisons les quelques notations utilisées par la suite.

NOTATIONS. De façon générale, on note

$$D(x \mapsto y)(x) = \frac{\partial y}{\partial x}$$

l'application linéaire tangente d'une application différentiable  $x \mapsto y$ . Lorsque le but est  $\mathbf{R}$  et la source  $\mathbf{R}^3$  comme c'est le cas ici pour  $[\mathbf{v} \mapsto a]$ , l'application linéaire tangente est un covecteur (une ligne de nombres) :

$$\frac{\partial a}{\partial \mathbf{v}} = \left[ \frac{\partial a}{\partial v_1} \quad \frac{\partial a}{\partial v_2} \quad \frac{\partial a}{\partial v_3} \right].$$

Lorsqu'il s'agit d'une application de  $\mathbf{R}$  dans  $\mathbf{R}^3$ , comme c'est le cas ici pour  $[t \mapsto \mathbf{v}]$  l'application linéaire est un vecteur (une colonne de nombres), de telle sorte que le produit de l'un par l'autre soit donné par la somme :

$$\frac{\partial a}{\partial \mathbf{v}} \cdot \frac{d\mathbf{v}}{dt} = \sum_{i=1}^3 \frac{\partial a}{\partial v_i} \frac{dv_i}{dt}.$$

Ces notations seront utilisées tout le long de ce livre. Nous omettrons parfois le point de multiplication.

<sup>28</sup>Les points de cette courbe sont des mouvements keplériens dont la trajectoire est elliptique, d'après nos hypothèses. Pour des raisons bien compréhensibles, ces ellipses sont appelées par Lagrange les *ellipses osculatrices* au mouvement réel.

<sup>29</sup>Ce théorème de Laplace dit que, sous certaines conditions sur les masses en présence, si les perturbations de la planète sont suffisamment petites, le grand axe de la famille d'ellipses osculatrices au mouvement perturbé ne contient pas de termes linéaires en temps mais uniquement des termes périodiques. Autrement dit, faiblement perturbé, le grand axe des ellipses osculatrices est borné dans le temps, ce qui est plutôt rassurant.

Revenons maintenant au système des planètes. Supposons donc qu'une planète subisse de façon continue une série de chocs infiniment petits. Ces chocs se traduisent par une variation instantanée de la vitesse, sans conséquence sur sa position. Si on désigne par  $a$  un élément quelconque de la planète (pas nécessairement le demi grand axe), le calcul différentiel donne :

$$\frac{da}{dt} = \frac{\partial a}{\partial \mathbf{v}} \cdot \frac{d\mathbf{v}}{dt}. \quad (\text{II.20})$$

En remarquant que le vecteur  $d\mathbf{v}/dt$  représente exactement la force perturbatrice  $X$  exercée sur la planète à l'instant  $t$  au point (vecteur)  $\mathbf{r}$ , la variation infinitésimale de l'élément  $a$ , sous l'effet de la perturbation, peut s'écrire :

$$\frac{da}{dt} = \frac{\partial a}{\partial \mathbf{v}} \cdot X. \quad (\text{II.21})$$

Le mouvement vrai est ainsi décrit par la courbe intégrale de cette équation, tracée dans l'espace des éléments de la planète. Cette famille d'ellipses est appelée *famille d'ellipses osculatrices* du mouvement perturbé.

Supposons maintenant que la force perturbatrice  $X$  dérive d'un potentiel  $\Omega$ , c'est-à-dire que, par définition<sup>30</sup> :

$$X = \frac{\partial \Omega}{\partial \mathbf{r}} \quad (\text{II.22})$$

Supposons de plus, comme Lagrange, que ce *potentiel de perturbation*  $\Omega$  n'est fonction que de  $\mathbf{r}$ . Ce qui, dit autrement, s'écrit :

$$\frac{\partial \Omega}{\partial \mathbf{v}} = 0. \quad (\text{II.23})$$

À ce stade, Lagrange emploie l'astuce suivante : ajouter  $0 = \partial \Omega / \partial \mathbf{v}$  à  $da/dt$ , pour éliminer dès le départ des termes symétriques ennuyeux et pour faire apparaître

---

<sup>30</sup>C'est vraisemblablement Lagrange qui a introduit cette propriété/définition des forces dérivant de potentiel. En termes mathématiques, sans connotation physique nous disons que la force  $X$  est la *différentielle* (ou la *dérivée extérieure*) d'une fonction scalaire  $\Omega$ . Cela pose un certain nombre de questions sur la nature géométrique de cet objet qu'est la force, mais nous n'aurons pas ce débat ici. Il est plus ou moins stérile et largement dépassé par le cadre moderne de la mécanique symplectique. Il suffit juste, pour ce qui nous concerne ici, de considérer les choses comme Lagrange lui-même les a considérées : simplement.

tout de suite la structure antisymétrique qui s'imposait par le calcul. Il aura fallu à Lagrange trois articles sur le même sujet (voir appendice A) pour arriver à ce raccourci et cette présentation épurée. Nous avons donc :

$$\begin{aligned}
\frac{da}{dt} &= \frac{\partial a}{\partial \mathbf{v}} \cdot X \\
&= \frac{\partial a}{\partial \mathbf{v}} \cdot \frac{\partial \Omega}{\partial \mathbf{r}} \\
&= \frac{\partial a}{\partial \mathbf{v}} \cdot \frac{\partial \Omega}{\partial \mathbf{r}} - \frac{\partial a}{\partial \mathbf{r}} \cdot \frac{\partial \Omega}{\partial \mathbf{v}} \\
&= \sum_{i=1}^3 \frac{\partial a}{\partial \mathbf{v}^i} \frac{\partial \Omega}{\partial \mathbf{r}^i} - \frac{\partial a}{\partial \mathbf{r}^i} \frac{\partial \Omega}{\partial \mathbf{v}^i}.
\end{aligned} \tag{II.24}$$

C'est maintenant, avec cette transformation astucieuse de Lagrange, que commence la véritable histoire, d'où sortira la géométrie symplectique. Mais il faut aller plus loin : puisque l'application  $(t, \mathbf{r}, \mathbf{v}) \mapsto (t, a, b, c, h, i, k)$  est un difféomorphisme, le potentiel de perturbation peut être considéré aussi bien comme une fonction de  $\mathbf{r}$  que comme une fonction du temps  $t$  et des éléments  $(a, b, c, h, i, k)$  de la planète. En utilisant les méthodes du calcul différentiel, que maîtrisait parfaitement Lagrange, on remplace l'expression de  $\partial\Omega/\partial\mathbf{r}$  :

$$\frac{\partial \Omega}{\partial \mathbf{r}} = \frac{\partial \Omega}{\partial a} \frac{\partial a}{\partial \mathbf{r}} + \frac{\partial \Omega}{\partial b} \frac{\partial b}{\partial \mathbf{r}} + \text{etc.}, \tag{II.25}$$

et de  $\partial\Omega/\partial\mathbf{v}$  :

$$\frac{\partial \Omega}{\partial \mathbf{v}} = \frac{\partial \Omega}{\partial a} \frac{\partial a}{\partial \mathbf{v}} + \frac{\partial \Omega}{\partial b} \frac{\partial b}{\partial \mathbf{v}} + \text{etc.}, \tag{II.26}$$

dans l'équation (II.24), pour obtenir une nouvelle expression de  $da/dt$  :

$$\frac{da}{dt} = (a, b) \frac{\partial \Omega}{\partial b} + (a, c) \frac{\partial \Omega}{\partial c} + \text{etc.} \tag{II.27}$$

où les parenthèses  $(a, b)$ ,  $(a, c)$ ,  $\dots$ , désignent les fonctions de  $(t, \mathbf{r}, \mathbf{v})$  définies par :

$$(a, b) = \sum_{i=1}^3 \frac{\partial a}{\partial \mathbf{v}^i} \frac{\partial b}{\partial \mathbf{r}^i} - \frac{\partial b}{\partial \mathbf{v}^i} \frac{\partial a}{\partial \mathbf{r}^i}. \tag{II.28}$$

Il en est de même pour les quatorze autres parenthèses<sup>31</sup>. Les termes  $\partial\Omega/\partial a$ ,  $\partial\Omega/\partial b$ ... intervenant dans cette formule peuvent être considérés comme les *forces de perturbations* rapportées aux variables  $(a, b, c, h, i, k)$ . Les coefficients des forces de perturbation exprimées dans les variables  $(a, b, c, h, i, k)$  sont appelés aujourd'hui *parenthèses de Lagrange*<sup>32</sup>.

L'expression formelle (II.21) de la variation  $da/dt$  est beaucoup plus simple que celle (II.27) à laquelle nous avons abouti après toutes ces transformations. On est en droit de se demander quel intérêt nous avons eu à effectuer ces transformations. La réponse est contenue dans le théorème suivant de Lagrange, où l'on considère le difféomorphisme  $(t, \mathbf{r}, \mathbf{v}) \mapsto (t, a, b, c, h, i, k)$ .

THÉORÈME. (LAGRANGE) *Les parenthèses  $(a, b)$ ,  $(a, c)$ , etc. considérées comme des fonctions de  $(t, a, b, c, h, i, k)$  ne sont fonction que des éléments  $(a, b, c, h, i, k)$ .*

À ce propos Lagrange écrit exactement [Lag8, tome II, p. 73] :

« Ainsi la variation de  $a$  sera représentée par une formule qui ne contiendra que les différences partielles de  $\Omega$  par rapport à  $b, c$ , etc., multipliées chacune par une fonction de  $a, b, c$ , etc., sans  $t$ . Et la même chose aura lieu à l'égard des variations des autres constantes arbitraires  $b, c, h$ , etc. »

NOTE. Lagrange a donné successivement plusieurs démonstrations de ce théorème, le généralisant et le simplifiant chaque fois davantage. Il l'énonce la première fois, dans le cadre du mouvement des planètes, dans son mémoire de 1808 [Lag5]. Il le généralise ensuite à tous les problèmes de la mécanique dans son mémoire de 1809 [Lag6], qu'il complète par l'inversion explicite de ses parenthèses dans son mémoire de 1810 [Lag7] (voir annexe A pour plus de détails). Dans son mémoire de 1809, Lagrange donne en annexe une construction épurée, simplifiée qui ne comporte plus que quelques pages. L'énoncé particulier que nous avons donné

<sup>31</sup>Pour 6 constantes du mouvement, il y a  $6 \times 6 = 36$  valeurs de parenthèses possibles ; par antisymétrie, les six parenthèses diagonales  $(x, x)$ , où  $x = a, \dots, k$ , sont nulles et les 30 autres s'organisent deux par deux  $(x, y) = -(y, x)$ , ce qui donne seulement 15 valeurs différentes possibles non nulles.

<sup>32</sup>Ce sont aussi les *crochets de Poisson* des fonctions coordonnées  $a, \dots, k$ .

plus haut est extrait de sa *Mécanique Analytique* publiée en 1811 [Lag8]. On peut voir dans ce théorème les prémisses du théorème connu aujourd'hui des étudiants sous la forme suivante : *le crochet de Poisson de deux constantes du mouvement est encore une constante du mouvement...*

Aussitôt énoncé son théorème, Lagrange fait remarquer<sup>33</sup> que la formule (II.27) donnant l'expression de la variation des éléments de la planète en fonction des forces de perturbations s'inverse, et note que :

$$\frac{\partial \Omega}{\partial a} = [a, b] \frac{db}{dt} + [a, c] \frac{dc}{dt} + \text{etc.}, \quad (\text{II.29})$$

où les crochets  $[a, b]$ ,  $[a, c]$ ,  $\dots$ , ne sont eux-mêmes fonction que des éléments  $(a, b, c, h, i, k)$ , et sont explicitement donnés par :

$$[a, b] = \sum_{i=1}^3 \frac{\partial \mathbf{r}^i}{\partial a} \frac{\partial \mathbf{v}^i}{\partial b} - \frac{\partial \mathbf{v}^i}{\partial a} \frac{\partial \mathbf{r}^i}{\partial b}, \quad \text{etc.} \quad (\text{II.30})$$

Dans cette dernière équation les vecteurs  $\mathbf{r}$  et  $\mathbf{v}$  sont considérés comme fonctions de  $t$  et des éléments  $(a, b, c, h, i, k)$ .

Ainsi le mouvement de la planète perturbée est décrit par l'équation différentielle (II.27) sur l'espace des mouvements de la planète non perturbée, ou si l'on préfère sur l'*espace des constantes d'intégration* du système non perturbé. C'est évidemment là l'origine du nom donné par Lagrange à sa méthode : *la méthode de la variation des constantes*. En effet, la variation des constantes d'intégration du système non perturbé décrit le mouvement réel du système perturbé.

NOTE. Cette méthode est évidemment de même nature que la méthode du même nom développée par Lagrange entre 1774 et 1779, à la fois pour comprendre la nature des *solutions particulières des équations différentielles* [Lag1, Lag3] que pour résoudre les systèmes différentiels linéaires inhomogènes [Lag2, Remarque 5,

<sup>33</sup>Cette présentation respecte la *Mécanique Analytique* de Lagrange même si historiquement Lagrange a d'abord défini ses crochets dans l'article de 1808, où ils apparaissent comme les coordonnées des forces  $\partial \Omega / \partial a$  en fonction des variations des éléments. Ils sont notés, dans cet article, comme des parenthèses.

pp. 159–165]. C’est dans ce dernier mémoire *Sur les suites récurrentes...*<sup>34</sup> que Lagrange expose de façon formelle sa méthode sur la variation des constantes, méthode qu’il n’avait fait qu’ébaucher dans [Lag1], mais qu’il avait déjà abondamment utilisée.

Dans le cas des équations linéaires inhomogènes [Lag2], la partie non homogène est traitée comme une perturbation de la partie linéaire. L’espace des solutions du système linéaire est un espace vectoriel dont chaque point est un ensemble de constantes d’intégration. Le terme non linéaire du système initial définit sur cet espace vectoriel un nouveau système différentiel, équivalent au premier, mais qui porte sur les *constantes d’intégration* du système linéaire. Il est intéressant de noter à ce propos cette remarque de Lagrange [Lag2, p. 163] :

« J’avoue que l’intégration des équations en  $a, b, c, \dots$  et  $x$  sera le plus souvent très difficile, du moins aussi difficile que celle de l’équation proposée [...] mais le grand usage de la méthode précédente est pour intégrer par approximation les équations dont on connaît déjà l’intégrale complète à peu près, c’est-à-dire en négligeant les quantités qu’on regarde comme très petites. »

Il achève sa remarque 5 de son article *Sur les suites récurrentes...* par ce paragraphe prémonitoire, treize ans avant son premier mémoire sur la variation des constantes appliquée au système des planètes :

« Il est visible au reste que cette méthode, que je ne fais qu’exposer ici en passant, peut s’appliquer également au cas où l’on aurait plusieurs équations différentielles entre plusieurs variables dont on connaîtrait les intégrales complètes approchées, c’est-à-dire en y négligeant des quantités supposées très petites. Elle sera par conséquent fort utile pour calculer les mouvements des planètes en tant qu’ils sont altérés par leur action mutuelle, puisqu’en faisant abstraction de cette action la solution complète du problème est connue ; et il est bon de remarquer que, comme dans ce cas les constantes  $a, b, c, \dots$  représentent

---

<sup>34</sup>Ce mémoire n’a que peu à voir avec la méthode de la variation des constantes. Lagrange le dit lui-même : « Quoique ce ne soit pas ici le lieu de nous occuper de cette matière, je vais néanmoins en traiter en peu de mots, me réservant de le faire ailleurs avec plus d’étendue. »

ce qu'on nomme les *éléments des planètes*, notre méthode donnera immédiatement les variations de ces éléments provenant de l'action que les planètes exercent les unes sur les autres. »

On peut se demander quelle est alors la différence entre cette méthode, introduite dans les années 1770, et son application au cas du système des planètes ? Elle relève principalement du type de système traité. En appliquant sa méthode générale de la variation des constantes aux systèmes différentiels spécifiques de la mécanique, Lagrange fait apparaître une structure particulière, qui n'existe pas dans le cas général et qui est à l'origine de la géométrie symplectique. Cette structure, caractérisée par les crochets et parenthèses qu'il a défini, Lagrange va savoir en tirer profit, comme il l'espérait, dans l'étude de la stabilité du grand axe des planètes, comme nous allons le voir maintenant.

### II.3 Application à la stabilité séculaire du grand axe

Nous sommes en mesure maintenant de déduire, de toutes ces transformations et manipulations algébriques, le théorème de Lagrange sur la stabilité du grand axe des planètes. Appliquons la formule (II.29) à l'époque  $c$  :

$$\frac{\partial \Omega}{\partial c} = [c, a] \frac{da}{dt} + [c, b] \frac{db}{dt} + \dots + [c, k] \frac{dk}{dt}. \quad (\text{II.31})$$

Lagrange a calculé l'ensemble de ces crochets. Les calculs sont pénibles et parfois même douteux, mais on peut le croire lorsqu'il affirme que les crochets  $[c, b]$ ,  $[c, h]$ ,  $[c, i]$ ,  $[c, k]$  sont nuls ; il reste alors :

$$[c, a] = -1/2a^2 \quad \text{d'où} \quad \frac{\partial \Omega}{\partial c} = -\frac{1}{2a^2} \frac{da}{dt}. \quad (\text{II.32})$$

Si on se rappelle alors que le demi-grand axe  $a$  est égal à  $-1/f$ , où la constante des forces vives  $f$  est le double de l'énergie  $H$ <sup>35</sup> du mouvement keplerien, on obtient :

$$\frac{dH}{dt} = -\frac{\partial \Omega}{\partial c}. \quad (\text{II.33})$$

---

<sup>35</sup>La lettre  $H$  a été choisie par Lagrange en l'honneur de Huygens et non de Hamilton, voir [Lag8, tome I, pages 217–226 et 267–270] ; voir également l'annexe A.

Cette formule est en réalité très générale et Lagrange l'établit pour tous les problèmes de mécanique analytique conservatifs [Lag6].

Comme nous l'avons déjà dit, le potentiel de perturbation  $\Omega$  (fonction de  $\mathbf{r}$ ) est considéré comme fonction de  $t$  et des éléments keplériens  $(a, b, c, h, i, k)$ . Mais le temps n'intervient dans  $\Omega$  que par  $t - c$ ; plus précisément  $\Omega$  n'est fonction que de  $(a, b, t - c, h, i, k)$ . En effet, dans des coordonnées orthonormées du plan de l'orbite, en prenant pour axe des  $x$  l'axe porté par le vecteur  $\mathbf{E}$  et en posant  $\mathbf{r} = (x, y)$  (voir fig. II.4), on a :

$$x = a\sqrt{1 - \frac{b}{a}} + a \cos(\theta) \quad \text{et} \quad y = \sqrt{ab} \sin(\theta), \quad (\text{II.34})$$

où l'anomalie excentrique  $\theta$  est donnée en inversant la formule (II.17) de Kepler. On peut préciser davantage les choses en notant  $\phi_E$  la fonction :

$$\phi_E : \theta \mapsto \theta - E \sin(\theta) \quad \text{avec} \quad E = \sqrt{1 - \frac{b}{a}}. \quad (\text{II.35})$$

Cette fonction est inversible (car  $E < 1$ ) et on peut écrire :

$$x = a\sqrt{1 - \frac{b}{a}} + a \cos \left[ \phi_E^{-1} \left( \frac{t - c}{a^{3/2}} \right) \right] \quad (\text{II.36})$$

et

$$y = \sqrt{ab} \sin \left[ \phi_E^{-1} \left( \frac{t - c}{a^{3/2}} \right) \right]. \quad (\text{II.37})$$

On en déduit, d'une part, une nouvelle expression de (II.33) donnant la variation de l'énergie  $H$  :

$$\frac{dH}{dt} = \frac{\partial \Omega}{\partial t}. \quad (\text{II.38})$$

On constate, d'autre part, que la fonction  $\Omega$  est périodique en  $t - c$  (voir (II.36) et (II.37)), de période  $2\pi a^{3/2}$ . Le potentiel peut alors se développer en série trigonométrique, qu'on appelle aujourd'hui *série de Fourier*. Il est intéressant de noter ce que Lagrange écrit explicitement<sup>36</sup> à ce propos [Lag5, pages 735–736] :

---

<sup>36</sup>Lagrange et Fourier étaient contemporains, quelle a été leur influence mutuelle ?

« Comme les valeurs des coordonnées peuvent être réduites en série de sinus et cosinus, il est facile de voir que la fonction  $\Omega$  pourra être réduite en une série de sinus et cosinus ; ces sinus et cosinus ayant pour coefficients des fonctions des éléments  $a, b, c$ , etc. »

Cette série de Fourier s'écrit aujourd'hui de la manière suivante :

$$\Omega = \sum_k A_k \exp \frac{ik(t-c)}{a^{3/2}}. \quad (\text{II.39})$$

Les coefficients  $A_k$  étant des fonctions uniquement des éléments de l'orbite  $a, b, h, i, k$ , l'équation (II.38) devient alors :

$$\frac{dH}{dt} = -\frac{\partial \Omega}{\partial c} = \sum_{k \neq 0} \frac{ikA_k}{a^{3/2}} \exp \frac{ik(t-c)}{a^{3/2}}. \quad (\text{II.40})$$

Ainsi que l'énonce Lagrange, la première approximation consiste à regarder dans la fonction  $\Omega$  tous ces éléments comme constants [Lag5, page 736] — *i.e.* à considérer, à l'intérieur des fonctions  $A_k$ , les éléments de l'orbite comme constants. Sans vouloir commenter la validité de cette affirmation, on obtient ensuite par intégration :

$$H(t) \sim H_0 + \sum_{k \neq 0} A_k \exp \frac{ik(t-c)}{a^{3/2}}. \quad (\text{II.41})$$

À ce premier ordre d'approximation, la fonction  $H$  (et donc le grand axe  $a = -1/2H$ ) ne contient pas de terme linéaire en  $t$  (qu'on appelle le *terme séculaire*<sup>37</sup>) mais seulement des termes périodiques. Nous venons de re-démontrer le théorème de stabilité du grand axe de Lagrange. Laissons le encore une fois s'exprimer [Lag5, page 736] :

THÉORÈME. (LAGRANGE) *Les grands axes des planètes ne peuvent être sujets qu'à des variations périodiques, et non à des variations croissant comme le temps.*

Ce théorème n'est qu'une application particulière des méthodes de la variation des constantes introduite par Lagrange. Elle ne concerne, telle qu'elle est

---

<sup>37</sup>Car leur présence entraîne des perturbations sensibles au long des siècles.

présentée ici, que la première approximation (démontrée la première fois, mais par d'autres méthodes, par Laplace en 1773). Son véritable théorème sur la stabilité séculaire des grands axes des planètes (où il étend véritablement le résultat de Laplace) est plus profond, subtil et délicat car il prend en compte le mouvement de toutes les planètes (consulter par exemple [Ste]). Nous ne le présenterons pas ici, car le calcul est plus compliqué même si le principe est le même : écrire le mouvement perturbé dans l'espace des solutions des mouvements non perturbés. Malgré tout, même si la méthode de Lagrange est correcte (nous en verrons plus loin une justification moderne), ses approximations sont douteuses et méritent davantage de soin. C'est ce qu'apporteront par la suite des mathématiciens comme Poincaré et ses successeurs.

L'importance de cette nouvelle méthode introduite par Lagrange, outre qu'elle formule de façon élégante les principes de la mécanique analytique — en introduisant la structure symplectique de l'espace des mouvements kepleriens — facilite aussi le calcul des autres *inégalités*<sup>38</sup>. C'est ce qui la rendu célèbre puisque Lagrange montre que la variation de l'angle du périhélie de Jupiter, observée par les astronomes (mais non encore expliquée à l'époque), est périodique. Il en calculera la période ( $\sim 900$  ans si on croit Sternberg [Ste]).

La résolution des inégalités du mouvement des planètes et le calcul des perturbations marque l'origine du calcul symplectique en mécanique. L'histoire a montré qu'il en est le cadre idéal.

## II.4 Structure de l'espace des mouvements kepleriens

Ces crochets  $[a, b]$ ,  $[a, c], \dots$ , fonctions seulement des éléments kepleriens  $a, b, c$  etc. possèdent trois propriétés remarquables qui sont à la base de l'axiomatique de la géométrie symplectique.

**1. Antisymétrie.** Ils sont *anti-symétriques* :

$$[a, b] = -[b, a], \quad [a, c] = -[c, a], \quad \text{etc.} \quad (\text{II.42})$$

---

<sup>38</sup>C'est ainsi qu'on appelait les variations des éléments de l'orbite dues aux perturbations extérieures.

**2. Inversibilité.** La matrice  $\omega$  définie par la famille de crochets :

$$\omega_{ab} = [a, b], \quad \omega_{ac} = [a, c], \quad \text{etc.} \quad (\text{II.43})$$

est inversible, et son inverse est la matrice des parenthèses de Lagrange :

$$(\omega^{-1})_{ab} = (a, b), \quad (\omega^{-1})_{ac} = (a, c), \quad \text{etc.} \quad (\text{II.44})$$

**3. Fermeture.** Pour tous les triplets d'éléments  $(a, b, c)$ ,  $(a, b, h)$ ,  $\dots$ ,  $(i, h, k)$  l'équation aux dérivées partielles suivante est vérifiée :

$$\frac{\partial[b, c]}{\partial a} + \frac{\partial[c, a]}{\partial b} + \frac{\partial[a, b]}{\partial c} = 0, \quad \text{etc.} \quad (\text{II.45})$$

Ces trois propriétés font de la matrice  $\omega$  ce qu'on appelle aujourd'hui une *forme symplectique*.

Sans vouloir nous attarder sur les définitions formelles, disons seulement qu'une *forme différentielle* définie sur un ouvert d'un espace numérique est une application qui à chaque point de cet ouvert associe une application multilinéaire alternée. Par exemple, une 2-forme  $\omega$  définie sur un ouvert de  $\mathbf{R}^{2n}$  sera caractérisée par  $n(n-1)/2$  fonctions  $\omega_{ij}$ , de telle sorte que :

$$\omega(x)(X, Y) = \sum_{i,j} \omega_{ij}(x) X^i Y^j, \quad (\text{II.46})$$

où  $x$  est un point de l'ouvert de définition,  $X = (X^i)$  et  $Y = (Y^i)$  deux vecteurs de  $\mathbf{R}^{2n}$ , les indices  $i$  et  $j$  variant de 1 à  $2n$ . On dit que la 2-forme différentielle  $\omega$  est *symplectique* si elle est *non dégénérée* en chaque point et si elle est *fermée*, c'est-à-dire<sup>39</sup> si :

$$\partial_i \omega_{jk} + \partial_j \omega_{ki} + \partial_k \omega_{ij} = 0, \quad (\text{II.47})$$

pour tout triplet d'indices  $i, j, k$ ; on note  $d\omega = 0$ .

Les trois propriétés que nous avons énoncées plus haut font des crochets de Lagrange les composantes, covariantes et contravariantes, d'une forme symplectique sur l'espace des mouvements keplériens de la planète. Les deux premières

<sup>39</sup>Cette formulation n'est pas très parlante. Dire qu'une forme différentielle  $\omega$  est fermée signifie précisément qu'elle est *localement exacte* : pour tout point  $x$  il existe un voisinage  $U$  et une forme différentielle  $\alpha$  tels que  $\omega|_U = d\alpha$ , c'est-à-dire :  $\omega_{ij} = \partial_i \alpha_j - \partial_j \alpha_i$ .

propriétés ont été soulignées explicitement par Lagrange, même s'il ne pouvait considérer à son époque ces crochets comme les éléments d'une matrice, *a fortiori* d'une 2-forme différentielle. Quant à la propriété de fermeture (la troisième propriété), il ne l'évoque pas. En effet, son importance n'est apparue que plus tard, avec la formalisation du calcul différentiel. Du point de vue de la mécanique, cette dernière propriété est la conséquence de l'existence du potentiel  $\Omega$  des forces de perturbation :  $X = \partial\Omega/\partial\mathbf{r}$ .

Lagrange a explicitement calculé la valeur de ses crochets, c'est-à-dire les composantes de la forme symplectique, au nombre de quinze<sup>40</sup>. Il en donne les expressions dans diverses cartes de l'espace des mouvements keplériens, c'est-à-dire pour divers choix d'éléments keplériens caractérisant les mouvements de la planète. Il n'y a pas grand intérêt à donner ici l'ensemble de ces expressions que l'on peut trouver dans [Lag5] et [Lag8].

REMARQUE. Lagrange note que l'on peut toujours choisir, comme constantes d'intégration, les positions et les vitesses à un instant donné plutôt que les éléments de la planète. L'expression des parenthèses et des crochets s'en trouve alors notablement simplifiée. En effet, dans ce cas les seuls crochets non nuls sont :

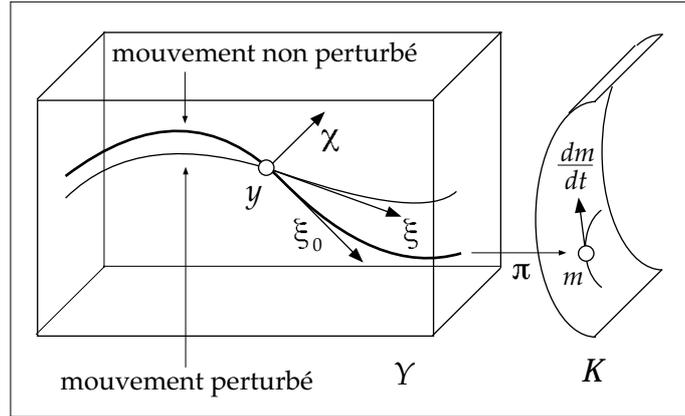
$$[\mathbf{v}_i, \mathbf{r}_i] = 1, \quad i = 1, 2, 3. \quad (\text{II.48})$$

Comme on le voit, les variables se regroupent par deux :  $\mathbf{r}_i$  avec  $\mathbf{v}_i$  et leurs crochets sont constants. Cette forme symplectique définie de façon générale sur  $\mathbf{R}^n \times \mathbf{R}^n$  est appelée aujourd'hui *forme symplectique canonique*.

Le *Théorème de Darboux* (voir annexe C) dit que toute forme symplectique possède au moins localement des coordonnées canoniques. Mais Lagrange, même s'il dit qu'« il y aurait toujours de l'avantage à utiliser ces constantes à la place des autres constantes  $a, b, c$ , etc. » [Lag8, volume II, page 76], n'a pratiquement pas utilisé ces coordonnées canoniques. En particulier, la carte  $(a, b, c, h, i, k)$  n'est pas canonique.

Revenons à la méthode de la variation des constantes telle qu'elle est présentée plus haut, et en particulier à la formule (II.20). Nous pouvons en donner

<sup>40</sup>Il y a *a priori* 36 composantes de la forme  $\omega$  mais pour des raisons d'antisymétrie, que nous avons déjà évoquées (note page 31), seules 15 de ces composantes sont indépendantes.

Figure II.5 : Projection de  $Y$  sur  $\mathcal{K}$ 

une justification en termes actuels. Considérons l'espace  $Y$  des conditions initiales du système étudié, c'est-à-dire l'espace des triplets  $y = (t, \mathbf{r}, \mathbf{v})$  où  $t \in \mathbf{R}$ ,  $\mathbf{r} \in \mathbf{R}^3 - \{0\}$  et  $\mathbf{v} \in \mathbf{R}^3$ . Les solutions de l'équation différentielle

$$\frac{d\mathbf{r}}{dt} = \mathbf{v} \quad \text{et} \quad \frac{d\mathbf{v}}{dt} = -\frac{\mathbf{r}}{r^3} + X, \quad (\text{II.49})$$

sont les courbes intégrales du feuilletage défini sur  $Y$  par :

$$y \mapsto \mathbf{R} \cdot \xi \quad \text{avec} \quad \xi = \begin{pmatrix} 1 \\ \mathbf{v} \\ -\mathbf{r}/r^3 + X \end{pmatrix}. \quad (\text{II.50})$$

Le vecteur  $\xi$  se décompose en  $\xi_0 + \chi$  :

$$\xi_0 = \begin{pmatrix} 1 \\ \mathbf{v} \\ -\mathbf{r}/r^3 \end{pmatrix} \quad \text{et} \quad \chi = \begin{pmatrix} 0 \\ 0 \\ X \end{pmatrix}. \quad (\text{II.51})$$

L'espace des mouvements kepleriens est l'espace quotient  $\mathcal{K} = Y/\mathbf{R} \cdot \xi_0$ , c'est-à-dire l'espace des courbes intégrales du feuilletage  $y \mapsto \mathbf{R} \cdot \xi_0$ . Considérons alors une feuille du feuilletage  $y \mapsto \mathbf{R} \cdot \xi$  passant par  $y = (t, \mathbf{r}, \mathbf{v})$ . Cette courbe se

projette sur l'espace des mouvements keplériens  $\mathcal{K}$  ; son équation est alors :

$$\frac{dm}{dt} = D\pi_y(\xi) = D\pi_y(\xi_0) + D\pi_y(\chi), \quad (\text{II.52})$$

où  $\pi : y \mapsto m$  est la projection de  $Y$  sur son quotient et  $D$  désigne l'application linéaire tangente. Or, par construction,  $D\pi_y(\xi_0) = 0$  ; il reste donc  $dm/dt = D\pi_y(\chi)$ . Un petit dessin vaut parfois mieux qu'un long discours (voir figure II.5). C'est la famille d'équations (II.21). Enfin, transformée en la famille d'équations (II.27), elle s'écrit encore<sup>41</sup> :

$$\frac{dm}{dt} = \omega^{-1}(d\Omega), \quad (\text{II.53})$$

où  $d\Omega$  est la différentielle de  $\Omega$ , sa dérivée extérieure, calculée au point  $m$  évidemment. La forme  $\omega$  peut être considérée comme une application de l'espace des vecteurs sur celui des covecteurs : l'application réciproque  $\omega^{-1}$  associe donc à une forme différentielle (par exemple  $d\Omega$ ) un vecteur tangent. C'est le sens qu'il faut donner à cette notation<sup>42</sup>.

Par analogie avec le cas euclidien, comme  $\omega$  est inversible, on appelle *gradient symplectique* de la fonction  $\Omega$  le champ de vecteurs  $\omega^{-1}(d\Omega)$ , et on note

$$\text{grad}(\Omega) \text{ ou } \text{grad}_\omega(\Omega) \stackrel{\text{def}}{=} \omega^{-1}(d\Omega). \quad (\text{II.54})$$

Avec ces conventions de langage, l'équation différentielle qui décrit la variation des constantes prend la forme suivante :

$$\frac{dm}{dt} = \text{grad}(\Omega). \quad (\text{II.55})$$

C'est sous cette forme qu'on la retrouve fréquemment dans les ouvrages actuels de mécanique. L'évolution du mouvement  $m$ , perturbé par le potentiel  $\Omega$ , est donc la courbe intégrale du gradient symplectique du potentiel de perturbation.

<sup>41</sup> $\omega$  est considérée comme une matrice,  $\omega^{-1}$  désigne la matrice inverse.

<sup>42</sup>Les physiciens utilisent parfois la notation  $\nabla\Omega$  pour désigner la dérivée extérieure de la fonction  $\Omega$ .

## II.5 Une remarque de Poincaré

La partie la moins précise du travail de Lagrange concerne sûrement la méthode d'approximation utilisée. Je voudrais à ce propos souligner qu'hormis ces méthodes d'approximation les conclusions de Lagrange sont rigoureusement établies, même si la présentation qu'il en a faite, et que j'ai essayé de reproduire ici, ne respecte pas les canons actuels de la rigueur mathématique. En ce sens, les transformations qu'il apporte aux équations initiales ne sont pas d'une grande utilité puisque celles qu'il obtient leur sont absolument équivalentes. Laissons parler Lagrange encore une fois :

« Ainsi on peut regarder les équations précédentes entre les nouvelles variables  $a, b, c$ , etc. comme les transformées des équations en  $x, y, z$  ; mais ces transformations seraient peu utiles pour la solution générale du problème. Leur grande utilité est lorsque la solution rigoureuse est impossible, et que les forces perturbatrices sont très petites ; elles fournissent alors un moyen d'approximation. »

Mais la justification de ces méthodes a occupé un grand nombre de mathématiciens après lui et non des moindres. Poincaré soulignait ainsi dans l'introduction de sa célèbre *Nouvelle mécanique céleste* [Poin] :

« Ces méthodes qui consistent à développer les coordonnées des astres suivant les puissances des masses, ont en effet un caractère commun qui s'oppose à leur emploi pour le calcul des éphémérides à longue échéance. Les séries obtenues contiennent des termes dits *séculaires*, où le temps sort des signes des sinus et cosinus, et il en résulte que leur convergence pourrait devenir douteuse si l'on donnait à ce temps  $t$  une grande valeur.

La présence de ces termes séculaires ne tient pas à la nature du problème, mais seulement à la méthode employée. Il est facile de se rendre compte, en effet, que si la véritable expression d'une coordonnée contient un terme en  $\sin \alpha mt$ ,  $\alpha$  étant une constante et  $m$  l'une des masses, on trouvera quand on voudra développer suivant les puissances de  $m$ ,

des termes séculaires  $\alpha mt - \alpha^3 m^3 t^3 / 6 + \dots$  et la présence de ces termes donnerait une idée très fautive de la véritable forme de la fonction étudiée. »

Cette objection est sans nul doute très pertinente et a conduit, notamment grâce aux travaux de Poincaré, au développement de la géométrie symplectique — en particulier en ce qui concerne son application à la mécanique. De nouvelles théories sont nées, comme par exemple la théorie des systèmes complètement intégrables et de leur perturbation qui a donné le fameux théorème<sup>43</sup> de Kolmogorov–Arnold–Moser, sur la stabilité de nombreux mouvements après perturbation (voir [Arn1] et [Arn2])

---

<sup>43</sup>Théorème difficile, qui nécessite beaucoup de précautions dans l'énoncé et d'expertise dans ses applications.



Figure II.6 : Galerie

---

# Les principes de la mécanique analytique

---

Comme nous l'avons montré au chapitre précédent, les origines de la géométrie symplectique remontent à Lagrange et à son souci de structurer les constructions de la mécanique qu'il avait développées pour répondre à des questions autant concrètes, comme celles, par exemple, sur la stabilité des planètes, que théoriques, comme celles posées par Euler sur le calcul des variations. Nous essayons dans ce chapitre d'en donner une présentation moderne synthétique, à partir du principe des travaux virtuels, ainsi que le fait Lagrange.

## III.1 Le principe des travaux virtuels

Il s'agit de décrire le mouvement d'une particule de masse  $m$  repérée, dans l'espace, par sa position  $r$ . Cette particule subit une force extérieure  $F$ , au sens de Newton, comme il a été expliqué plus haut. Il s'agit donc de résoudre l'équation différentielle ordinaire :

$$F(t, r(t), v(t)) = m \frac{dv(t)}{dt}, \quad v(t) = \frac{dr(t)}{dt}. \quad (\text{III.1})$$

Changeons de siècle et interprétons les équations de Newton à partir du « principe des travaux virtuels » de d'Alembert : les équilibres d'un point matériel, soumis à

l'influence de forces  $F_i$ , sont les solutions de l'équation :

$$\sum_i F_i \cdot \delta r = 0, \quad (\text{III.2})$$

où  $\delta r$  est une variation<sup>1</sup> de  $r$ , compatible avec les contraintes existantes. Nous verrons plus loin, par des exemples, comment se traite la question des contraintes dans ce genre de problème.

Ce principe des équilibres se nomme *principe des travaux virtuels*. Pour employer le mode de pensée et le langage des mécaniciens : le *travail* d'une force  $F$  est, par définition, le produit (scalaire) de la force par le déplacement de son point d'appui ; la somme  $\sum_i F_i \cdot \delta r$  représente donc le travail de l'ensemble des forces  $F_i$ , agissant en  $r$ , pour un déplacement *virtuel* infinitésimal  $\delta r$ . Supposer l'équilibre, c'est supposer que le travail du système de forces est infinitésimalement nul pour tout déplacement virtuel du point d'appui  $\delta r$ , compatible avec les contraintes.

Comme tout principe de ce genre, le principe des travaux virtuels a pour but de remplacer la répétition de l'analyse par la méthode, permettant ainsi à notre esprit de se dégager des problèmes déjà résolus pour se concentrer sur les nouveaux. L'expression de l'équilibre par la formule  $\sum_i F_i \cdot \delta r = 0$  trouve son intérêt dans le traitement des contraintes ; par exemple, si l'on veut traiter l'équilibre d'un point matériel sur une surface d'équation  $f = \text{const.}$ , l'équation  $\sum_i F_i \cdot \delta r = 0$  couplée à l'équation de contrainte  $\text{grad } f \cdot \delta r = 0$  donnera immédiatement comme résultat : «  $\sum_i F_i$  proportionnel à  $\text{grad } f$  », où le gradient est ici le gradient ordinaire associé à la métrique euclidienne, comme le montre l'exemple particulier suivant.

EXEMPLE. Considérons une bille de métal roulant librement sur une sphère aimantée  $S$  (*contrainte* à ne pouvoir s'en détacher). Supposons que cette bille soit soumise à une gravitation  $F$  constante, dont la direction s'appelle la verticale. Soit  $r$  le point occupé par la bille sur la sphère  $S$ . Ce point sera en équilibre, d'après le principe des travaux virtuels, si et seulement si  $F \cdot \delta r = 0$  pour toute variation  $\delta r$  du point de contact  $r$ , compatible avec la contrainte, c'est-à-dire pour tout vecteur tangent  $\delta r$  au point  $r \in S$ . Il faut donc introduire cette contrainte

<sup>1</sup>étant donné un point  $x$  d'une variété  $X$ , ce que nous appelons variation (au sens moderne) est un vecteur tangent  $\delta x$  à  $X$ , au point  $x$ , formellement :  $\delta x \in T_x X$ . Cette notation ancienne a été réhabilitée par Souriau [Sou2].

dans l'équation du problème : être tangent à la sphère  $S$  au point  $r$ , c'est être orthogonal à  $r$ , autrement dit  $F \cdot \delta r = 0$  pour tout  $\delta r$  tel que  $r \cdot \delta r = 0$ . On déduit donc que  $F$  et  $r$  sont nécessairement colinéaires. Il y a deux points où  $F$  et  $r$  sont colinéaires : les points haut et bas de la sphère  $S$ . Ce sont les deux points d'équilibre de la bille. Un autre type de considérations permet de qualifier le point haut d'*équilibre instable* et le point bas d'*équilibre stable*. Voilà illustré, dans ce cas particulier, comment la méthode repose l'esprit.

Lagrange va traiter le problème de dynamique qui se pose à lui — trouver les solutions des équations de Newton, c'est-à-dire trouver un mouvement, une dynamique — en le ramenant à un problème de statique<sup>2</sup>, de recherche d'équilibre. Pour cela, il considère le vecteur

$$I = -m \frac{dv}{dt} \quad (\text{III.3})$$

comme la *force de l'inertie* à laquelle est soumise la particule lors de son mouvement. Le système des équations de Newton devient alors une simple condition d'équilibre, les forces extérieures équilibrant les forces d'inertie à chaque instant. Le temps s'impose ainsi comme un nouveau paramètre : l'équilibre n'est plus statique mais dynamique. En appliquant alors le principe des travaux virtuels les équations de Newton deviennent :

$$(F + I) \cdot \delta r = 0, \quad (\text{III.4})$$

en tout point  $r$  du mouvement et pour tout  $\delta r$ , compatible avec les contraintes.

---

<sup>2</sup>Lagrange écrit explicitement [Lag8, volume I p. 223] :

« Le *Traité de Dynamique* de d'Alembert, qui parut en 1743, mit fin à ces espèces de défis, en offrant une méthode directe et générale pour résoudre, ou du moins pour mettre en équation tous les problèmes de la Dynamique que l'on peut imaginer. Cette méthode réduit toutes les lois du mouvement des corps à celles de leur équilibre, et ramène ainsi la Dynamique à la Statique. »

Pour ramener les lois de la dynamique à celles de la statique, Lagrange ne s'est pas borné à reprendre les méthodes de ses prédécesseurs mais il les a généralisées en les simplifiant et en les unifiant.

Comme Lagrange, supposons maintenant que le mobile se déplace sans contrainte et que la force  $F$  dérive<sup>3</sup>

d'un potentiel<sup>4</sup>  $U$  :

$$F = -\frac{\partial U}{\partial r}. \quad (\text{III.5})$$

On transforme ensuite l'équation des travaux virtuels par quelques manipulations algébriques :

$$\begin{aligned} (F + I).\delta r &= -\frac{\partial U}{\partial r}(\delta r) - m\frac{dv}{dt} \cdot \delta r \\ &= -\frac{\partial U}{\partial r}(\delta r) - \frac{d}{dt}(mv\delta r) + mv\frac{d}{dt}(\delta r) \\ &= -\frac{\partial U}{\partial r}(\delta r) - \frac{d}{dt}(mv\delta r) + mv\delta v \\ &= \delta\left(\frac{mv^2}{2} - U\right) - \frac{d}{dt}(mv\delta r). \end{aligned}$$

Définissons alors :

$$L = \frac{mv^2}{2} - U, \quad (\text{III.6})$$

fonction de  $(t, r, v)$  définie sur  $\mathbf{R} \times \mathbf{R}^3 \times \mathbf{R}^3$ , ou tout au moins sur un ouvert de cet espace. Cette fonction s'appelle le *lagrangien* du problème, elle a été introduite par Lagrange à la fin du dix-huitième siècle, dans sa *Mécanique Analytique* [Lag8, Tome I, pp. 299–300], où elle est notée  $Z$ . L'équation de Newton devient après ce

---

<sup>3</sup>Lagrange remarque que cette propriété, vérifiée pour les systèmes de points en interaction gravitationnelle, lui permet de démontrer le théorème de conservation de la force vive. Il en a ensuite fait une condition des lois de la mécanique elle-même. Lagrange note  $Pdp + Qdq + Rdr + \dots$  le travail des forces extérieures ; il écrit précisément [Lag8, Tome I, pp. 290–291] :

« ... Mais cette opération devient encore plus facile, lorsque les forces sont telles, que la somme des moments, c'est-à-dire la quantité

$$Pdp + Qdq + Rdr + \dots$$

est intégrable, ce qui, comme nous l'avons déjà observé, est proprement le cas de la nature. »

<sup>4</sup>La force  $F$  est alors considérée comme un covecteur plutôt que comme un vecteur.

passage à travers le filtre du principe de d'Alembert :

$$\delta L = \frac{d}{dt}(mv\delta r). \quad (\text{III.7})$$

On a l'habitude de présenter cette équation sous sa version intégrale, connue sous le nom de « principe de moindre action » (de Maupertuis, de d'Alembert...). On intègre les deux membres de cette équation le long d'un intervalle de temps quelconque  $[t_a, t_b]$ , et l'on suppose que la variation  $\delta r$  est nulle aux extrémités ; on obtient la condition équivalente suivante :

$$\delta \int_{t_a}^{t_b} L dt = 0, \quad \text{pour tout } t \mapsto \delta r, \quad \text{tel que } \delta r(t_a) = \delta r(t_b) = 0. \quad (\text{III.8})$$

Le lecteur peut vérifier lui-même que les solutions de cette *équation aux variations* sont exactement les solutions de l'équation de Newton originale. Nous avons simplement interprété les solutions de l'équation de Newton comme les *courbes extrémales* de l'*action lagrangienne* :

$$A([t \mapsto r]) = \int_{t_a}^{t_b} L(t, r(t), v(t)) dt \quad \text{avec} \quad v(t) = \frac{dr}{dt}. \quad (\text{III.9})$$

Ainsi, le principe du calcul des variations consiste à sélectionner, selon un principe d'*extremum*, les *mouvements réels* de la particule, parmi tous ses *mouvements virtuels*, c'est-à-dire tous les chemins possibles  $t \mapsto r$ .

REMARQUE. Telle quelle cette construction n'a que peu d'intérêt, sinon celui d'offrir un principe — que l'on peut apprécier diversement — aux lois de la mécanique newtonienne. Mathématiquement parlant, le passage des équations de Newton aux équations aux variations de d'Alembert–Lagrange n'offre aucun intérêt, le problème restant inchangé et la difficulté identique. Peut-être le choix de tel lagrangien plutôt qu'un autre a-t-il pu sembler, dans certaines conditions, plus justifié que le choix explicite de l'expression d'une force ? Ou bien l'utilisation de la terminologie quasi-mystique « moindre action » peut-elle forcer la foi ? Le fait est que cette construction a le mérite de nous conduire à la *structure symplectique* de l'espace des solutions de ces équations, et de mettre en évidence la structure très particulière de cet espace qui aurait pu rester cachée si nous en étions restés à l'expression initiale des équations de Newton.

Nous allons maintenant abandonner la lecture rigoureuse des œuvres de Lagrange pour adopter un ton un peu plus moderne. Elargissons dès à présent le cadre des équations de la mécanique en changeant l'espace vectoriel  $\mathbf{R}^3$  pour une variété quelconque  $X$ , et définissons précisément les différents objets que nous venons d'introduire :

DÉFINITION. *On appelle lagrangien (d'ordre 1), sur une variété différentielle  $X$ , toute fonction différentiable  $L$  définie sur un ouvert  $V$  de  $\mathbf{R} \times TX$  à valeurs réelles. On appelle action d'un chemin  $\gamma : [a, b] \rightarrow X$  l'intégrale :*

$$A(\gamma) = \int_a^b L(t, \gamma(t), \dot{\gamma}(t)) dt. \quad (\text{III.10})$$

*On appelle solution du problème variationnel défini par le lagrangien  $L$ , tout chemin  $\gamma$  extrémal pour l'action  $A$ , c'est-à-dire tel que :*

$$\delta A = 0 \quad \forall \delta\gamma \quad \text{avec} \quad \delta\gamma(a) = \delta\gamma(b) = 0. \quad (\text{III.11})$$

Puisque nous introduisons ici pour la première fois cette notation, il est peut-être nécessaire d'en préciser le sens. Une variation  $\delta\gamma$  d'un chemin  $\gamma$  est le chemin obtenu en dérivant une famille à un paramètre  $\gamma_s$  par rapport à  $s$ , pour la valeur  $s = 0$ , autrement dit :

$$\delta\gamma = \left. \frac{\partial \gamma_s}{\partial s} \right|_{s=0}. \quad (\text{III.12})$$

Cette variation est donc, elle aussi, un chemin mais tracé dans le fibré tangent  $TX$ , au-dessus de  $\gamma$ , c'est-à-dire tel que :  $\delta\gamma(t) \in T_{\gamma(t)}X$ . Puisque  $A$  est fonction de  $\gamma$ , la famille  $\gamma_s$  définit une famille d'actions  $A_s = A(\gamma_s)$ , la variation  $\delta A$  étant évidemment le résultat de la dérivation :

$$\delta A = \left. \frac{\partial A(\gamma_s)}{\partial s} \right|_{s=0}. \quad (\text{III.13})$$

On présente souvent les équations de ce problème variationnel sous la forme des équations obtenues en calculant naïvement la variation de l'action :

$$\begin{aligned} \delta A &= \int_a^b \left[ \frac{\partial L}{\partial r}(\delta r) + \frac{\partial L}{\partial v}(\delta v) \right] dt \\ &= \int_a^b \left[ \frac{\partial L}{\partial r} - \frac{d}{dt} \left( \frac{\partial L}{\partial v} \right) \right] (\delta r) dt + \int_a^b \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial v}(\delta r) \right) \right] dt. \end{aligned} \quad (\text{III.14})$$

Le dernier terme de cette identité est nul puisque la variation est supposée à support compact ; il reste donc le système :

$$\delta A = 0 \quad \Leftrightarrow \quad \frac{\partial L}{\partial r} - \frac{d}{dt} \left( \frac{\partial L}{\partial v} \right) = 0. \quad (\text{III.15})$$

Ces équations sont communément appelées *équations d'Euler-Lagrange*.

REMARQUE. Malheureusement si, comme on le verra plus loin, on peut donner facilement un sens intrinsèque au terme  $\partial L/\partial v$ , il n'en est pas de même du terme  $\partial L/\partial r$ . En effet, s'il avait un sens ce terme désignerait la dérivée partielle du lagrangien  $L$  à  $v$  fixé :

$$\frac{\partial L}{\partial r}(u) = \lim_{\varepsilon \rightarrow 0} \frac{L(r + \varepsilon u, v) - L(r, v)}{\varepsilon}. \quad (\text{III.16})$$

Or — il est important de le souligner — si  $v \in T_r X$  alors  $v \notin T_{r+\varepsilon u} X$ , autrement dit : cette formule (III.16) n'a pas de sens, ou alors localement, dans les coordonnées d'une carte. On pourrait, pour éviter ce problème, identifier les espaces tangents voisins (définir ce qu'on appelle une *connexion*) mais il n'est jamais bon d'introduire une structure notoirement étrangère au problème que l'on traite. Nous allons voir immédiatement comment nous pouvons avantageusement nous en passer.

## III.2 Homogénéisation du lagrangien

Même si cela est équivalent, il est parfois commode de chercher les solutions d'un problème variationnel sous forme de courbes non paramétrées dans l'espace des conditions initiales  $(t, x, v)$  plutôt que sous la forme de courbes paramétrées :  $t \mapsto x(t)$ , dans l'espace  $X$ . L'espace des  $(t, r, v)$  est l'espace naturel de définition du lagrangien  $L$  et il est donc tout aussi naturel d'y rechercher les solutions du problème variationnel associé. Nous le noterons  $Y$  et l'appellerons l'*espace d'évolution du système*, ou encore *espace des conditions initiales*.

Choisissons un paramétrage auxiliaire  $s \mapsto (t(s), x(s), v(s))$  de la courbe  $\{(t, x(t), v(t))\}_t$ , définie dans  $Y$  par le chemin  $\gamma : t \mapsto x(t)$ . Introduisons les

dérivées de  $(t, x, v)$  par rapport à  $s$  et notons :

$$\dot{t} = \frac{dt}{ds}, \quad \dot{x} = \frac{dx}{ds}, \quad \dot{v} = \frac{dv}{ds}, \quad (\text{III.17})$$

de telle sorte que :

$$A = \int_{s_a}^{s_b} L\left(t, x, \frac{\dot{x}}{\dot{t}}\right) \dot{t} ds = \int_{s_a}^{s_b} l(q, \dot{q}) ds, \quad (\text{III.18})$$

avec :

$$q = \begin{pmatrix} t \\ x \end{pmatrix}, \quad \dot{q} = \begin{pmatrix} \dot{t} \\ \dot{x} \end{pmatrix}, \quad \text{et} \quad l(q, \dot{q}) = L\left(t, x, \frac{\dot{x}}{\dot{t}}\right) \dot{t}. \quad (\text{III.19})$$

Notons par la lettre  $Q$  l'espace de la variable  $q = (t, x)$  lorsque  $x$  parcourt  $X$  et  $t$  parcourt  $\mathbf{R}$ . Le lagrangien  $l$  est maintenant défini sur l'espace tangent<sup>5</sup>  $TQ$ , un ouvert de  $\mathbf{R}^4 \times \mathbf{R}^4$  ; il ne dépend pas du paramètre  $s$  et il est évidemment homogène de degré 1. Nous pouvons lui appliquer la formule d'Euler :

$$l(q, \dot{q}) = p(\dot{q}), \quad \text{où} \quad p = \frac{\partial l}{\partial \dot{q}}. \quad (\text{III.20})$$

La notation  $\partial l / \partial \dot{q}$  désigne l'application linéaire tangente de la restriction  $l|_{T_q Q}$ . C'est donc une application linéaire bien définie, de  $T_q Q$  à valeur réelle, c'est-à-dire un élément de l'espace vectoriel dual  $T_q^* Q$ . Nous noterons  $P$  l'application  $(q, \dot{q}) \mapsto (q, p)$ , et l'appellerons *application de Legendre*<sup>6</sup> :

$$P : TQ \rightarrow T^*Q, \quad P(q, \dot{q}) = (q, p) \quad p = \frac{\partial l}{\partial \dot{q}} = D(l|_{T_q Q})_{\dot{q}} \in T_q^* Q. \quad (\text{III.21})$$

Arrêtons nous un instant sur cette formule. À ce stade, nous pouvons oublier comment nous en sommes arrivés là, oublier notre particule de départ, l'homogénéisation du lagrangien, et considérer que :

<sup>5</sup>Les bases de la géométrie différentielle ordinaire sont supposées connues du lecteur, sinon nous renvoyons aux ouvrages classiques.

<sup>6</sup>Attention, cette dénomination n'est pas usuelle.

1.  $q$  est un point courant d'une variété quelconque  $Q$ ,
2.  $\dot{q}$  est un vecteur tangent à  $Q$  au point  $q$ ,
3.  $l$  est une fonction réelle, homogène<sup>7</sup> de degré 1, définie sur un ouvert de l'espace tangent  $TQ$  privé de la section nulle.

Ce sera le cadre général du calcul lagrangien des variations comme nous l'entendons. Cet espace  $Q$  est appelé en mécanique *espace des configurations*, car chaque point  $q \in Q$  représente un état possible du système étudié : une *configuration*. Dans le cas d'une particule, c'est, comme nous l'avons vu,  $\mathbf{R} \times \mathbf{R}^3$ . Ce sera  $\mathbf{R} \times (\mathbf{R}^3)^N$  dans le cas d'un système à  $N$  particules en interaction,  $\mathbf{R} \times SO(3)$  dans le cas d'un solide etc. Dans chacun de ces cas, un mouvement virtuel est une courbe (non paramétrée) de  $Q$ , et parmi eux se trouvent les mouvements réels.

Revenons maintenant à l'écriture de l'action lagrangienne  $A$  le long de la courbe  $s \mapsto (q, \dot{q})$  :

$$A = \int_{s_a}^{s_b} p(\dot{q}) ds = \int_{s_a}^{s_b} p\left(\frac{dq}{ds}\right) ds. \quad (\text{III.22})$$

Considérons la forme<sup>8</sup> de Liouville  $\lambda$  définie, sur l'espace cotangent<sup>9</sup>  $T^*Q$  à  $Q$ , par :

$$\forall y = (q, p) \in T^*Q, \quad \forall \delta y \in T_y T^*Q \quad : \quad \lambda(\delta y) = p(\delta q), \quad (\text{III.23})$$

où  $\delta q$  est la projection du vecteur  $\delta y$  sur  $Q$  ; c'est donc un vecteur tangent à  $Q$  au point  $q$ , et que nous évaluons sur la 1-forme  $p$ .

<sup>7</sup>Nous dirons homogène pour positivement homogène, c'est-à-dire homogène par multiplication d'un nombre strictement positif

<sup>8</sup> $\lambda$  est ce que l'on appelle une *forme différentielle* sur  $Q$ , c'est-à-dire une fonction différentiable définie sur  $TQ$  à valeurs réelles, et linéaire en restriction à chaque espace tangent  $T_q Q$  lorsque  $q$  parcourt  $Q$ .

<sup>9</sup>Rappelons que l'espace cotangent  $T^*Q$  à  $Q$ , est l'espace de toutes les formes linéaires définies sur chaque espace tangent  $T_q Q$ , lorsque  $q$  parcourt  $Q$ . Cet espace est naturellement *fibré* par la projection *point base* qui consiste à associer à une forme linéaire  $\alpha$  le point base de l'espace tangent sur lequel elle est définie. On note formellement  $\pi : T^*Q \rightarrow Q$ . Rappelons si nécessaire qu'une forme linéaire  $\alpha$  définie sur un espace vectoriel réel  $E$  est une application linéaire  $\alpha : E \rightarrow \mathbf{R}$ . Une forme différentielle  $\alpha$  sur  $Q$  est une application particulière (une section) de  $Q$  dans  $T^*Q$ .

Soit  $\varpi$  l'image réciproque de  $\lambda$  sur  $TQ$ , par l'application de Legendre définie plus haut :

$$\varpi = P^*\lambda. \quad (\text{III.24})$$

Cette forme  $\varpi$  est appelée *forme de Cartan* associée au lagrangien  $l$ .

Introduisons le *relevé holonome* du chemin  $\gamma$ , qui décrit le mouvement virtuel du système, sur lequel se fait l'intégration du lagrangien :

DÉFINITION. Soit  $\gamma : s \mapsto q$  un chemin dans  $Q$  ; nous appellerons relevé holonome de  $\gamma$  le chemin  $s \mapsto (q, \dot{q})$  dans  $TQ$  tel que  $\dot{q} = dq/ds$ .

L'action  $A$  s'écrit alors sous forme condensée :

$$A = \int_{\bar{\gamma}} \varpi. \quad (\text{III.25})$$

Toute variation de  $\gamma$  entraîne une variation de l'action, donnée par la formule suivante (voir annexe F) :

$$\delta A = \int_{\bar{\gamma}} d\varpi(\delta\bar{\gamma}) + \int_{\bar{\gamma}} \varpi(\delta\bar{\gamma}), \quad (\text{III.26})$$

où  $\delta\bar{\gamma}$  est la variation du relevé holonome  $\bar{\gamma}$  associée à la variation  $\delta\gamma$ . Il est peut-être utile de préciser le sens des notations utilisées :

1.  $\gamma$  est en réalité *une courbe* de  $Q$  : c'est un arc  $s \mapsto q$ , *modulo* reparamétrage.
2.  $\delta\gamma$  est une *variation* de  $\gamma$  : c'est une courbe  $s \mapsto \delta q$  du tangent  $TQ$ , telle que  $\delta q \in T_q Q$  ; on dit que c'est un *relevé* de  $\gamma$  dans  $TQ$ .
3. La courbe  $\delta\bar{\gamma}$  n'est pas un relevé quelconque dans  $TTQ$  de  $\bar{\gamma}$  : c'est le relevé holonome de la variation  $\delta\gamma$ , autrement dit :

$$\delta\bar{\gamma} : s \mapsto \frac{d}{ds}[\delta\gamma(s)]. \quad (\text{III.27})$$

Nous dirons que c'est une *variation holonome* de  $\bar{\gamma}$ .

4.  $d\varpi(\delta\bar{\gamma})$  est le contracté de  $d\varpi$  avec la variation  $\delta\bar{\gamma}$  : c'est une 1-forme centrée sur la courbe  $\bar{\gamma}$ .

Compte tenu de ces précisions, la variation  $\delta A$  s'écrit encore :

$$\delta A = \int_{s_a}^{s_b} d\varpi \left( \delta\bar{\gamma}(s), \frac{d}{ds}\bar{\gamma}(s) \right) ds + [\varpi(\delta\bar{\gamma}(s))]_{s_a}^{s_b}. \quad (\text{III.28})$$

REMARQUE. Cette écriture est relativement compliquée du fait du nombre d'espaces mis en jeu par cette construction, qui va de  $Q$  (l'espace de configuration) et son tangent  $TQ$  pour la définition du lagrangien et de la forme de Cartan, jusqu'à l'espace tangent  $TTQ$  de  $TQ$  qui héberge les variations nécessaires. Mais cette relative complexité ne doit pas nous effrayer, car en réalité les choses sont simples. Utilisons une carte locale de  $Q$ ,  $q = (q_i)$ ; soit  $p = (p_i)$  les coordonnées de  $p$  dans la base duale de la base associée à cette carte. Si le chemin  $\gamma$  est contenu complètement dans le domaine de cette carte, la variation de l'action est donnée par :

$$\delta A = \int_{s_a}^{s_b} \sum_{i=1}^n [\delta p_i \frac{dq_i}{ds} - \frac{dp_i}{ds} \delta q_i] ds + \sum_{i=1}^n [p_i \delta q_i]_{s_a}^{s_b}. \quad (\text{III.29})$$

La lettre  $n$  désigne la dimension de l'espace  $Q$ .

Revenons à la formule précédente (III.28) de la variation de l'action, et notons que :

PROPOSITION. *Une condition suffisante pour que  $\gamma$  soit une solution du problème est que  $\bar{\gamma}$  soit tangent en tout point au noyau de  $d\varpi$ , c'est-à-dire qu'il soit contenu dans son feuilletage caractéristique.*

DÉMONSTRATION. *En effet, si  $\bar{\gamma}$  est dans le feuilletage caractéristique de  $d\varpi$ , c'est-à-dire si :*

$$\frac{d}{ds}\bar{\gamma}(s) \in \ker d\varpi, \quad (\text{III.30})$$

*alors la variation  $\delta A$  de l'action est nulle pour toute variation arbitraire de  $\bar{\gamma}$  (formule III.28), et a fortiori pour des variations holonomes. ■*

Il faut noter que le feuilletage caractéristique de  $d\varpi$  est au moins de dimension deux sur  $TQ$ ; en effet,  $l$  étant homogène de degré un sur  $TQ$ , l'application de Legendre

est homogène de degré zéro, c'est-à-dire invariante sous l'action des dilatations positives  $(q, \dot{q}) \mapsto (q, \alpha\dot{q})$ , avec  $\alpha > 0$  :

$$P(q, \alpha\dot{q}) = P(q, \dot{q}) \quad \forall (q, \dot{q}) \in TQ - Q \quad \forall \alpha > 0. \quad (\text{III.31})$$

L'invariance de l'application de Legendre  $P$  nous conduit naturellement à introduire un nouvel espace, sur lequel  $P$  est définie : l'espace des directions tangentes  $SQ = (TQ - Q)/]0, \infty[$ .

### III.3 L'espace des directions tangentes

La forme de Cartan  $\varpi$  est donc invariante par les dilatations positives agissant sur chaque espace tangent  $T_qQ - \{0\}$  en tout point  $q \in Q$  ; d'autre part, les demi-droites engendrées par ces dilations sont dans le noyau de  $\varpi$ . On en conclut (par la formule de Cartan par exemple) que ces demi-droites sont dans le noyau de  $d\varpi$ . Puisque  $d\varpi$  est antisymétrique, son rang est pair. La dimension de  $TQ$  étant paire, son noyau est de dimension paire et donc au moins égal à deux.

Mais nous venons de montrer que nous pouvons nous débarrasser de cette direction inessentielle dans le noyau de  $d\varpi$ , qui ne correspond qu'à la liberté qui nous est donnée par l'homogénéité du lagrangien  $l$  de reparamétriser les solutions du problème.

Nous considérerons donc l'espace des *demi-droites tangentes*, que nous appelons aussi *directions tangentes*, c'est-à-dire le quotient de  $TQ$ , privé de la section nulle, par les dilatations positives  $(q, \dot{q}) \mapsto (q, \alpha\dot{q})$ ,  $\alpha > 0$ . Cet espace est encore une variété fibrée sur  $Q$  ; chaque fibre au dessus d'un point  $q$  est l'espace des demi-droites tangentes en  $q$  ; nous noterons  $[\dot{q}]$  la demi-droite engendrée par  $\dot{q}$ ,  $\dot{q} \neq 0$ . Nous noterons  $S_qQ$  l'espace des demi-droites tangentes au point  $q$ ,  $S$  étant mis pour *sphère* puisque  $S_qQ$  est topologiquement une sphère. Nous noterons  $SQ$  le quotient total  $(TQ - Q)/]0, \infty[$ , et nous l'appellerons l'espace des directions tangentes ; il est de dimension  $2n - 1$  ; nous noterons  $\pi$  la projection de  $TQ - Q$  sur  $SQ$  :

$$SQ = (TQ - Q)/]0, \infty[ \quad \pi : \begin{array}{ccc} TQ - Q & \longrightarrow & SQ \\ \pi(q, \dot{q}) & \longmapsto & (q, [\dot{q}]) \end{array} \quad (\text{III.32})$$

Puisque l'application de Legendre est invariante par dilatation, elle se factorise à travers  $SQ$  :

$$\begin{array}{ccc} TQ - Q & & \\ \pi \downarrow & \searrow P & \\ SQ & \xrightarrow{P_S} & T^*Q \end{array}$$

$$P_S : SQ \rightarrow T^*Q \quad P = P_S \circ \pi : P_S(q, [\dot{q}]) = P(q, \dot{q}) = (q, p). \quad (\text{III.33})$$

Nous noterons  $\varpi_S$  l'image directe sur  $SQ$ , par  $\pi$ , de la forme de Cartan  $\varpi$  définie sur  $TQ - Q$ , mais nous conserverons l'appellation *forme de Cartan* :

$$\varpi_S = \pi_* \varpi \quad \text{ou} \quad \varpi = \pi^* \varpi_S. \quad (\text{III.34})$$

À tout chemin  $\gamma$ , nous pouvons associer son relevé holonome sur  $SQ$ , obtenu par projection du relevé holonome sur  $TQ$  : c'est la courbe des directions tangentes de  $\gamma$ . Nous le noterons  $[\bar{\gamma}]$ . L'expression (III.28) de la variation de l'action le long du chemin  $\gamma$  peut-être considérée comme s'appliquant au relevé holonome de  $\gamma$  dans  $SQ$  :

$$\delta A = \int_{s_a}^{s_b} d\varpi_S \left( \delta[\bar{\gamma}(s)], \frac{d}{ds} [\bar{\gamma}(s)] \right) ds + [\varpi([\delta\bar{\gamma}(s)])]_{s_a}^{s_b}. \quad (\text{III.35})$$

La variété  $SQ$  étant de dimension impaire, le noyau de  $d\varpi_S$  est au moins de dimension un. Supposons qu'il soit de dimension constante exactement égale à un, alors on a :

**PROPOSITION.** *Soit  $l$  un lagrangien homogène de degré un défini sur le tangent d'un espace de configuration  $Q$ . Si la dimension du feuilletage caractéristique de  $d\varpi_S$ , définie sur  $SQ$ , est constante et égale à 1, les solutions du problème variationnel associé à  $l$  sont les courbes intégrales du feuilletage caractéristique de  $d\varpi_S$ .*

**DÉMONSTRATION.** *Les courbes cherchées sont les solutions d'un système différentiel du premier ordre sur  $TQ$  comme le montrent les équations d'Euler-Lagrange,*

écrites dans une carte locale de  $TQ$  :

$$\delta A = 0 \quad \Leftrightarrow \quad \frac{dq}{ds} = \dot{q} \quad \frac{d\dot{q}}{ds} = \frac{d}{ds} \left( \frac{\partial l}{\partial \dot{q}} \right).$$

Ainsi, par chaque point  $(q, [\dot{q}]) \in SQ$  passe une solution unique, mais comme nous l'avons vu plus haut, la courbe intégrale du feuilletage caractéristique passant par ce point est solution. C'est donc la solution cherchée. La condition suffisante énoncée plus haut est ainsi devenue nécessaire. ■

Parce que cette situation particulière est importante, on introduit un nouveau terme de vocabulaire :

**DÉFINITION.** On appelle forme présymplectique toute 2-forme fermée dont le noyau est de dimension constante.

**REMARQUE.** Je voudrais attirer l'attention du lecteur sur le seul point délicat de cette proposition. Le problème que nous avons à résoudre est un cas particulier du problème général suivant : soit  $X$  une variété différentielle et  $\beta$  une 1-forme sur  $X$ , appelons *action d'un chemin*  $\gamma$  de  $X$  l'intégrale de  $\beta$  sur  $\gamma$

$$A = \int_{\gamma} \beta, \quad (\text{III.36})$$

La variation de l'action, associée à une variation  $\delta\gamma$ , est donnée par la formule générale (voir annexe F) :

$$\delta A = \int_{\gamma} d\beta(\delta\gamma) + [\beta(\delta\gamma)]_{\partial\gamma}. \quad (\text{III.37})$$

Il est évident que l'action est extrémale si et seulement si  $\gamma$  est tangent en tout point au feuilletage caractéristique de  $d\beta$ . Mais dans ce cas la variation du chemin  $\delta\gamma$  est arbitraire : c'est un relevé quelconque de  $\gamma$  dans  $TX$ . Or, dans le cas qui nous occupe (celui d'un lagrangien  $l$  défini sur  $TQ$ ) la variation est holonome. Nous obtenons un résultat identique seulement parce que nous avons fait l'hypothèse que le feuilletage caractéristique de  $d\varpi_S$  est constant, de dimension 1, et parce que les courbes que nous cherchons sont des solutions d'un système différentiel ordinaire.

Cela implique en particulier que, dans ce cas, les solutions du problème général, pour  $X = TQ$  et  $\beta = \varpi_S$ , sont nécessairement les relevés holonomes de chemins dans  $Q$ .

Ainsi que nous venons de le voir, dans tous les cas, l'espace intéressant est l'espace des directions tangentées  $SQ$ , ou plutôt l'ouvert de définition  $Y$  de la forme de Cartan, que nous noterons alors  $\varpi_Y$ . Comme cet espace est important nous lui donnerons un nom.

**DÉFINITION.** *Nous appellerons espace d'évolution du système défini par le lagrangien  $l$ , le plus grand ouvert  $Y$  de  $SQ$ , sur lequel est définie l'application de Legendre  $P_S$ , et donc la forme de Cartan  $\varpi_S$ .*

### III.4 La structure symplectique

Nous pouvons faire apparaître maintenant la structure symplectique promise, définie sur l'espace des mouvements du système. Nous renvoyons à l'annexe C pour une brève description des objets formels de la géométrie symplectique.

Nous considérerons le cas évoqué au paragraphe précédent d'une 2-forme  $d\varpi_Y$  présymplectique. Rappelons qu'alors une solution, un mouvement du système, est une courbe intégrale de son feuilletage caractéristique. L'espace des mouvements est donc l'ensemble de ces courbes intégrales et nous le noterons :

$$\mathcal{M} = Y / \ker d\varpi_Y. \quad (\text{III.38})$$

Cet espace de feuilles n'a pas toujours de bonnes propriétés topologiques, mais si c'est le cas : si  $\mathcal{M}$  est une variété telle que la projection associée

$$\pi : Y \rightarrow \mathcal{M} \quad (\text{III.39})$$

soit une submersion, alors  $\mathcal{M}$  est naturellement muni d'une unique 2-forme fermée  $\omega$  vérifiant :

$$\pi^*\omega = d\varpi_Y, \quad \ker \omega = \{0\}. \quad (\text{III.40})$$

Rappelons que (voir annexe C) :

DÉFINITION. On appelle forme symplectique toute 2-forme fermée non dégénérée, et variété symplectique toute variété munie d'une forme symplectique.

PROPOSITION. Soit  $Q$  une variété et  $l$  un lagrangien homogène de degré 1. Soit  $\varpi_Y$  la forme de Cartan associée, définie sur l'espace d'évolution  $Y \subset SQ$ . Si  $d\varpi_Y$  est présymplectique et si l'espace des solutions  $\mathcal{M} = Y / \ker d\varpi_Y$  est une variété, alors il existe sur  $\mathcal{M}$  une forme symplectique  $\omega$  définie par  $\pi^*\omega = d\varpi_Y$ .

DÉMONSTRATION. Pour définir  $\omega$  en un point  $m$  de  $\mathcal{M}$ , il suffit de donner la valeur qu'elle prend sur deux vecteurs tangents  $\delta m$  et  $\delta' m$ , soit :

$$\omega(\delta m, \delta' m) = d\varpi_Y(\delta y, \delta' y), \quad (\text{III.41})$$

où

$$\pi(y) = m, \pi_*(\delta y) = \delta m, \pi_*(\delta' y) = \delta' m. \quad (\text{III.42})$$

Il faut simplement vérifier que cette valeur ne dépend pas du représentant  $y$  de  $m$  ni des vecteurs tangents  $\delta y$  et  $\delta' y$ , pourvu que  $\pi_*(\delta y) = \delta m$  et  $\pi_*(\delta' y) = \delta' m$ . Nous avons supposé que la projection  $\pi$  est une submersion, ce qui implique en particulier que l'on peut trouver une section  $\varphi$  de la projection  $\pi$ , définie sur un voisinage  $U$  de  $m$ , qui envoie  $m$  sur  $y$ . L'image  $\varphi(U)$  de cette section est une sous-variété transverse au feuilletage, qui coupe chaque feuille qu'elle rencontre en un point et un seul. Le théorème de redressement du flot affirme qu'il existe alors un voisinage  $V \subset Y$  de  $y$ , difféomorphe au produit d'un intervalle  $] - \varepsilon, \varepsilon[$  par  $\varphi(U)$ , qui identifie les segments du type  $] - \varepsilon, \varepsilon[ \times \{y\}$  à l'intersection de la feuille passant par  $y$  avec  $V$ . Grâce à cette identification, le vecteur  $\delta y$  qui relève  $\delta m$  s'écrit  $\delta y = (\delta s, \delta m)$ , et de même  $\delta' y = (\delta' s, \delta' m)$ . Il est clair ainsi que  $d\varpi_Y(\delta y, \delta' y)$  ne dépend pas des composantes  $\delta s$  et  $\delta' s$  qui sont dans le noyau de  $d\varpi_Y$ . Il reste à vérifier l'indépendance par rapport au point  $y$  choisi, ce qui est encore une conséquence du théorème du redressement du flot. Soit  $z$  un autre point de la feuille  $m$ ,  $W$  et  $\psi$  l'analogue, pour le point  $z$ , de  $V$  et  $\varphi$ . Pour que  $\delta z$  relève  $\delta m$  il faut évidemment que  $\delta z = (*, \delta m)$ , où  $*$  est quelconque. ■

REMARQUE. Il faut à la fois apprécier cette structure qui nous est offerte avec les solutions du problème et relativiser son importance, c'est-à-dire comprendre

ce qu'elle représente du problème dont elle est issue. Il faut noter en particulier que ce n'est pas parce que  $\omega$  est la projection (image directe) d'une forme exacte qu'elle est elle-même exacte. Nous verrons quelques exemples du contraire. Mais il est vrai que localement, pour tout mouvement  $m$ , il existe un voisinage  $U$  de  $m$  sur lequel  $\omega$  est exacte; nous pouvons d'ailleurs prendre le même voisinage que précédemment sur lequel  $\omega = \varphi^*(d\varpi_Y)$ . Si nous pensons que le rôle d'une 2-forme est d'être intégrée sur une deux-chaîne, et si cette deux-chaîne est suffisamment petite pour être contenue dans un voisinage de trivialisatation comme  $U$  alors

$$\int_{\sigma} \omega = \int_{\varphi(\sigma)} d\varpi_Y = \int_{\varphi(\partial\sigma)} \varpi_Y.$$

Notons  $\gamma$  le contour d'intégration,  $\gamma = \partial\varphi(\sigma) = \varphi(\partial\sigma)$ ; alors :

$$\int_{\sigma} \omega = \int_{\gamma} \varpi_Y = \int_{\gamma} l = A(\gamma). \quad (\text{III.43})$$

Autrement dit : la *forme symplectique c'est l'action*; ou plutôt, c'est tout ce qu'il est possible de se rappeler de l'action originale du système lorsque l'on ne considère que l'espace de ses mouvements. Nous verrons plus loin comment cette remarque nous permet de démontrer facilement, par exemple, la formule de Crofton sur la mesure des droites qui coupent un domaine, et donc de l'étendre au cas général d'un problème variationnel.

Examinons maintenant sur quelques exemples ce que nous donne cette construction.

EXEMPLE. (LE POINT MATÉRIEL) Revenons au point matériel de masse  $m$ ,  $Q = \mathbf{R} \times \mathbf{R}^3$ . Dans les variables  $(t, r, v) \in \mathbf{R} \times \mathbf{R}^3 \times \mathbf{R}^3$ , le lagrangien  $L$  vaut :

$$L = \frac{mv^2}{2} - U. \quad (\text{III.44})$$

Il induit sur  $TQ$  le lagrangien :

$$l = \frac{m\dot{r}^2}{2\dot{t}} - U\dot{t}, \quad q = (t, r) \in Q, \quad \dot{q} = (\dot{t}, \dot{r}) \in T_qQ, \quad \dot{t} \neq 0. \quad (\text{III.45})$$

L'application  $(q, \dot{q}) \mapsto p$  est donnée par :

$$(q, \dot{q}) \mapsto p = \left[ -\frac{m\dot{r}^2}{2\dot{t}^2} - U \quad \frac{m\dot{r}}{\dot{t}} \right]. \quad (\text{III.46})$$

L'espace d'évolution est défini par :

$$Y = \{(q, \dot{q}) \mid \dot{t} \neq 0\} \sim \mathbf{R} \times T\mathbf{R}^3, \quad (\text{III.47})$$

et s'identifie avec  $\mathbf{R} \times T\mathbf{R}^3$  en faisant  $\dot{t} = 1$  ; la forme de Cartan devient :

$$\varpi_Y = mv \cdot dr - hdt \quad \text{avec} \quad h = \frac{mv^2}{2} + U. \quad (\text{III.48})$$

La fonction  $h$  est appelée *hamiltonien du système*. En se rappelant que  $\partial U / \partial r = -F$ , la dérivée extérieure  $d\varpi_Y$  vaut :

$$d\varpi_Y(\delta y, \delta' y) = \langle m\delta v - F\delta t, \delta' r - v\delta' t \rangle - \langle m\delta' v - F\delta' t, \delta r - v\delta t \rangle, \quad (\text{III.49})$$

où  $\delta y$  et  $\delta' y$  sont deux vecteurs tangents à  $Y$  au point  $y = (t, r, v)$ . Comme on peut le constater, on retrouve bien les équations de Newton comme équations du noyau de  $d\varpi_Y$  :

$$\delta y = \frac{dy}{ds} \in \ker d\varpi_Y \quad \Leftrightarrow \quad \delta r = v\delta t \quad \text{et} \quad m\delta v = F\delta t. \quad (\text{III.50})$$

Les courbes intégrales du feuilletage caractéristique de  $d\varpi_Y$ , ou ce qui est équivalent, leur projection  $y \mapsto (t, r)$  sur  $\mathbf{R} \times \mathbf{R}^4$ , sont les mouvements de la particule.

EXEMPLE. (LA VARIÉTÉ DES DROITES) L'exemple de la variété des droites est un cas particulier d'espace de géodésiques que nous verrons en toute généralité plus loin. Soit  $Q = \mathbf{R}^n$  et  $l$  le lagrangien homogène défini sur  $TQ \sim \mathbf{R}^n \times \mathbf{R}^n$  par :

$$l(q, \dot{q}) = \|\dot{q}\|. \quad (\text{III.51})$$

Le symbole double barre  $\|\|$  désigne la norme euclidienne usuelle. L'application de Legendre est donnée par :

$$P(q, \dot{q}) = (q, \bar{u}) \quad \text{avec} \quad u = \frac{\dot{q}}{\|\dot{q}\|}, \quad (\text{III.52})$$

où  $\bar{u}$  désigne le transposé du vecteur  $u$  pour la métrique euclidienne. Nous voyons que nous pouvons naturellement identifier l'espace  $SQ$  avec le produit direct  $\mathbf{R}^n \times S^{n-1}$ ,  $Y = SQ$ . La forme de Cartan s'écrit :

$$\varpi_Y = \bar{u}dq, \quad y = (q, u) \in Y. \quad (\text{III.53})$$

Sa dérivée extérieure est donnée par :

$$d\varpi_Y(\delta y, \delta' y) = \langle \delta u, \delta' q \rangle - \langle \delta' u, \delta q \rangle. \quad (\text{III.54})$$

Un vecteur  $\delta y$  tangent à  $Y$  au point  $y$  est dans le noyau de  $d\varpi_Y$  si :

$$\langle \delta u, \delta' q \rangle - \langle \delta' u, \delta q \rangle = 0 \quad \forall \delta' u \in T_u S^{n-1}, \forall \delta' q. \quad (\text{III.55})$$

Pour résoudre ce type d'équation avec contrainte, on fixe le vecteur tangent  $\delta y$  et on considère  $d\varpi(\delta y, \cdot)$  comme une 1-forme linéaire, que l'on peut noter  $d\varpi(\delta y)$ . Ensuite on remarque que  $\delta' u \in T_u S^{n-1}$  est équivalent à  $\bar{u}\delta' u = 0$ , l'équation précédente devenant alors :

$$\delta' u \in \ker \bar{u} \Rightarrow (\delta' q, \delta' u) \in \ker d\varpi(\delta y). \quad (\text{III.56})$$

On utilise maintenant le *théorème des multiplicateurs de Lagrange* :

**THÉORÈME.** (LAGRANGE) *Soit  $A : E \rightarrow F$  et  $B : E \rightarrow G$  deux opérateurs linéaires si  $\ker A \subset \ker B$  alors il existe  $C : F \rightarrow G$ , linéaire, tel que  $B = CA$ .*

Cela nous permet d'écrire :

$$\delta y = (\delta q, \delta u) \in \ker d\varpi \quad \Rightarrow \quad \exists \alpha \in \mathbf{R} : \delta q = \alpha u, \delta u = 0. \quad (\text{III.57})$$

Autrement dit, le feuilletage caractéristique de  $d\varpi$  est engendré par le champ de vecteurs  $\xi$  suivant :

$$\ker d\varpi = \mathbf{R}\xi \quad \text{avec} \quad \xi(q, u) = (u, 0). \quad (\text{III.58})$$

La courbe intégrale  $\Delta(q, u)$  passant par le point  $y = (q, u)$  est la droite affine de  $\mathbf{R}^n$ , de vecteur directeur  $u$  et passant par  $q$  :

$$\Delta(q, u) = \{(q + su, u) \in Y \mid s \in \mathbf{R}\}. \quad (\text{III.59})$$

L'espace  $\mathcal{M}$  des solutions du problème variationnel en question est donc l'espace des droites affines de  $\mathbf{R}^n$ . Pour écrire la forme symplectique, caractérisons la droite  $\Delta(q, u)$  par  $u$ , évidemment, puisqu'il est conservé, et par la projection orthogonale  $x$  de  $q$  sur  $u$ , autrement dit :

$$\Delta(u, q) \sim (u, x) \quad \text{avec} \quad x = [\mathbf{1} - u\bar{u}]q. \quad (\text{III.60})$$

Cela nous permet d'identifier  $\mathcal{M}$  avec le fibré tangent à la sphère  $S^{n-1}$  :

$$\mathcal{M} \simeq TS^{n-1}, \quad (\text{III.61})$$

qui peut aussi être considéré comme une sous-variété de  $Y$ , section de la projection canonique  $\pi : Y \rightarrow \mathcal{M}$ . Pour déterminer la forme symplectique, il suffit de restreindre la forme  $d\varpi$  à  $TS^{n-1}$  :

$$\omega(\delta m, \delta' m) = \langle \delta u, \delta' x \rangle - \langle \delta' u, \delta x \rangle \quad \text{avec} \quad m = (u, x) \in TS^{n-1}. \quad (\text{III.62})$$

Cet exemple est le plus simple des exemples d'espaces de géodésiques, mais nous en verrons d'autres par la suite.

EXEMPLE. (THÉORÈME DE CROFTON) Le théorème de Crofton (dans une version simplifiée) démontre l'existence d'une mesure  $\mu$ , unique à un facteur multiplicatif près, définie sur l'espace des droites du plan  $\mathbf{R}^2$ , invariante par le groupe des déplacements euclidiens  $x \mapsto Ax + b$ , où  $A \in O(2)$  et  $b \in \mathbf{R}^2$ . Cette situation est facile à généraliser à  $\mathbf{R}^n$ . On peut montrer ensuite (c'est le théorème proprement dit) que la mesure des droites qui coupent un domaine convexe compact  $\Omega$  est proportionnelle au périmètre du bord de  $\Omega$ . Nous allons voir comment ce théorème, à l'apparence mystérieuse, est une conséquence, presque triviale, de la nature symplectique de la variété des droites.

Les constructions précédentes nous indiquent qu'il existe, sur l'espace  $\mathcal{D}_n$  des droites orientées de  $\mathbf{R}^n$ , une forme symplectique naturelle  $\omega$ . Puisque la variété  $\mathcal{D}_n$  est de dimension  $2(n-1)$ , et que la forme  $\omega$  est non-dégénérée, la puissance extérieure  $\text{vol}_\omega = \omega^{\wedge(n-1)}$  est une forme volume de  $\mathcal{D}_n$ . Ce volume oriente donc naturellement la variété des droites de  $\mathbf{R}^n$ . Dans notre cas  $n = 2$ , le volume est la forme symplectique elle-même. Il faut noter à ce propos que la variété des droites (orientées!) du plan qui est, nous l'avons vu, l'espace tangent au cercle  $S^1$ , est difféomorphe au cylindre<sup>10</sup>  $S^1 \times \mathbf{R}$ . Considérons maintenant un domaine  $\Omega$  strictement convexe et compact du plan  $\mathbf{R}^2$ , bordé par une courbe fermée  $\Sigma = \partial\Omega$ . Soit  $\tilde{\Omega}$  l'ensemble des droites qui coupent  $\Omega$ ; c'est un domaine compact de la variété

<sup>10</sup>À ce propos, le lecteur peut vérifier que la variété des droites non-orientées est le ruban de Möbius, lui-même non-orientable.

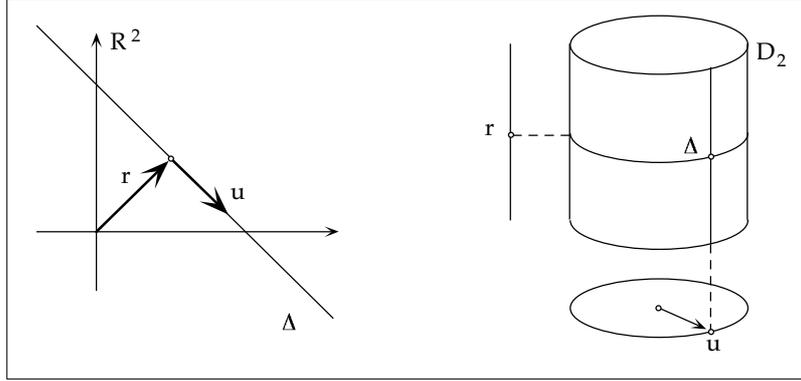


Figure III.1 : La variété des droites du plan

des droites  $\mathcal{D}_2$ . On peut en effet obtenir  $\tilde{\Omega}$  comme l'image, dans  $\mathcal{D}_2$ , du produit  $\Sigma \times \Sigma$  par l'application  $j_\Omega$  qui à deux points  $(x, y) \in \Sigma \times \Sigma$  associe la droite unique qui passe par ces deux points. Cette application est un plongement et n'est définie que sur le produit  $\Sigma \times \Sigma$  privé de la diagonale  $\Sigma$ . Son image est un cylindre dont le bord est constitué de deux composantes : les sous-variétés des droites tangentes à  $\Sigma$ , orientées dans un sens et dans l'autre. Ces deux bords sont obtenus comme les différentes limites,  $\lim_{y \rightarrow x} j_\Omega(y, x)$  et  $\lim_{y \rightarrow x} j_\Omega(x, y)$ , quand  $x$  parcourt  $\Sigma$ . La mesure des droites qui coupent le domaine  $\Omega$  est donc l'intégrale

$$\mu(\tilde{\Omega}) = \frac{1}{2} \int_{\tilde{\Omega}} \omega = \frac{1}{2} \int_{\tilde{\Omega}} d\lambda = \frac{1}{2} \int_{\partial \tilde{\Omega}} \lambda \quad (\text{III.63})$$

où  $\lambda$  est la primitive  $\bar{u}dx$  de  $\omega$ ,  $u \in S^1$  et  $x \in \mathbf{R}^2$  avec  $u \cdot x = 0$ . Le coefficient  $1/2$  vient de ce que nous ne considérons que les droites *modulo* orientation. Rappelons que le couple  $(u, x)$  représente la droite  $D = \{x + tu \mid t \in \mathbf{R}\}$ , orientée par  $u$ . Orientons la courbe  $\Sigma$ , c'est-à-dire choisissons un vecteur unitaire  $r \mapsto u$ , où  $r \in \Sigma$ . Le bord  $\partial \tilde{\Omega}$  du domaine des droites qui coupent  $\Omega$  est donc la réunion des images des deux applications :

$$r \mapsto (u, x = [\mathbf{1} - u\bar{u}]r) \text{ et } r \mapsto (-u, x = [\mathbf{1} - u\bar{u}]r). \quad (\text{III.64})$$

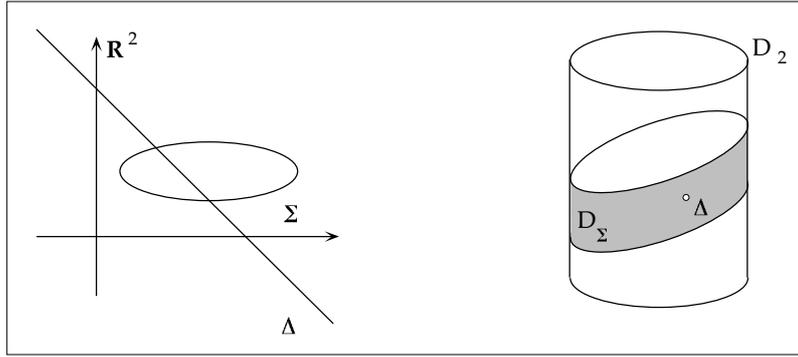


Figure III.2 : Le domaine des droites qui coupent un convexe

Ayant convenablement orienté le cylindre  $\mathcal{D}_2$ , il vient alors :

$$\mu(\tilde{\Omega}) = \int_{\Sigma} \bar{u} d([\mathbf{1} - u\bar{u}]r) = \int_{\Sigma} \bar{u} dr - \int_{\Sigma} \bar{u} d(u\bar{u}r); \quad (\text{III.65})$$

or :

$$\int_{\Sigma} \bar{u} d(u\bar{u}r) = \int_{\Sigma} (\bar{u}r)\bar{u} du + \int_{\Sigma} d(\bar{u}r), \quad (\text{III.66})$$

et  $\bar{u} du = 0$  car  $u$  est unitaire ; comme la courbe  $\Sigma$  est supposée fermée :

$$\int_{\Sigma} d(\bar{u}r) = \bar{u}r |_{\partial\Sigma} = 0. \quad (\text{III.67})$$

Il reste donc enfin :

$$\mu(\tilde{\Omega}) = \frac{1}{2} \int_{\tilde{\Omega}} \omega = \int_{\Sigma} \bar{u} dr = \text{Long}(\Sigma). \quad (\text{III.68})$$

C'est ce que nous voulions vérifier : la mesure des droites qui coupent un domaine strictement convexe compact à bord lisse est proportionnelle au périmètre de son bord. Ce qui précède est une spécialisation, dans ce cas de la variété des droites du plan, de l'argument général de la remarque de la page 60. Dans le cas général des droites de  $\mathbf{R}^n$ , on montre de la même manière que la mesure des droites qui coupent un domaine strictement convexe, compact à bord lisse est proportionnelle au volume de son bord. Les domaines non strictement convexes sont obtenus, comme d'habitude, par des procédés de limite.

EXERCICE. *Considérons un domaine non nécessairement convexe  $D$ , obtenu par plongement  $C^\infty$  du disque  $D^2 \subset \mathbf{R}^2$  dans  $\mathbf{R}^2$  affine. Soit  $\nu_D$  la fonction suivante, définie sur la variété des droites affines non orientées de  $\mathbf{R}^2$  :  $\nu_D(\Delta)$  est la moitié du nombre des points d'intersection de la droite affine  $\Delta \subset \mathbf{R}^2$  avec le bord  $\Sigma = \partial D$ , ce que l'on note aussi :*

$$\nu_D(\Delta) = \sharp(\Delta \cap \Sigma).$$

*En adaptant les méthodes du cas convexe, montrer que, sous certaines conditions techniques sur le bord  $\Sigma$ , l'intégrale de la fonction  $\nu_D$  sur l'espace des droites non orientées de  $\mathbf{R}^2$  donne encore la longueur du bord  $\Sigma$ .*

### III.5 Le point de vue hamiltonien

L'intérêt de la construction précédente d'homogénéisation du lagrangien (s'il n'est pas déjà homogène) réside en partie dans ce qu'elle introduit naturellement l'espace des directions tangentes, qui est essentiel dans la description du problème, sa résolution et l'introduction de la structure symplectique sur l'espace des solutions. Mais c'est une variété abstraite, obtenue par quotient, et difficile à manipuler concrètement. Il est plus agréable de travailler sur le fibré tangent. Une solution consisterait à choisir une section de la projection<sup>11</sup>  $\pi : TQ - Q \rightarrow SQ$ . Cela est possible en effet puisque la fibration principale  $\pi$  est à groupe contractile. On pourrait aussi munir la variété  $Q$  d'une métrique riemannienne auxiliaire et choisir le sous-fibré unitaire pour cette métrique. Mais cette méthode introduirait un objet extérieur au problème, ce qui n'est pas une bonne idée. Une autre méthode, plus naturelle, consiste à choisir un niveau du lagrangien  $l$ , par exemple, si la valeur 1 est atteinte :

$$\Sigma = l^{-1}(1). \tag{III.69}$$

Mais pour que  $\Sigma$  soit une bonne section de  $\pi$ , il faut s'assurer de deux choses :

1. Chaque demi-droite  $\alpha\dot{q}$ ,  $\alpha > 0$  et  $\dot{q} \neq 0$ , coupe  $\Sigma$  en un point et un seul.

---

<sup>11</sup>Dans ce paragraphe, nous supposons que le lagrangien  $l$  est défini sur tout  $TQ - Q$ , c'est-à-dire  $Y = SQ$ . Il est facile de l'adapter au cas où  $Y$  est seulement un ouvert de  $SQ$ .

2. L'hypersurface  $\Sigma$  est une sous-variété de  $TQ - Q$ , c'est-à-dire que son application linéaire tangente  $dl$  ne s'annule nulle part.

La première condition s'interprète simplement :

PROPOSITION. *Pour que  $\Sigma = l^{-1}(1)$  soit une section de la projection  $\pi : TQ - Q \rightarrow SQ$ , il faut et il suffit que le lagrangien  $l$  soit toujours strictement positif.*

DÉMONSTRATION. *Supposons  $l > 0$  ; soit  $v \in T_qQ - Q$  et  $\alpha = l(v)$  ; puisque  $\alpha > 0$  notons  $u = v/\alpha$  ; par homogénéité  $l(u) = 1$  et donc  $u \in \Sigma$ . Soit  $v'$  un autre point de la demi-droite  $\mathbf{R}_+^*v$ , alors  $v' = \beta v$ ,  $\beta > 0$  ; si  $l(v) = l(v')$  alors  $\beta = 1$ ,  $v = v'$ . Ainsi, si  $l > 0$ ,  $\Sigma$  coupe chaque demi-droite en un point et un seul. Réciproquement, soit  $v \in T_qQ - Q$ , soit  $u$  le point d'intersection de la demi-droite dirigée par  $v$  et de  $\Sigma$ , alors  $v = \alpha u$  avec  $\alpha > 0$  et donc  $l(v) = l(\alpha u) = \alpha l(u) = \alpha$  est strictement positif. ■*

Remarquons que si  $l$  est une fonction strictement positive, alors  $p = \partial l / \partial \dot{q}$  ne s'annule jamais, puisque  $l = p(\dot{q})$ . On en déduit donc  $dl \neq 0$  et l'on a la proposition suivante :

PROPOSITION. *Si  $\Sigma = l^{-1}(1)$  est une section de la projection  $\pi : TQ - Q \rightarrow SQ$ , alors c'est une sous-variété lisse de  $TQ - Q$ .*

Nous regrouperons les deux propositions précédentes en une seule :

PROPOSITION. *Pour que  $\Sigma = l^{-1}(1)$  soit une section lisse de la projection  $\pi : TQ - Q \rightarrow SQ$ , il faut et il suffit que le lagrangien  $l$  soit toujours strictement positif.*

Si cette condition de positivité sur le lagrangien  $l$  est réalisée, nous pouvons identifier  $SQ$ , c'est-à-dire  $Y$  à  $W$  ; c'est ce que nous ferons ici. En particulier, la forme différentielle  $\varpi_Y$  devient la restriction  $\varpi | Y$ , sa dérivée extérieure la restriction de sa dérivée extérieure, etc.

Mais que devient, dans ce cas, la condition d'immersion sur l'application de Legendre qui nous permet de munir l'espace des solutions du problème de sa structure symplectique ?

Notons d'abord que l'application de Legendre est une immersion si et seulement si sa restriction à chaque fibre  $Y_q = T_q Q \cap Y = l_q^{-1}(1)$  est une immersion. La fibre  $Y_q$  est une sous-variété d'un espace vectoriel ; il est donc légitime d'évoquer sa convexité :

**PROPOSITION.** *Soit  $l$  un lagrangien homogène strictement positif défini sur  $TQ - Q$ , et  $Y = l^{-1}(1)$ . La restriction de l'application de Legendre  $P$  à  $Y$  est une immersion si et seulement si chaque fibre  $Y_q = Y \cap T_q Q$ , au dessus de  $q \in Q$ , est strictement convexe, c'est alors un plongement.*

**DÉMONSTRATION.** *Notons  $E = T_q Q$ ,  $S = Y_q$ ,  $x$  un point courant de  $E$  et  $p = dl_q(x)$  la valeur de l'application de Legendre au point  $x$ . Considérons l'application  $H$  qui à  $x \in S$  associe l'hyperplan tangent  $h = \ker p = T_x S$ , orienté par le vecteur sortant  $x$ . Puisque l'application de Legendre est une immersion,  $H$  est un difféomorphisme local de  $S$  dans la variété des hyperplans orientés de  $E$ , c'est-à-dire dans la sphère  $S^{n-1}$ , où  $n = \dim Q$ . D'autre part, l'application  $H$  est surjective : par compacité de  $S$ , on peut trouver, pour tout hyperplan  $h$ , un hyperplan affine parallèle à  $h$  qui ne coupe pas  $S$  ; on le rapproche ensuite radialement de  $S$  : il existe un moment où il sera tangent à  $S$ . L'application  $H$  est donc un revêtement de  $S^{n-1}$ . Si  $n > 2$ ,  $S^{n-1}$  est simplement connexe,  $H$  est alors un difféomorphisme global et  $S$  est convexe. Pour  $n = 2$ , on remarque d'une part que ce que nous avons dit implique que  $S$  est localement convexe ; d'autre part,  $S$  étant radiale, le degré de  $H$  ne peut être qu'égal à 1, et donc  $H$  est encore un difféomorphisme global. ■*

Dans les conditions de la proposition précédente, tout ce qui a été dit sur  $Y$  peut être transposé sur l'image  $Y^* = P(Y)$  ; la forme de Cartan  $\varpi$  est alors remplacée par la forme de Liouville  $\lambda$  (voir la définition III.23) :

**PROPOSITION.** *Soit  $l$  un lagrangien homogène strictement positif, défini sur une variété  $Q$ . Si l'application de Legendre  $P$ , restreinte à  $Y$ , est une immersion alors les solutions du problème variationnel associé sont en bijection avec les caractéristiques, sur l'hypersurface  $Y^* = P(Y)$ , de la dérivée extérieure  $d\lambda$  de la forme de Liouville  $\lambda$  du cotangent  $T^*Q$ .*

REMARQUE. La variété  $Y^*$  peut être réalisée comme le niveau 1 d'une fonction  $l^*$  définie sur  $T^*Q$ . Nous pouvons d'ailleurs choisir pour  $l^*$  l'application réelle homogène de degré 1 définie sur  $T^*Q - Q$  qui vaut 1 sur  $Y^*$ . En voici une construction géométrique. De façon générale, soit  $E$  un espace vectoriel,  $l$  une fonction réelle, strictement positive, homogène de degré un définie sur  $E$  telle que sa différentielle restreinte à  $S = l^{-1}(1)$  soit une immersion. Associons à toute forme linéaire  $p \in E^*$  le point  $\dot{q}$  de  $E$  tel que  $T_{\dot{q}}S = \ker p$ . Ce point est unique par stricte convexité de  $S$ , posons alors :

$$l^*(p) = p(\dot{q}), \quad \ker p = T_{\dot{q}}(S). \quad (\text{III.70})$$

Cette fonction est évidemment positive, homogène de degré 1, et vaut 1 sur  $S$ . Il suffit maintenant de poser  $E = T_qQ$  et de faire varier  $q \in Q$  : la fonction  $l^*$  est bien homogène de degré 1 et vaut 1 sur  $Y^*$ . Nous pouvons faire de cette construction une définition :

DÉFINITION. *Nous appellerons transformée homogène de Legendre de la fonction  $l$  la fonction  $l^*$  définie par la formule précédente (III.70), c'est-à-dire :*

$$l^*(p) = p(\dot{q}), \quad \ker p = T_{\dot{q}}(S).$$

Cette transformée de Legendre n'est pas la transformée de Legendre ordinaire de  $l$ , qui est infinie dans le cas des fonctions homogènes de degré un, mais, comme nous l'avons vu, elle est peut être construite de manière analogue.

Cette construction est le point de vue dual du point de vue lagrangien dont nous avons parlé jusqu'à présent : on l'appelle *point de vue hamiltonien*. C'est un cas particulier de la construction générale suivante. Soit  $T^*Q$  l'espace cotangent de  $Q$ ,  $\lambda$  sa forme de Liouville, sa dérivée extérieure  $d\lambda$  est alors symplectique ; en effet, elle est évidemment fermée puisque exacte mais elle est aussi non dégénérée. Cela se vérifie directement, si  $(q, p)$  désigne un point courant de  $T^*Q$ ,  $\lambda = pdq$  et  $d\lambda$  s'exprime, dans une carte locale, de la manière suivante :

$$d\lambda = \sum_{i=1}^n dp_i \wedge dq_i, \quad (\text{III.71})$$

où  $n = \dim Q$  ;  $d\lambda$  est clairement non dégénérée. Considérons maintenant une variété symplectique quelconque  $(X, \omega)$  et une fonction réelle  $f$  définie sur  $X$ . Soit

$\Sigma$  une hypersurface de niveau  $\Sigma = f^{-1}(c)$  et supposons que  $\Sigma$  soit une sous-variété de  $X$ , c'est-à-dire  $df \neq 0$  en tout point de  $\Sigma$ . La restriction de  $\omega$  à  $\Sigma$  est dégénérée, ne serait-ce que parce que  $\dim \Sigma = 2n - 1$  est impaire. On peut exprimer plus précisément son noyau ; il faut résoudre l'équation :

$$\delta y \in \ker \omega | \Sigma \Leftrightarrow \omega(\delta y, \delta' y) = 0, \forall \delta' y \text{ tel que } df(\delta' y) = 0. \quad (\text{III.72})$$

En appliquant la méthode des multiplicateurs de Lagrange, on obtient :

$$\ker \omega | \Sigma = \mathbf{R} \operatorname{grad}_{\omega}(f) \quad \text{avec} \quad \operatorname{grad}_{\omega}(f) = \omega^{-1}(df). \quad (\text{III.73})$$

La notation  $\operatorname{grad}_{\omega}$  mérite peut-être quelque précision :  $\omega$  est une forme non dégénérée, que l'on peut interpréter, en tout point  $x \in X$ , comme un isomorphisme linéaire entre  $T_x X$  et  $T_x^* X$  : au vecteur  $v \in T_x X$  on associe la forme linéaire  $\omega_x(v, \cdot)$  ;  $\omega_x^{-1}$  est l'isomorphisme linéaire inverse. Le vecteur  $\operatorname{grad}_{\omega}(f)$  est appelé *gradient symplectique* de  $f$ .

Ainsi, le feuilletage caractéristique de  $f$  sur  $\Sigma$  est de dimension 1, et ses feuilles sont les courbes intégrales du champ de vecteur  $\operatorname{grad}_{\omega}(f)$ . C'est ce cadre général que nous avons retrouvé en construisant la sous-variété  $Y^* = P(Y)$ , associé au lagrangien  $l$ , et appliqué à  $X = T^*Q$  et  $f = l^*$ .

La construction que nous venons d'exposer n'est évidemment pas la plus générale qui nous permette d'identifier l'espace  $SQ$  quotient du tangent  $TQ$  avec une sous-variété  $W \subset TQ$ . Les conditions exigées sur le lagrangien  $l$  sont très particulières et pas toujours vérifiées. Si l'on prend par exemple l'homogénéisé d'un lagrangien  $L(t, x, v)$ , comme dans le paragraphe III.2,

$$l(q, \dot{q}) = L(t, x, \dot{x}/\dot{t})\dot{t}$$

où  $q = (t, x)$  et  $\dot{q} = (\dot{t}, \dot{x})$ , la condition de positivité de  $l$ , c'est-à-dire de  $L$  (car on prend  $\dot{t} > 0$ ), n'est généralement pas vérifiée : pour une particule dans un champ de forces  $L(t, x, v) = mv^2/2 + U(t, x)$  la positivité de  $L$  est équivalente à celle du potentiel  $U$ . Mais dans ce cas la sous-variété privilégiée qui réalise l'espace des directions, ou plutôt un ouvert de cet espace, est la sous-variété d'équation  $\dot{t} = 1$ . Comme cette situation est la plus générale des situations particulières, précisons-la davantage. Revenons avant l'homogénéisation du lagrangien : soit

$X$  une variété et un lagrangien  $L$  défini sur  $\mathbf{R} \times TX$ ,  $q = (t, x)$  et  $\dot{q} = (\dot{t}, \dot{x})$  les variables de l'homogénéisation, et  $l = L(t, x, \dot{x}/\dot{t})\dot{t}$  le lagrangien homogénéisé. Choisissons naturellement comme section de l'espace des directions tangentes  $SQ$  la nappe d'équation  $\dot{t} = 1$ ; l'application de Legendre s'exprime sans difficulté :

$$P(t, x, v) = \left[ l - \frac{\partial l}{\partial v}(v) \quad \frac{\partial l}{\partial v} \right], \quad (\text{III.74})$$

de telle sorte que la forme de Cartan s'écrive :

$$\varpi(\delta y) = p(\delta x) - h\delta t, \quad \text{avec} \quad p = \frac{\partial l}{\partial v} \quad \text{et} \quad h = \frac{\partial l}{\partial v}(v) - l. \quad (\text{III.75})$$

On reconnaît en  $h$  la transformée de Legendre du lagrangien  $L$  et la forme classique de la forme de Cartan du problème variationnel associé. Dans ce cas, si la 2-forme  $d\varpi$  est présymplectique, les sections à temps constant sont des cartes de l'espace des solutions. Si le problème est suffisamment régulier, ces cartes sont globales, autrement dit l'espace des solutions est alors isomorphe à l'espace tangent  $TX$ , muni de l'image réciproque de la dérivée extérieure de la forme de Liouville  $d\lambda$  par l'application  $(x, v) \mapsto p$ ; on peut écrire :

$$\omega = dp \wedge dx. \quad (\text{III.76})$$

C'était justement la situation de l'exemple de la particule.

---

## Quelques exemples

---

Le cas le plus simple de lagrangien homogène est celui défini par la norme associée à une métrique riemannienne positive : la variété  $Q$  est munie d'une métrique riemannienne  $g$  et  $l(q, \dot{q}) = \|\dot{q}\|$ . Les solutions d'un tel problème variationnel sont appelées *géodésiques* de la variété  $Q$  : ce sont les extrémals de la longueur. Dans le cas (rare) où l'espace des géodésiques est une variété, c'est donc une variété symplectique. Nous allons l'expliciter ici, pour deux exemples simples : les géodésiques de la sphère  $S^2$  et celles du disque de Poincaré  $H^2$ . Nous avons déjà vu au paragraphe précédent comment la variété des droites affines de l'espace  $\mathbf{R}^2$ , ses géodésiques, était munie d'une structure de variété symplectique. Ces trois exemples : le plan, la sphère et le disque réalisent les trois cas de courbure constante : nulle, positive et négative.

### IV.1 Les géodésiques de la sphère

L'exemple bien connu<sup>1</sup> des géodésiques de la sphère nous servira ici d'illustration à la méthode générale que nous venons de présenter. La sphère  $S^2$  est la sous-variété des vecteurs unitaires de  $\mathbf{R}^3$  :

$$S^2 = \{x \in \mathbf{R}^3 \mid \|x\| = 1\}. \quad (\text{IV.1})$$

---

<sup>1</sup>Si nous marchons sur la Terre toujours *tout droit* dans une direction donnée, nous passerons inévitablement par les antipodes de notre point de départ, et donc par les antipodes de chacun des points par lesquels nous passons.

en un point  $x$ , l'espace tangent s'identifie naturellement avec le sous-espace vectoriel de  $\mathbf{R}^3$ , orthogonal à  $x$  :

$$T_x S^2 \simeq \{v \in \mathbf{R}^3 \mid \langle x, v \rangle = 0\}. \quad (\text{IV.2})$$

La sphère  $S^2$  est munie de sa métrique usuelle, induite par la métrique de  $\mathbf{R}^3$ . Autrement dit, pour tout couple de vecteurs  $v, v'$  orthogonaux à  $x$ , on pose simplement :

$$g_x(v, v') = \langle v, v' \rangle. \quad (\text{IV.3})$$

Le lagrangien est donné par :

$$l(x, v) = \|v\|. \quad (\text{IV.4})$$

On en déduit l'application de Legendre  $P$  :

$$P : (x, v) \mapsto \left( x, \frac{\bar{v}}{\|v\|} \right), \quad (\text{IV.5})$$

où la barre désigne le vecteur transposé. La sous-variété  $Y$  d'équation  $l = 1$  est le fibré unitaire tangent à la sphère  $S^2$ , c'est-à-dire :

$$\begin{aligned} Y &= \{(x, u) \in TS^2 \mid \|v\| = 1\} \\ &= \{(x, u) \in \mathbf{R}^3 \times \mathbf{R}^3 \mid \langle x, x \rangle = \langle u, u \rangle = 1 \text{ et } \langle x, u \rangle = 0\}, \end{aligned} \quad (\text{IV.6})$$

et la forme de Cartan est simplement :

$$\varpi(\delta y) = \langle u, \delta x \rangle, \quad \text{avec } y = (x, u). \quad (\text{IV.7})$$

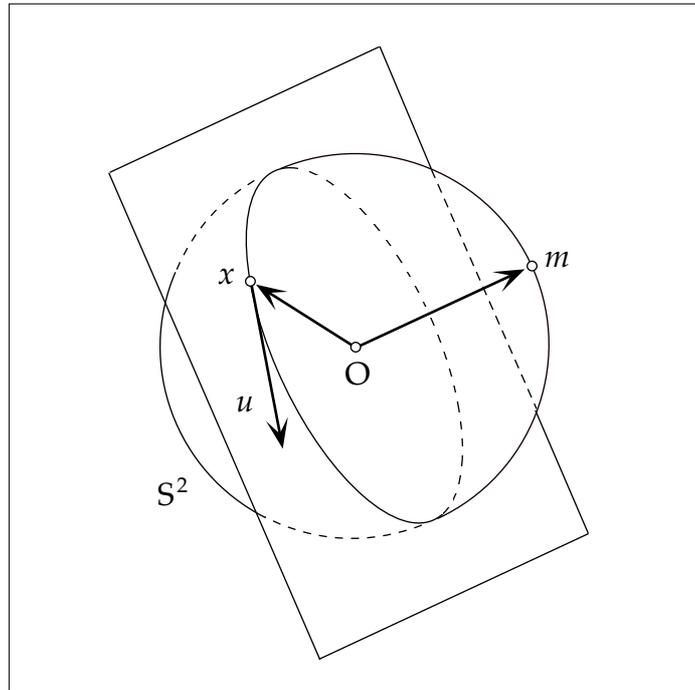
La forme présymplectique  $d\varpi$  définie sur  $Y$ , dont on doit calculer le noyau, s'écrit :

$$d\varpi(\delta y, \delta' y) = \langle \delta u, \delta' x \rangle - \langle \delta' u, \delta x \rangle. \quad (\text{IV.8})$$

En utilisant la méthode des multiplicateurs de Lagrange exposée plus haut, on obtient les équations du noyau :

$$\frac{dy}{ds} \in \ker d\varpi \quad \Leftrightarrow \quad \frac{dx}{ds} = \alpha u, \quad \frac{du}{ds} = -\alpha x, \quad \alpha \in \mathbf{R}. \quad (\text{IV.9})$$

Sans méthode générale pour résoudre cette équation, on peut remarquer que le

Figure IV.1 : Les géodésiques de la sphère  $S^2$ 

plan  $m$  engendré par  $x$  et  $u$  est constant le long de chaque solution. En effet, le plan  $m$  étant entièrement caractérisé par le produit vectoriel  $x \wedge u$ , on peut écrire directement :

$$m = x \wedge u \quad \text{et} \quad \frac{dm}{ds} = 0. \quad (\text{IV.10})$$

La solution est donc entièrement contenue dans l'intersection de  $S^2$  avec  $m$ . Comme cette intersection est une courbe et que nous cherchons une courbe, cela ne peut être que celle-là. Autrement dit,  $m$  représente la solution de condition initiale  $(x, u)$ . Réciproquement, tout plan de  $\mathbf{R}^3$  définit une solution ; donc l'espace des géodésiques de la sphère  $S^2$  est l'espace des plans orientés de  $\mathbf{R}^3$ . Il faut noter à ce propos que l'involution  $(x, u) \mapsto (x, -u)$  ne préserve pas la 2-forme  $d\varpi$ , mais l'inverse. La solution de condition initiale  $(x, u)$  est donc différente de la solution

$(x, -u)$ . Il est important de préciser que, comme courbe de sur  $S^2$ , la solution de condition initiale  $(x, u)$  est orientée par l'orientation du plan  $m$ .

Autrement dit, la bonne réalisation de la variété des géodésiques (orientées) de  $S^2$  est celle que nous avons donnée plus haut, définie par l'application :

$$\mu : Y = US^2 \rightarrow \mathbf{R}^3, \quad (x, u) \mapsto m = x \wedge u. \quad (\text{IV.11})$$

La variété des géodésiques de  $S^2$  est l'image par  $\mu$  de  $Y$ , c'est-à-dire :

$$\mathcal{G}_{S^2} = \mu(US^2) = S^2. \quad (\text{IV.12})$$

Il reste à exprimer la forme symplectique associée. C'est un exercice d'algèbre linéaire de montrer que :

$$\omega_m(\delta m, \delta' m) = -\langle m, \delta m \wedge \delta' m \rangle = \det(m, \delta m, \delta' m). \quad (\text{IV.13})$$

C'est l'élément de surface de la sphère  $S^2$ . Autrement dit, la variété des géodésiques sur la sphère  $S^2$  munie du produit scalaire ordinaire est la sphère unité de  $\mathbf{R}^3$  munie de la restriction de la forme d'aire canonique (au signe près).

EXERCICE. *Que se passe-t-il pour une métrique homothétique  $ag$ ,  $a > 0$  ? Vérifier le signe de la formule précédente (IV.13).*

EXERCICE. *Traiter le cas des géodésiques de la sphère unité de  $\mathbf{R}^n$ .*

## IV.2 Les géodésiques du disque de Poincaré

L'exemple des géodésiques du disque de Poincaré est moins intuitif que celui de la sphère  $S^2$  et l'intuition nous faisant défaut, la méthode générale est ici notre seul guide.

Nous allons donner maintenant l'un des innombrables modèles de l'espace hyperbolique de dimension 2 et la construction de ses géodésiques. Cet espace a  $\text{PSL}(2, \mathbf{R})$  comme groupe d'isométries directes. C'est le groupe des matrices réelles  $2 \times 2$  de déterminant  $+1$  modulo  $\mathbf{Z}/2\mathbf{Z}$  :

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ tel que } ad - bc = +1 \text{ et } A \sim -A. \quad (\text{IV.14})$$

Nous allons partir directement de ce groupe pour reconstruire à la fois l'espace hyperbolique et sa variété des géodésiques.

L'algèbre de Lie de  $\mathrm{SL}(2, \mathbf{R})$ , c'est-à-dire son espace tangent en l'identité (dans l'espace vectoriel des matrices  $2 \times 2$ ), est le sous-espace vectoriel des matrices  $2 \times 2$  de trace nulle, noté  $\mathrm{sl}(2, \mathbf{R})$ . Tout élément  $X$  de  $\mathrm{sl}(2, \mathbf{R})$  s'écrit donc :

$$X = \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix}. \quad (\text{IV.15})$$

On définit sur  $\mathrm{sl}(2, \mathbf{R})$  la forme bilinéaire suivante :

$$\langle X, X' \rangle = \frac{1}{2} \mathrm{Tr}(XX') = \alpha\alpha' + \frac{\beta\gamma' + \gamma\beta'}{2}. \quad (\text{IV.16})$$

Il est facile de vérifier qu'elle est non dégénérée et de signature  $++-$ . Il suffit par exemple d'identifier la matrice  $X$  avec le vecteur, noté par la même lettre,  $X = (x, y, z)$  défini par :

$$x = \alpha \quad y = \frac{\beta + \gamma}{2} \quad z = \frac{\beta - \gamma}{2}. \quad (\text{IV.17})$$

La réciproque de cette correspondance est évidemment donnée par :

$$\alpha = x \quad \beta = y + z \quad \gamma = y - z. \quad (\text{IV.18})$$

La forme bilinéaire définie plus haut devient :

$$\langle X, X' \rangle = xx' + yy' - zz'. \quad (\text{IV.19})$$

On dit que c'est un pseudo-produit scalaire, encore appelé « produit scalaire lorentzien ». Par analogie avec avec le cas euclidien, on note  $\|X\|^2 = \langle X, X \rangle$  bien que cette forme ne soit pas définie positive.

La sous-variété de  $\mathrm{sl}(2, \mathbf{R})$  d'équation  $\|X\|^2 = -1$ , c'est-à-dire  $x^2 + y^2 - z^2 = -1$ , est un hyperboloïde de révolution (autour de  $oz$ ) à deux nappes, asymptote au cône  $C$  d'équation  $z^2 = x^2 + y^2$ . Nous noterons  $H^2$  la nappe contenant la matrice

$$X_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \sim \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad (\text{IV.20})$$

représentée par le vecteur  $(0, 0, 1)$ .

La forme bilinéaire restreinte à  $H^2$  est alors définie positive. On peut le vérifier directement ou encore utiliser l'action adjointe de  $\mathrm{SL}(2, \mathbf{R})$ , c'est-à-dire

$$(A, X) \mapsto AXA^{-1}. \quad (\text{IV.21})$$

Il est clair que le produit pseudo-scalaire est invariant par cette action et donc que  $H^2$  est stable. Par un argument sur la dimension, il est possible de montrer que  $H^2$  est homogène sous l'action de  $\mathrm{SL}(2, \mathbf{R})$ , autrement dit que pour tout couple de points  $X, X'$  dans  $H^2$  il existe une matrice  $A \in \mathrm{SL}(2, \mathbf{R})$  telle que  $AX = X'$ . Il suffit alors de vérifier que le produit pseudo-scalaire est positif sur l'espace tangent à  $H^2$  au point  $X_0$ . Soit  $UH^2$  le fibré unitaire tangent à  $H^2$  :

$$UH^2 = \{(X, U) \mid X \in H^2 \text{ et } \langle X, U \rangle = 0\}. \quad (\text{IV.22})$$

De la même façon que pour la sphère, la forme de Cartan s'écrit :

$$\varpi(\delta Z) = \langle U, \delta X \rangle = \frac{1}{2} \mathrm{Tr}(U\delta X), \quad \text{avec } Z = (X, U). \quad (\text{IV.23})$$

On en déduit l'expression de la forme présymplectique :

$$\omega(\delta Z, \delta' Z) = \frac{1}{2} \mathrm{Tr}(\delta U \delta' X - \delta' U \delta X). \quad (\text{IV.24})$$

En utilisant la même méthode que pour la sphère, on obtient les équations différentielles des géodésiques de cette métrique riemannienne sur  $H^2$  :

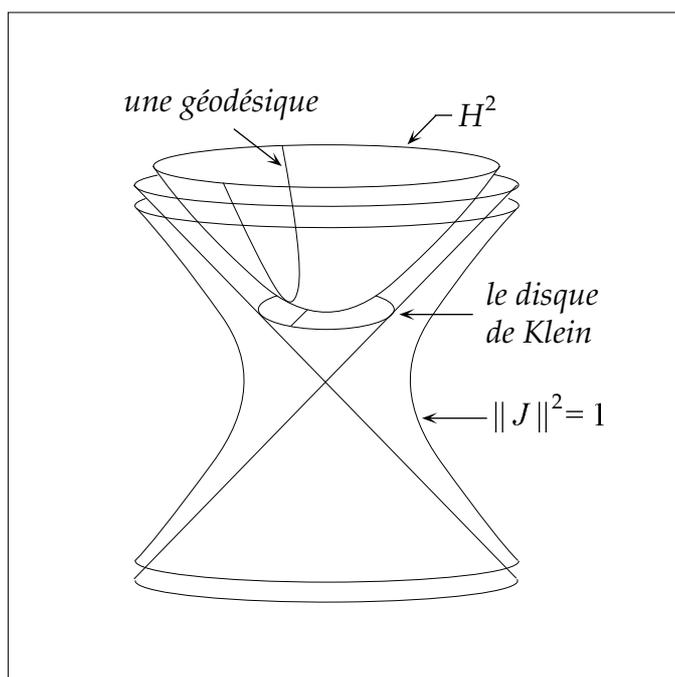
$$\frac{dX}{ds} = \alpha U \quad \text{et} \quad \frac{dU}{ds} = \alpha X, \quad \alpha > 0. \quad (\text{IV.25})$$

On peut remarquer sur ces équations que la courbure de  $H^2$  est égale à  $-1$ . Considérons alors la matrice  $J \in \mathfrak{sl}(2, \mathbf{R})$  définie par :

$$J = \frac{1}{2}[U, X] = \frac{1}{2}(UX - XU); \quad (\text{IV.26})$$

elle est conservée le long des géodésiques ; en effet

$$\frac{dJ}{ds} \propto \left[ \frac{dU}{ds}, X \right] + \left[ U, \frac{dX}{ds} \right] \propto [X, X] + [U, U] = 0. \quad (\text{IV.27})$$

Figure IV.2 : Les géodésiques de l'espace hyperbolique  $H^2$ .

D'autre part,  $J$  est à la fois orthogonal (pour le produit pseudo-scalaire) à  $X$  et à  $U$  (utiliser les propriétés de la trace). La trajectoire géodésique appartient donc à l'intersection de  $H^2$  et du plan orthogonal à  $J$ , et comme cette intersection est une courbe, c'est donc la géodésique elle-même. Ainsi les géodésiques de  $H^2$  sont obtenues comme les intersections de  $H^2$  avec les plans passant par l'origine.

Notons que l'application  $(X, U) \mapsto J$  nous donne un moyen de réaliser l'espace des géodésiques de  $H^2$  comme une sous-variété de  $\mathfrak{sl}(2, \mathbf{R})$ . En effet, on peut vérifier<sup>2</sup> que  $\|J\|^2 = 1$  : il suffit de se placer au point  $X_0$ . D'autre part, tout plan orthogonal à un  $J$  tel que  $\|J\|^2 = 1$  coupe  $H^2$  (on résoud l'équation  $J = [U, X]$ ). L'espace des géodésiques de  $H^2$  s'identifie donc avec la sous-variété

<sup>2</sup>On peut faire le calcul général ou bien utiliser la transitivité de  $SL(2, \mathbf{R})$  sur  $H^2$  et se placer au point  $X_0$ .

des matrices  $J$  de  $\mathfrak{sl}(2, \mathbf{R})$  d'équation  $\|J\|^2 = 1$ , c'est-à-dire  $x^2 + y^2 - z^2 = 1$ . C'est un hyperboloïde de révolution à une nappe, asymptote au cône  $C$ , dont les traces avec les plans d'équation  $z = c$  sont des cercles de rayon  $\sqrt{1 + c^2}$ . C'est un autre espace homogène (orbite adjointe) de  $\mathrm{SL}(2, \mathbf{R})$ . La variété des géodésiques de  $H^2$  est donc difféomorphe au cylindre  $S^1 \times \mathbf{R}$ .

On peut retrouver cette identification avec le cylindre en considérant la *représentation de Klein* qui consiste à projeter  $H^2$  sur le disque  $K^2$  d'équations  $z = 1$  et  $x^2 + y^2 < 1$  suivant les droites passant par l'origine et intérieures au cône  $C$ . Les géodésiques sont alors des segments de droites, traces sur  $K^2$  des plans passant par l'origine. L'espace des géodésiques est donc difféomorphe à la variété des droites, c'est-à-dire au cylindre

$$\mathcal{G}_{H^2} = \{J \in \mathbf{R}^3 \mid \|J\|^2 = 1\} \simeq TS^1. \quad (\text{IV.28})$$

Le lecteur perspicace aura remarqué sans doute l'analogie entre cette construction et celle de la variété des géodésiques de la sphère  $S^2$ , notamment dans l'introduction du vecteur  $J$ , équivalent de  $m$  dans le cas  $S^2$ . En effet, dans ce cas la fonction  $(x, u) \mapsto m$  est le *moment cinétique* relatif au groupe des isométries  $\mathrm{SO}(3)$  pour la sphère  $S^2$ , et  $(X, U) \mapsto J$  est l'équivalent du moment cinétique du groupe  $\mathrm{SL}(2, \mathbf{R})$  pour  $H^2$  — en réalité, si  $L$  désigne le *moment* associé à l'action de  $\mathrm{SL}(2, \mathbf{R})$  sur  $H^2$ , alors  $L = [X, U] = -2J$ . Dans ces deux cas c'est l'abondance de symétries qui rend la construction de l'espace des géodésiques aussi simple (et algébrique).

EXERCICE. *Montrer que la forme symplectique s'écrit :*

$$\omega_J(\delta J, \delta' J) = \frac{1}{2} \langle J, [\delta J, \delta' J] \rangle = \frac{1}{4} \mathrm{Tr}(J[\delta J, \delta' J]). \quad (\text{IV.29})$$

*Utiliser le fait que  $(U, J, X)$  est une base « orthonormée » de  $\mathfrak{sl}(2, \mathbf{R})$  et décomposer les vecteurs dans cette base. Utiliser cette méthode pour les géodésiques de la sphère  $S^2$ , en identifiant  $\mathbf{R}^3$  avec l'algèbre de Lie de  $\mathrm{SO}(3)$ , ensemble des matrices réelles  $3 \times 3$  antisymétriques, grâce au produit vectoriel :*

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto j(X) = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} \quad (\text{IV.30})$$

### IV.3 Les mouvements du solide

Bien que l'exemple des mouvements d'un solide dans  $\mathbf{R}^3$  ne soit pas du même genre que les précédents, il n'est toutefois pas sans rapport. C'est, de plus, le terrain privilégié de la théorie symplectique des systèmes complètement intégrables. Comme on ne trouve pas toujours les développements qui aboutissent aux équations du mouvement des solides dans les ouvrages classiques, nous les détaillons ici.

Définissons d'abord ce qu'est un solide. Intuitivement, c'est un ensemble de points  $r = (r_1, \dots, r_N)$  dont les distances mutuelles sont fixes :

$$\|r_i - r_j\| = cst, \quad \forall i, j = 1 \dots N. \quad (\text{IV.31})$$

Chacun de ces points est soumis à une force  $F_i = F_i(t, r_1, \dots, r_N)$  et vérifie donc les équations de Newton :

$$\frac{dr_i}{dt} = v_i, \quad m_i \frac{dv_i}{dt} = F_i, \quad F_i = -\frac{\partial U}{\partial r_i}, \quad (\text{IV.32})$$

mais en respectant la contrainte solide. Comment traiter cette contrainte, c'est ce que nous allons voir.

La construction que nous avons donnée à partir du principe des travaux virtuels, qui conduit à transformer cette famille d'équations en un problème variationnel, est encore valable, avec comme lagrangien :

$$L(t, r, v) = \sum_{i=1}^N \frac{m_i v_i^2}{2} - U, \quad (\text{IV.33})$$

où

$$r = (r_1, \dots, r_N) \quad \text{et} \quad v = (v_1, \dots, v_N) \in \mathbf{R}^{3N}. \quad (\text{IV.34})$$

Nous cherchons toujours les solutions du problème :

$$\delta a = 0, \quad \text{avec} \quad a = \int_a^b L(t, r(t), v(t)) dt, \quad (\text{IV.35})$$

mais avec les contraintes suivantes sur les variations :

$$\delta \|r_i - r_j\| = 0. \quad (\text{IV.36})$$

Si l'on veut éviter maintenant des calculs pénibles et quasiment inextricables, il faut remarquer que la condition de rigidité signifie qu'à chaque instant il existe une isométrie  $g_t$  de l'espace euclidien  $\mathbf{R}^3$  telle que :

$$r_i(t) = g_t(r_i), \quad \forall i = 1, \dots, N. \quad (\text{IV.37})$$

Rappelons que le groupe des isométries de  $\mathbf{R}^3$ , que nous noterons  $E$ , est appelé *groupe des déplacements euclidiens*, qu'un élément  $g \in E$  est défini par une rotation  $A \in \text{SO}(3)$  et un vecteur  $T \in \mathbf{R}^3$ , et que son action sur  $\mathbf{R}^3$  est donnée par :

$$g = (A, T) \in E, \quad r \in \mathbf{R}^3 \quad g(r) = Ar + T. \quad (\text{IV.38})$$

En d'autres termes, la contrainte solide signifie simplement que le point

$$r = (r_1, \dots, r_N)$$

évolue sur une orbite du groupe  $E$  agissant sur  $\mathbf{R}^{3N}$ .

Si le nombre de points est suffisant et s'ils ne sont pas tous alignés, alors le déplacement  $g_t$  qui envoie les  $r_i$  sur  $r_i(t)$  est unique. Autrement dit, à tout chemin  $t \mapsto r = (r_1, \dots, r_N)$  dans  $\mathbf{R}^{3N}$  est associé un seul chemin  $t \mapsto g(t) = g_t$  dans  $E$ . De plus ce chemin est différentiable comme l'est  $t \mapsto r$ .

Toute variation  $\delta r$  compatible avec les contraintes est obtenue par une variation quelconque du chemin  $t \mapsto g_t$  :

$$\delta r_i(t) = \delta g(t)(r_i). \quad (\text{IV.39})$$

En choisissant une origine  $r$  de l'orbite, et puisque

$$v(t) = \frac{dg(t)}{dt}(r), \quad (\text{IV.40})$$

on peut écrire l'action  $a$  directement sur le groupe  $E$

$$a = \int_a^b \tilde{L}(t, g(t), \nu(t)) dt, \quad (\text{IV.41})$$

où  $\tilde{L}$  est l'image réciproque de  $L$  par l'application définie sur  $\mathbf{R} \times TE$  à valeurs dans  $\mathbf{R} \times T\mathbf{R}^{3N}$  :

$$\hat{r} : (t, g, \nu) \mapsto (t, g(r), \nu(r)), \quad g \in E, \quad \nu \in T_g E, \quad (\text{IV.42})$$

c'est-à-dire :

$$\tilde{L}(t, g, \nu) = L(t, g(r), \nu(r)). \quad (\text{IV.43})$$

Dans le cas qui nous préoccupe le vecteur  $\nu$  est défini par deux vecteurs  $\alpha \in T_A \text{SO}(3)$  et  $\tau \in \mathbf{R}^3$  de telle sorte que :

$$\nu = (\alpha, \tau), \quad r_i \in \mathbf{R}^3 \quad \Rightarrow \quad \nu(r_i) = \alpha r_i + \tau. \quad (\text{IV.44})$$

Nous sommes donc ramenés à la situation générale d'un problème variationnel avec, comme variété de configuration, le groupe de Lie des déplacements euclidiens de l'espace  $\mathbf{R}^3$ . Nous avons donc à notre disposition toutes les constructions qui ont été décrites jusqu'à présent, en particulier l'expression de la forme de Cartan  $\varpi$  sous sa forme non homogène (III.75) :

$$\varpi(\delta y) = \tilde{p}(\delta g) - \tilde{h}\delta t, \quad \tilde{y} = (t, g, \nu), \quad (\text{IV.45})$$

où :

$$\tilde{p} = d\tilde{L}_g \in T_g^* E, \quad \text{et} \quad \tilde{h}(t, g, \nu) = h(t, g(r), \nu(r)). \quad (\text{IV.46})$$

On obtient immédiatement la valeur de  $\tilde{L}$  :

$$\tilde{L}(t, g, \nu) = \frac{1}{2} \text{Tr} \alpha J \bar{\alpha} + \langle \alpha c, \tau \rangle + \frac{m}{2} \|\tau\|^2 - \tilde{U}, \quad (\text{IV.47})$$

la matrice  $J$  étant la « matrice des moments » :

$$J = \sum_{i=1}^N m_i r_i \bar{r}_i, \quad (\text{IV.48})$$

le vecteur  $c$  et le scalaire  $m$

$$c = \sum m_i r_i, \quad m = \sum_{i=1}^N m_i, \quad (\text{IV.49})$$

étant respectivement le centre de gravité du solide et sa masse totale. Enfin,  $\tilde{U}(t, g) = U(t, g(r))$  est le potentiel des forces résultantes sur le solide.

La forme linéaire  $p$  a deux composantes, que nous noterons  $p_{\text{SO}(3)}$  et  $p_{\mathbf{R}^3}$ , qui ont pour valeur :

$$p_{\text{SO}(3)}(\delta A) = \text{Tr}[(J\bar{\alpha} + \tau\bar{c})\delta A] \quad \text{et} \quad p_{\mathbf{R}^3}(\delta\tau) = \langle m\tau + \alpha c, \delta\tau \rangle. \quad (\text{IV.50})$$

La construction de la forme de Cartan dans le cas général des solides libres n'est pas difficile mais fastidieuse, et nous nous contenterons, à partir de maintenant, du cas des mouvements du solides autour d'un point fixe. Le groupe qui intervient alors n'est plus le groupe des déplacements euclidiens mais son sous-groupe  $SO(3)$ . Les termes dépendants de  $\tau$  disparaissent et il reste :

$$\tilde{L}(t, g, \nu) = \frac{1}{2} \text{Tr} \alpha J \bar{\alpha} - \tilde{U}. \quad (\text{IV.51})$$

La matrice  $J$  est définie positive, et définit donc une métrique riemannienne sur  $SO(3)$ ; notons :

$$\langle \alpha, \alpha' \rangle_J = \text{Tr} \alpha J \bar{\alpha}', \quad \tilde{L} = \frac{1}{2} \|\alpha\|_J^2 - \tilde{U}. \quad (\text{IV.52})$$

La forme de Cartan devient :

$$\varpi(\delta y) = \langle \alpha, \delta A \rangle_J - \tilde{h} \delta t, \quad \text{avec} \quad \tilde{h} = \frac{1}{2} \|\alpha\|_J^2 + \tilde{U}, \quad y = (t, A, \alpha). \quad (\text{IV.53})$$

Notons alors :

$$\tilde{F} = -\text{grad}_J \tilde{U} \in T_A SO(3), \quad (\text{IV.54})$$

où  $\text{grad}_J$  désigne le gradient au sens de la norme définie par  $J$ ,  $\tilde{F}$  étant la force induite sur  $SO(3)$  par le potentiel  $\tilde{U}$ . La dérivée extérieure de  $\varpi$  s'écrit simplement :

$$d\varpi(\delta y, \delta' y) = \left\langle \delta A - \alpha \delta t, \delta' \alpha - \tilde{F} \delta' t \right\rangle_J - \left\langle \delta' A - \alpha \delta' t, \delta \alpha - \tilde{F} \delta t \right\rangle_J. \quad (\text{IV.55})$$

On peut déjà remarquer sur cette expression que si la force exercée sur le solide est nulle (donc  $\tilde{F} = 0$ ), les équations que l'on obtient sont celles des géodésiques sur le groupe de  $SO(3)$  pour la métrique définie par  $J$ . En particulier, si  $J = \mathbf{1}$  les mouvements solides sont les groupes à un paramètre de  $SO(3)$ . Autrement dit, les mouvements d'un solide s'interprètent comme les mouvements d'un point  $A$  dans l'espace des matrices  $M(\mathbf{R}^3)$ , muni de la métrique définie par  $J$ , soumis à une force  $\tilde{F}$ , et astreint à se mouvoir sur la sous-variété  $SO(3) \subset M(\mathbf{R}^3)$ .

Notons encore que si le terme de force est suffisamment régulier, la variété symplectique des mouvements solides est le tangent du groupe  $SO(3)$  muni de la forme symplectique image réciproque de la forme de Liouville par la métrique  $J$ . Il suffit en effet de faire  $t = cst$  et on obtient :

$$\omega(\delta x, \delta' x) = \langle \delta A, \delta' \alpha \rangle_J - \langle \delta' A, \delta \alpha \rangle_J, \quad x = (A, \alpha) \in T SO(3). \quad (\text{IV.56})$$

Le traitement des problèmes de ce type est généralement difficile du fait de la métrique et du plongement de la sous-variété de configuration. Mais dans le cas d'un groupe, comme ici, on peut trivialisier l'espace tangent ;  $T\text{SO}(3)$  est isomorphe au produit  $\text{SO}(3) \times \text{so}(3)$  grâce à la forme de Maurer-Cartan définie par :

$$\theta : T_A \text{SO}(3) \rightarrow \text{so}(3), \theta(A, \alpha) = (A, Z = A^{-1}\alpha). \quad (\text{IV.57})$$

Dans cette trivialisisation, en notant  $x = (a, Z) \in \text{SO}(3) \times \text{so}(3)$  et  $\delta x$  un vecteur tangent à  $X$ , la forme de Cartan devient :

$$\varpi(\delta x) = \text{Tr}A^{-1}\delta A J \bar{Z} - \left(\frac{1}{2} \text{Tr}Z J \bar{Z} + \tilde{U}\right) \delta t, \quad (\text{IV.58})$$

et sa dérivée extérieure :

$$\begin{aligned} d\varpi(\delta x, \delta' x) &= \text{Tr}(\xi' - Z\delta' t)J(\delta \bar{Z} - \bar{f}\delta t) \\ &- \text{Tr}(\xi - Z\delta t)J(\delta' Z - \bar{f}\delta' t) \\ &+ \text{Tr}[\xi', \xi]J\bar{Z}, \end{aligned} \quad (\text{IV.59})$$

où on a défini :

$$\xi = A^{-1}\delta A, \xi' = A^{-1}\delta' A \in \text{so}(3), \quad f = A^{-1}\tilde{F}. \quad (\text{IV.60})$$

On identifie enfin l'algèbre de Lie  $\text{so}(3)$  avec  $\mathbf{R}^3$  grâce à l'opérateur  $j$  (définition IV.30), et on obtient :

$$\begin{aligned} d\varpi(\delta x, \delta' x) &= \langle \Omega' - \zeta\delta' t, \delta\zeta - \phi\delta t \rangle_I - \langle \Omega - \zeta\delta t, \delta'\zeta - \phi\delta' t \rangle_I \\ &+ \langle \zeta, \Omega' \wedge \Omega \rangle_I \end{aligned} \quad (\text{IV.61})$$

où  $\langle \cdot, \cdot \rangle_I$  désigne le produit scalaire dans  $\mathbf{R}^3$  associé au *tenseur d'inertie* du solide

$$I = J - \text{Tr}J, \quad (\text{IV.62})$$

et où  $\Omega, \Omega', \zeta$  et  $\phi$  représentent respectivement  $\xi, \xi', Z$  et  $f$  :

$$\xi = j(\Omega) \quad \xi' = j(\Omega') \quad Z = j(\zeta) \quad f = j(\phi). \quad (\text{IV.63})$$

Le calcul du noyau de  $d\varpi$  est alors immédiat : il est engendré par le champ de vecteurs solution des équation :

$$\Omega = \zeta \quad \text{et} \quad \frac{d\zeta}{dt} = \phi + I^{-1}(I\zeta \wedge \zeta). \quad (\text{IV.64})$$

Ce sont les équations du mouvement du solide. On a l'habitude de les présenter dans une base diagonalisant le tenseur d'inertie, soit  $\lambda_i$  ( $i = 1, 2, 3$ ) les valeurs propres de  $I$ , le système précédent est équivalent au suivant :

$$\frac{d\zeta_i}{dt} = \frac{\lambda_j - \lambda_k}{\lambda_i} \zeta_j \zeta_k, \quad (i, j, k) = \text{pc}(1, 2, 3). \quad (\text{IV.65})$$

EXERCICE. Généraliser cette construction pour un groupe de Lie quelconque  $G$  et établir les équations du mouvement.

EXERCICE. Traiter le problème de 4 points dans  $\mathbf{R}^3$  contraints à se mouvoir en préservant le volume du tétraèdre qu'ils délimitent; utiliser l'action du groupe  $\text{SL}(3, \mathbf{R})$ .

REMARQUE. Lorsque l'on veut décrire le solide par une distribution continue de points et non plus par une distribution finie, on le représente par un compact  $S \subset \mathbf{R}^3$ , muni d'une densité de matière, c'est-à-dire une mesure  $dn$ . L'espace de configuration n'est plus  $Q = \mathbf{R} \times \mathbf{R}^{3N}$  mais  $Q = \mathbf{R} \times C^\infty(S, \mathbf{R}^3)$ . Chaque application  $r : S \rightarrow \mathbf{R}^3$  devient un état du solide. Le lagrangien du problème est modifié :

$$L(t, r, v) = \int_S \left[ \frac{1}{2} m \|v\|^2 - U \right] dn, \quad (\text{IV.66})$$

où  $v$ , comme  $r$ , est une application de  $S$  dans  $T_r \mathbf{R}^3 = \mathbf{R}^3$ , représentant la distribution des vitesses des constituants du solide;  $m$  est une fonction réelle positive définie sur  $S$ , représentant la distribution de masse. Les forces sont locales, ce qui explique le terme de potentiel sous forme d'intégrale. On en déduit la forme de Cartan associée :

$$\varpi(\delta y) = \int_S m \langle v, \delta r \rangle dn - h \delta t, \quad h = \int_S \left[ \frac{1}{2} m \|v\|^2 + U \right] dn. \quad (\text{IV.67})$$

Il faut noter que cette forme de Cartan est définie sur un espace de dimension infinie  $\mathbf{R} \times C^\infty(S, T\mathbf{R}^3)$ , où  $C^\infty(S, T\mathbf{R}^3) = TC^\infty(S, \mathbf{R}^3)$ . La contrainte solide se traite de la même manière que pour le cas fini. Les configurations solides sont encore une orbite du groupe d'Euclide, agissant naturellement sur  $C^\infty(S, \mathbf{R}^3)$  par composition :

$$g \in E, \quad r \in C^\infty(S, \mathbf{R}^3) : \quad g(r) = g \circ r. \quad (\text{IV.68})$$

Les mouvements du solide continu sont les solutions des mêmes équations que pour une distribution finie de points, mais avec la matrice des moments suivante :

$$J = \int_S m r \bar{r} \, dn. \quad (\text{IV.69})$$

On peut traiter, de la même manière, d'autres types de contrainte en remplaçant le groupe des déplacements euclidiens par un autre sous-groupe des difféomorphismes de  $\mathbf{R}^3$ . Par exemple, pour traiter les mouvements d'un fluide incompressible on choisira le sous-groupe des difféomorphismes qui préservent le volume. Une grande partie de la théorie des systèmes complètement intégrables trouve son origine dans la recherche des solutions du mouvement des solides. Le lecteur pourra consulter avec profit le beau livre de Michèle Audin [Aud1].

---

# Invariance du lagrangien et théorème de Noether

---

Comme nous l'avons vu dans les exemples précédents, les groupes de symétries des problèmes nous ont permis de réaliser simplement l'espace des solutions du système et de lui fournir sa structure symplectique. Cela n'est pas dû au hasard, c'est une conséquence du théorème de Emmy Noether. Ce théorème associe à tout groupe de symétrie du lagrangien une famille d'intégrales premières. Dans certaines conditions, ces intégrales premières sont suffisantes pour séparer les solutions, comme pour le cas des exemples que nous avons rencontrés précédemment. Nous verrons ensuite à quelle classe de lagrangiens, « presque » invariants, nous pouvons étendre cette construction. Dans ce cas, il apparaît une classe de cohomologie non triviale, associée à cette *presque invariance*. Nous retrouverons cette classe de cohomologie au chapitre suivant, dans l'étude générale de l'application moment.

## V.1 Lagrangien invariant par un groupe de symétries

Considérons d'abord un groupe de Lie  $G$  qui agit sur une variété de configuration  $Q$ , et son action dérivée sur  $TQ$  :

$$g : (q, \dot{q}) \mapsto (g(q), dg_q(\dot{q})). \quad (\text{V.1})$$

Soit  $L$  un lagrangien homogène défini sur  $TQ$ . Supposons que l'action de  $G$  préserve le lagrangien :

$$\forall g \in G \quad L(q, \dot{q}) = L(g(q), dg_q(\dot{q})). \quad (\text{V.2})$$

L'application de Legendre est alors équivariante sous l'action de  $G$  :

$$\forall \delta \dot{q} \in T_q Q \quad dL_q(\delta \dot{q}) = dL_{g(q)}(dg_q(\delta \dot{q})), \quad (\text{V.3})$$

c'est-à-dire :

$$P(g(q), dg_q(\dot{q})) = P(q, \dot{q}) dg_q^{-1}. \quad (\text{V.4})$$

La forme de Cartan  $\varpi$  est donc invariante sous l'action dérivée de  $G$  :

$$\forall g \in G \quad g^* \varpi = \varpi. \quad (\text{V.5})$$

En effet, soit  $x = (q, \dot{q}) \in TQ$  et  $\delta x \in T_x Q$ , évaluons  $g^* \varpi$  sur  $\delta x$  :

$$\begin{aligned} (g^* \varpi)_x(\delta x) &= \varpi_{g(x)}(dg_x(\delta x)) \\ &= P(g(q), dg_q(\dot{q}))(dg_q(\delta \dot{q})) \\ &= P(q, \dot{q}) dg_q^{-1}(dg_q(\delta \dot{q})) \\ &= P(q, \dot{q})(\delta \dot{q}) \\ &= \varpi_x(\delta x) \end{aligned} \quad (\text{V.6})$$

Soit  $Z$  un vecteur de l'algèbre de Lie  $\mathfrak{G}$  de  $G$ , soit  $Z_Q$  et  $Z_{TQ}$  les champs de vecteurs fondamentaux associés, définis sur  $Q$  et  $TQ$  :

$$Z_Q(q) = D(g \mapsto g_Q(q))(\mathbf{1})(Z), \quad Z_{TQ}(x) = D(g \mapsto g_{TX}(x))(\mathbf{1})(Z). \quad (\text{V.7})$$

Puisque  $\varpi$  est invariante sous l'action de  $G$ , sa dérivée de Lie par rapport à tout champ de vecteur fondamental est nulle :

$$\forall Z \in \mathfrak{G} : \quad \mathcal{L}_Z \varpi = 0 \quad (\text{V.8})$$

en appliquant à  $\varpi$  la formule de Cartan :

$$\mathcal{L}_Z \varpi = d\varpi(Z_{TQ}, \cdot) + d[\varpi(Z_{TQ})], \quad (\text{V.9})$$

on obtient :

$$d\varpi(Z_{TQ}(x), \cdot) = -d[\mu \cdot Z], \quad \mu(x) = [Z \mapsto \varpi(Z_{TQ}(x))]. \quad (\text{V.10})$$

L'application  $\mu$  est évidemment à valeur dans le dual  $\mathfrak{G}^*$  de l'algèbre de Lie de  $G$ , puisque  $(Z + Z')_{TQ} = Z_{TQ} + Z'_{TQ}$ . Elle est évidemment différentiable :  $\mu \in C^\infty(Y, \mathfrak{G}^*)$ , et s'exprime simplement grâce à l'application de Legendre. Cette construction donne lieu à la définition suivante.

DÉFINITION. *L'application  $\mu$  définie par :*

$$\mu(q, \dot{q}) \cdot Z = P(q, \dot{q})(Z_Q(q)), \quad (\text{V.11})$$

*est appelée application moment de l'action du groupe  $G$  sur  $Y$ .*

Par linéarité, l'action de  $G$  sur  $TQ$  passe sur l'espace des directions tangentées  $Y = SQ$ ; d'autre part, l'application moment  $\mu$  est clairement invariante par les dilatations : elle est donc définie aussi sur  $Y$ , son intérêt majeur résidant alors dans le théorème suivant.

THÉORÈME. (EMMY NETHER) *Soit  $Q$  une variété différentiable,  $L$  un lagrangien homogène défini sur  $TQ$ ,  $\varpi$  la forme de Cartan associée. Soit  $G$  un groupe de Lie agissant sur  $Q$  et préservant le lagrangien  $L$ . Si la dérivée extérieure de la forme de Cartan définie sur  $Y = SQ$  est présymplectique, alors l'application moment est constante sur les solutions du problème variationnel associé à  $L$ .*

DÉMONSTRATION. *On sait que les solutions du problème variationnel sont les caractéristiques de  $d\varpi$  sur  $Y$  ; soit  $s \mapsto y$  un paramétrage d'une caractéristique  $m$ ,  $dy/ds \in \ker d\varpi$  ; grâce à la formule (V.10) on a :*

$$\forall Z \in \mathfrak{G}^* \quad 0 = d\varpi(Z_Y, dy/ds) = -d[\mu(y) \cdot Z](dy/ds), \quad (\text{V.12})$$

*donc  $\mu$  est constant le long de  $m$ . ■*

REMARQUE. L'application moment est la construction, dans le cadre de la mécanique variationnelle, qui associe symétries et invariants.

EXERCICE. Calculer l'application moment dans les exemples des géodésiques de la sphère  $S^2$  et du disque de Poincaré  $H^2$ . Que peut-on dire sur la construction qui a été donnée des espaces des géodésiques, dans chacun de ces cas ?

REMARQUE. La démonstration du théorème de Noether ne nécessite pas que l'action du groupe  $G$  sur  $TQ$  soit relevée d'une action sur la base  $Q$  ; elle s'adapte donc parfaitement au cas d'un groupe agissant sur  $TQ$  qui préserve le lagrangien sans autre condition, mais la formule du moment (V.11) n'est alors plus valable telle quelle.

Si le lagrangien  $L$  est invariant sous l'action de  $G$ , nous avons vu que la forme de Cartan est elle-même invariante. Plus précisément on a

PROPOSITION. *La forme de Cartan est invariante sous l'action du groupe  $G$  si et seulement si le lagrangien est invariant.*

DÉMONSTRATION. *C'est une conséquence immédiate de la formule (V.4), appliquée au vecteur  $\delta\dot{q} = \dot{q}$  :*

$$P(q, \dot{q})(\dot{q}) = P(g(q), dg_q(\dot{q}))dg_q(\dot{q}) \Rightarrow L(q, \dot{q}) = L(g(q), dg_q(\dot{q})), \quad (\text{V.13})$$

puisque  $L(q, \dot{q}) = P(q, \dot{q})(\dot{q})$ . ■

## V.2 Lagrangien « presque » invariant

Dans le paragraphe précédent, nous avons noté que l'invariance du lagrangien par un groupe  $G$ , agissant sur  $Q$ , entraîne l'invariance de la forme de Cartan et donc l'existence d'une application moment constantes sur les solutions du problème. En d'autres termes, l'application moment est une fonction définie, non seulement sur l'espace des conditions initiales  $Y$ , mais aussi sur l'espace des solution  $\mathcal{M} = Y/\ker d\varpi$  ; nous la noterons de la même lettre  $\mu$  :

$$\mu : \mathcal{M} \mapsto \mathcal{G}^*. \quad (\text{V.14})$$

On peut supposer que  $\mathcal{M}$  est une variété, bien que cela ne soit pas tout à fait nécessaire. Le groupe  $G$  préservant  $d\varpi$  préserve son feuilletage caractéristique et

agit donc en réalité sur l'espace des solutions en préservant sa structure symplectique  $\omega = \pi_* d\varpi$ , où  $\pi$  est la projection naturelle de  $Y$  sur son quotient :

$$\forall g \in G : \quad g^* \omega = \omega. \quad (\text{V.15})$$

L'application moment  $\mu$  vérifie alors, sur  $\mathcal{M}$ , l'identité suivante :

$$\forall Z \in \mathcal{G} : \quad \omega(Z_{\mathcal{M}}, \cdot) = -d\mu \cdot Z. \quad (\text{V.16})$$

Cette formule est d'ailleurs la définition originale de l'application moment, dans le cadre général des actions de groupes de Lie sur les variétés symplectiques.

Et comme on peut le constater, il n'est pas nécessaire pour en arriver là que le groupe  $G$  préserve la forme de Cartan  $\varpi$ , c'est-à-dire le lagrangien  $L$ . Il suffit que pour tout  $Z \in \mathcal{G}$ , la forme différentielle  $d\varpi(Z_{TQ}, \cdot)$  soit exacte. Ce sera le cas, en particulier s'il existe une fonction  $F$  définie sur  $G \times TQ$  à valeurs réelles telle que :

$$g^* \varpi = \varpi + dF(g, \cdot). \quad (\text{V.17})$$

En notant que

$$\mathcal{L}_Z \varpi = df \cdot Z \quad (\text{V.18})$$

où l'application  $f$ , définie sur  $TQ$  à valeurs dans  $\mathcal{G}^*$ , est donnée par :

$$f = D(g \mapsto F(g, \cdot))(1), \quad (\text{V.19})$$

et en utilisant la formule de Cartan pour la dérivée de Lie de  $\varpi$ , on déduit l'expression du moment  $\mu$  :

$$\mu \cdot Z = \varpi(Z_{TQ}) - f \cdot Z = P \cdot Z_Q(q) - f \cdot Z, \quad \forall Z \in \mathcal{G}. \quad (\text{V.20})$$

Tout cela donne lieu à cette proposition/définition :

**PROPOSITION.** *S'il existe une fonction  $F : G \times TQ \rightarrow \mathbf{R}$  telle que, pour tout  $g \in B : g^* \varpi = \varpi + dF(g, \cdot)$ , alors pour tout  $Z \in \mathcal{G} : d\varpi(Z_{\mathcal{M}}, \cdot) = -d\mu \cdot Z$ . La fonction  $\mu$  est constante sur les caractéristiques de  $d\varpi$ . Elle est encore appelée application moment.*

REMARQUE. (DÉFAUT D'INVARIANCE) Cette application  $F$  traduit en réalité le défaut d'invariance de l'action  $a$ , associée à  $L$ . En effet, on peut toujours fixer la constante additive de  $F$  de telle sorte que :

$$a(\gamma) = \int_{\tilde{\gamma}} \varpi \quad \Rightarrow \quad a(g_*(\gamma)) = a(\gamma) + F(g)(q, \dot{q}) \quad (\text{V.21})$$

où  $\gamma$  est un chemin quelconque qui relie un point fixe  $q_0 = \gamma(0)$  à  $q = \gamma(1)$ , tel que  $\dot{q}_0 = \dot{\gamma}(0) = 0$  et  $\dot{q} = \dot{\gamma}(1)$ . Montrer que la différence  $a(g_*(\gamma)) - a(\gamma)$  est indépendante du chemin  $\gamma$  est d'ailleurs une méthode pour vérifier que l'on est dans les conditions supposées.

Cette application  $F$  n'est pas quelconque ; en utilisant la loi de groupe on déduit son comportement par rapport à la multiplication dans  $G$  :

$$(gg')^* \varpi = g'^*(g^* \varpi) \quad \Rightarrow \quad d[F(gg') - F(g) \circ g - F(g')] = 0. \quad (\text{V.22})$$

Si on suppose alors  $Q$  connexe, il existe une fonction  $c$  définie sur le produit  $G \times G$  à valeurs réelles telle que :

$$F(gg') = F(g) \circ g' + F(g') + c(g, g'). \quad (\text{V.23})$$

Mais là encore, cette fonction n'est pas quelconque ; grâce à l'associativité de l'action du groupe, on établit l'identité :

$$c(gg', g'') + c(g, g') = c(g, g'g'') + c(g', g''). \quad (\text{V.24})$$

Cette fonction  $c$  est donc un *deux-cocycle* du groupe  $G$  à valeurs réelles :

$$c \in Z^2(G, \mathbf{R}). \quad (\text{V.25})$$

Ce cocycle attaché au défaut d'invariance du lagrangien n'est pas bien défini. La fonction  $F$  n'étant définie qu'à une constante près, nous aurions pu choisir :

$$F'(g) = F(g) + k(g), \quad k \in C^\infty(G, \mathbf{R}). \quad (\text{V.26})$$

Le cocycle associé  $c'$  est alors relié à  $c$  par :

$$c'(g, g') = c(g, g') + k(gg') - k(g) - k(g') = c(g, g') + dk(g, g'), \quad (\text{V.27})$$

où  $dk$  est le *cobord* de  $k$ ;  $c'$  est donc cohomologue à  $c$ . Ce n'est pas le cocycle  $c$  qui est défini par le défaut d'invariance du lagrangien mais sa classe  $\sigma$  :

$$\sigma = [c] \in H^2(G, \mathbf{R}). \quad (\text{V.28})$$

Ce cocycle, ou plutôt son cocycle dérivé, est aussi relié à la variance du moment  $\mu$  par le groupe  $G$ . Plaçons nous dans le cadre général d'une 2-forme fermée  $\omega$  définie sur une variété connexe  $X$ , munie d'une action d'un groupe de Lie  $G$  et possédant un moment  $\mu$ , défini par la formule (V.16). Soit  $Z \in \mathfrak{G}$  et  $Z_X$  son champ de vecteur fondamental sur  $X$ , notons la variance du champ  $Z_X$  sous l'action de  $G$  :

$$Z_X(g(x)) = dg_x[\text{ad}(g^{-1})(Z)_X(x)], \quad (\text{V.29})$$

où  $\text{ad}$  désigne l'action adjointe de  $G$  sur son algèbre de Lie; elle est définie par :

$$\text{ad}(g)(Z) = D(L_g \circ R_{g^{-1}})(\mathbf{1})(Z), \quad (\text{V.30})$$

où  $R_g$  et  $L_g$  désignent les multiplications à droite et à gauche par  $g$  dans  $G$ . Prenons l'image réciproque par  $g$ , de chaque terme de l'équation (V.16) :

$$g^*(\omega(Z_X, \cdot)) = -dg^*\mu \cdot Z. \quad (\text{V.31})$$

En appliquant ensuite la formule (V.30) et en utilisant l'invariance de  $\omega$  par  $G$  on obtient :

$$g^*(\omega(Z_X, \cdot)) = \omega(\text{ad}(g^{-1})(Z)_X, \cdot), \quad (\text{V.32})$$

d'où on déduit, en appliquant la définition du moment :

$$d[(g^*\mu - \text{ad}^*(g)(\mu)) \cdot Z] = 0, \quad \forall Z \in \mathfrak{G}, \quad (\text{V.33})$$

où  $\text{ad}^*(g)$  désigne l'action coadjointe de  $G$  sur le dual  $\mathfrak{G}^*$  de l'algèbre de Lie de  $G$  :

$$\text{ad}^*(g)(\mu) : Z \mapsto \mu \cdot \text{ad}(g^{-1})(Z). \quad (\text{V.34})$$

L'application  $g^*\mu - \text{ad}^*(g)(\mu)$  est donc constante sur  $X$ . Il existe ainsi une application  $\Theta$  définie sur  $G$  à valeur dans le dual  $\mathfrak{G}^*$  telle que :

$$g^*\mu = \text{ad}^*(g)(\mu) + \Theta(g). \quad (\text{V.35})$$

Sa variance par rapport au groupe  $G$  est donnée par :

$$\Theta(gg') = \text{ad}^*(g)[\Theta(g')] + \Theta(g). \quad (\text{V.36})$$

On reconnaît ici un un-cocycle du groupe  $G$  à valeurs dans le dual de son algèbre de Lie. Il existe une relation entre ce un-cocycle  $\Theta$  et le deux-cocycle  $c$  dont il a été question précédemment (éq. V.23), donnée par :

$$\Theta(g)(Z) = \frac{\partial}{\partial t} \{c(g^{-1}e^{tZ}, g) - c(g, g^{-1}e^{tZ})\}_{t=0}. \quad (\text{V.37})$$

Notons  $x = (q, \dot{q}) \in TQ$ . Grâce à la formule (V.30) et à l'hypothèse (éq. V.17), on établit une première valeur de  $\Theta = g^*\mu - \text{ad}^*(g)(\mu)$  :

$$\Theta(g) \cdot Z = dF(g, \cdot)_x(\text{ad}(g^{-1})(Z)_{TQ}(x)) + f(x) \cdot \text{ad}(g^{-1})(Z) - f(g(x)) \cdot Z. \quad (\text{V.38})$$

On calcule alors la variance de la fonction  $f$ , dérivée de  $F$  en l'identité :

$$f(g(x)) \cdot Z = dF(\cdot, x)_g[d(R_g)_1(Z)] - dc(\cdot, g)_1(Z), \quad (\text{V.39})$$

puis, grâce à la formule (V.23) de  $F$ , on établit ensuite :

$$\begin{aligned} dF(g, \cdot)_x(\text{ad}(g^{-1})(Z)_{TQ}(x)) &= dF(\cdot, x)_g[d(R_g)_1(Z)] \\ &- f(x) \cdot \text{ad}(g^{-1})(Z) \\ &- dc(g, \cdot)_1(\text{ad}(g^{-1})(Z)). \end{aligned} \quad (\text{V.40})$$

Il ne nous reste plus qu'à assembler les morceaux pour trouver :

$$\Theta(g) \cdot Z = dc(\cdot, g)_1(Z) - dc(g, \cdot)_1(\text{ad}(g^{-1})(Z)), \quad (\text{V.41})$$

ce qui, après quelques manipulations supplémentaires, donne le résultat annoncé. On dit que  $\Theta$  est le *cocycle dérivé* de  $c$ . Cette construction est générale à la théorie de la cohomologie des groupes.

REMARQUE. Lorsque le lagrangien  $L$  est invariant par le groupe  $G$ , le cocycle  $c$  est alors trivial; le moment  $\mu$  est donc équivariant, ce qui peut être vérifié directement.

EXEMPLE. Grâce à cette construction, Souriau [Sou2] a montré que la masse totale d'un système dynamique galiléen libre peut être définie comme la classe de cohomologie associée au défaut d'équivariance du moment, pour l'action du groupe de Galilée. Précisons qu'un *système dynamique galiléen libre* est une variété symplectique  $(X, \omega)$  munie d'une *action hamiltonienne* du groupe de Galilée et qu'une action symplectique est dite *hamiltonienne* si elle possède un moment  $\mu$ . Le groupe de Galilée est défini par son action sur  $\mathbf{R} \times \mathbf{R}^3$  par :

$$q = \begin{pmatrix} r \\ t \end{pmatrix} \mapsto g(q) = \begin{pmatrix} Ar + bt + c \\ t + e \end{pmatrix}, \quad g = \begin{pmatrix} A & b & c \\ 0 & 1 & e \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{V.42})$$

avec  $A \in \text{SO}(3)$ ,  $b, c \in \mathbf{R}^3$ ,  $e \in \mathbf{R}$ , la loi de groupe étant la loi de multiplication des matrices.

Considérons le cas particulier d'une assemblée de  $N$  particules sans interactions de lagrangien homogénéisé :

$$L(q, \dot{q}) = \sum_{i=1}^N \frac{m_i \|\dot{x}_i\|^2}{2t}. \quad (\text{V.43})$$

On peut vérifier directement, en utilisant les résultats de l'exemple du point matériel (page 61), que la dérivée extérieure de la forme de Cartan associée est invariante par le groupe de Galilée. En utilisant la remarque sur le défaut d'invariance (page 93), on établit immédiatement que la fonction  $F$  est donnée par :

$$F(g, q) = \left\langle A \sum_{i=1}^N m_i r_i, b \right\rangle + \sum_{i=1}^N m_i \frac{b^2 t}{2}. \quad (\text{V.44})$$

On déduit ensuite l'expression du cocycle  $c$  associé :

$$c = \sum_{i=1}^N m_i c_0, \quad \text{avec} \quad c_0(g, g') = \langle Ac', b \rangle + \frac{b^2 e'}{2}. \quad (\text{V.45})$$

Bargmann a montré que la cohomologie du groupe de Galilée est de dimension 1 ; on peut vérifier, d'autre part, que le cocycle  $c_0$  représente une base de cette cohomologie. Le cocycle  $c$  est donc nécessairement proportionnel à  $c_0$ , mais on

remarque que le coefficient de proportionnalité est la masse totale du système. Il est donc légitime de définir, de manière générale, la masse totale d'un système dynamique galiléen isolé comme la classe de cohomologie du cocycle  $c$  associé à l'action du groupe de Galilée, indépendamment du fait que la variété symplectique  $(X, \omega)$  soit obtenu comme l'espace des solutions d'un problème variationnel.

On montre d'autre part que la cohomologie du groupe de Poincaré (le groupe des isométries de l'espace de Minkowski) est nulle. Cela explique en partie la difficulté qu'il y a à définir la masse d'un système relativiste.

L'*homomorphisme de Calabi* est un homomorphisme réel du groupe des difféomorphismes qui préservent l'aire du plan  $\mathbf{R}^2$ . La valeur de cet homomorphisme sur un difféomorphisme est donc invariante par conjugaison : c'est l'*invariant de Calabi*. Bien qu'il sorte légèrement du cadre de ce paragraphe, nous allons montrer comment cet homomorphisme est relié à la construction générale que nous avons exposée plus haut.

EXEMPLE. (L'INVARIANT DE CALABI) Considérons une variété symplectique exacte  $(X, \omega)$  avec  $\omega = d\varpi$ . Nous supposons  $X$  simplement connexe :  $\pi_1(X) = 0$ . Soit  $G$  un groupe de Lie agissant sur  $X$  par symplectomorphismes :  $g^*\omega = \omega$ . Puisque  $X$  est simplement connexe, il existe une fonction  $F : G \times X \rightarrow \mathbf{R}$ , telle que pour tout  $g \in G$  :  $g^*\varpi = \varpi + dF(g)$ . Nous pouvons même choisir  $F$  différentiable : choisissons une origine  $o \in X$  ; alors l'intégrale

$$\int_{\gamma} g^*\varpi - \varpi$$

ne dépend de  $\gamma$  que par son extrémité  $x = \gamma(1)$ . En effet, pour toute variation  $\delta\gamma$  nulle au bord ( $\delta\gamma(0) = \delta\gamma(1) = 0$ ), on a :

$$\begin{aligned} \delta \int_{\gamma} g^*\varpi - \varpi &= \int_{\gamma} d[g^*\varpi - \varpi](\delta\gamma) && \text{(V.46)} \\ &= \int_{\gamma} [g^*d\varpi - d\varpi](\delta\gamma) \\ &= \int_{\gamma} [g^*\omega - \omega](\delta\gamma) \\ &= 0. \end{aligned}$$

Cette intégrale est donc constante sur les composantes connexes des arcs d'extrémité  $x$ , c'est-à-dire sur les classes d'homotopie des arcs pointés en  $o$ . Or nous avons justement supposé  $X$  simplement connexe ; cette intégrale ne dépend donc que de  $x$ . Nous pouvons alors définir  $F$ , sans ambiguïté, par :

$$F(g, x) = \int_o^x g^* \varpi - \varpi. \quad (\text{V.47})$$

EXERCICE. *Montrer que la fonction  $F(g, \cdot)$  est différentiable.*

En appliquant encore la formule de la variation de l'intégrale le long du chemin  $\gamma$ , mais cette fois avec  $\delta\gamma(1) = \delta x$ , on vérifie que  $F$  est une solution de l'équation :

$$g^* \varpi = \varpi + dF(g), \quad (\text{V.48})$$

toutes les autres se déduisant de  $F$  par une constante additive. Soit  $c$  le cocycle réel de  $G$  qui lui est associé par la formule (V.23).

Choisissons maintenant pour  $X$  l'espace vectoriel  $\mathbf{R}^n \times \mathbf{R}^n$ , muni de sa forme symplectique ordinaire et choisissons pour  $G$  le groupe des difféomorphismes symplectiques à support compact :

$$g \in G \Leftrightarrow \exists K \text{ compact} : g | \mathbf{R}^n \times \mathbf{R}^n - K = \mathbf{1}. \quad (\text{V.49})$$

Bien que  $G$  ne soit pas un groupe de Lie de dimension finie, ce qui a été dit plus haut est toujours vrai. Nous pouvons remarquer que pour  $g$  donné la fonction  $F(g)$  est constante en dehors du support de  $g$  ; en effet si  $x$  et  $x'$  sont deux points à l'extérieur du support de  $g$ , ils peuvent être joints par un chemin toujours contenu à l'extérieur du support de  $g$  sur lequel  $g^* \varpi = \varpi$ , et donc :

$$\forall x, x' \notin \text{Supp}(g) : F(g, x) = F(g, x'). \quad (\text{V.50})$$

On peut définir alors une application  $k : G \rightarrow \mathbf{R}$  par :

$$k(g) = F(g, x), \quad x \notin \text{Supp}(g). \quad (\text{V.51})$$

Cette fonction est différentiable dans un sens que nous ne précisons pas ici et qui est suffisant pour justifier ce qui suit. Définissons alors une nouvelle solution  $\tilde{F}$  de l'équation (V.48) par :

$$\tilde{F}(g, x) = F(g, x) - k(g); \quad (\text{V.52})$$

elle vérifie, par construction :

$$\forall g \in G, \forall x \notin \text{Supp}(g) : \tilde{F}(g, x) = 0. \quad (\text{V.53})$$

En particulier, le cocycle associé  $\tilde{c}$  est nul :

$$\begin{aligned} \tilde{c}(g, g') &= \tilde{F}(gg', x) - \tilde{F}(g, g'(x)) - \tilde{F}(g', x) \\ &= \tilde{F}(gg', \infty) - \tilde{F}(g, g'(\infty)) - \tilde{F}(g', \infty) = 0. \end{aligned} \quad (\text{V.54})$$

Pour tout  $g \in G$ ,  $\tilde{F}(g)$  est une fonction réelle à support compact que l'on peut donc intégrer sur  $X = \mathbf{R}^n \times \mathbf{R}^n$ , soit :

$$\kappa(g) = \int_X \tilde{F}(g, x) \text{vol}, \quad \text{vol} = \omega^{\wedge n}. \quad (\text{V.55})$$

Cette fonction  $\kappa$  est un homomorphisme du groupe des difféomorphismes symplectiques à valeurs réelles ; en effet :

$$\kappa(gg') = \int_X \tilde{F}(gg', x) \text{vol} = \int_X \tilde{F}(g, g'(x)) \text{vol} + \int_X \tilde{F}(g', x) \text{vol}, \quad (\text{V.56})$$

mais puisque  $g^*\omega = \omega$  alors  $g^* \text{vol} = \text{vol}$ , et par un changement de variable :

$$\int_X \tilde{F}(g, g'(x)) \text{vol} = \int_{g(X)} \tilde{F}(g, x) g^* \text{vol} = \int_X \tilde{F}(g, x) \text{vol}, \quad (\text{V.57})$$

d'où on conclut :

$$\kappa(gg') = \int_X \tilde{F}(g, x) \text{vol} + \int_X \tilde{F}(g', x) \text{vol} = \kappa(g) + \kappa(g'). \quad (\text{V.58})$$

Cet homomorphisme est appelé *invariant de Calabi*.

EXERCICE. On considère  $\mathbf{R}^2$  muni de sa forme symplectique ordinaire  $\omega = dx \wedge dy$ . On note  $(\rho, \theta)$  les coordonnées polaires.

1. Donner l'expression de la forme symplectique  $\omega$  en coordonnées polaires.

On considère une fonction réelle positive  $\phi$ , définie sur l'intervalle  $]0, \infty[$ , à support compact, contenue dans l'intervalle  $]1, 2[$ .

2. Vérifier que l'application  $g : (\rho, \theta) \mapsto (\rho, \theta + \phi(\rho))$  définit un symplectomorphisme (à support compact) de  $\mathbf{R}^2$ .
3. Calculer  $\kappa(g)$ , et en déduire que l'invariant de Calabi n'est pas trivial.
4. Généraliser à  $\mathbf{R}^n \times \mathbf{R}^n$ .

EXERCICE. Essayer d'établir le cadre le plus général de la construction de l'invariant de Calabi.

---

# L'application moment

---

L'application moment que nous avons introduite sous plusieurs formes dans le chapitre précédent est un outil essentiel dans l'étude des actions symplectiques de groupe de Lie. Il n'est pas nécessaire, en réalité, de considérer les variétés symplectiques, mais seulement les variétés munies d'une 2-forme fermée. Comme le théorème de Noether le fait remarquer, l'intérêt du moment en mécanique est justement d'être conservé le long des caractéristiques de  $\omega$ . Mais le moment n'est pas vraiment un objet de la géométrie des 2-formes fermées invariantes. Je voudrais insister sur cet aspect : on présente souvent le moment comme un objet de géométrie symplectique, mais c'est une erreur. Comme nous allons le montrer dans ce chapitre, le moment est un objet de la géométrie des 1-formes différentielles invariantes.

Nous considérerons une 2-forme fermée  $\omega$  définie sur une variété  $X$ , munie d'une action différentiable d'un groupe de Lie  $G$ , supposé connexe. On pourra être conduit à étendre la notion de groupe de Lie à la dimension infinie en considérant des sous-groupes du groupe des difféomorphismes préservant la 2-forme  $\omega$ . Dans ce cas, nous considérerons comme algèbre de Lie du groupe l'algèbre de Lie des champs de vecteurs dont la dérivée de Lie annule  $\omega$ . Nous le préciserons à chaque fois. Mais avant tout, donnons une première définition formelle, dans ce cadre général, de l'application moment.

*DÉFINITION. Soit  $X$  une variété différentiable munie d'une 2-forme fermée  $\omega$ , invariante sous l'action d'un groupe de Lie  $G$ . On dit que l'action de  $G$  est hamil-*

tonienne si, pour tout  $Z \in \mathcal{G}$ , la 1-forme fermée  $\omega(Z_X, \cdot)$  est exacte.

On peut choisir la primitive de  $\omega(Z_X, \cdot)$  de telle sorte qu'elle dépende linéairement de  $Z$ , comme le dit la proposition suivante, qui est aussi la définition originale de l'application moment introduite par J.-M. Souriau [Sou2]. Nous verrons au paragraphe suivant une autre définition du moment qui évite l'introduction d'une base de l'algèbre de Lie.

**PROPOSITION. (APPLICATION MOMENT)** *Si l'action de  $G$  est hamiltonienne, il existe une application  $\mu \in C^\infty(X, \mathcal{G}^*)$  telle que :*

$$\omega(Z_X, \cdot) = -d\mu \cdot Z. \quad (\text{VI.1})$$

*Une telle application est appelée application moment.*

**DÉMONSTRATION.** *Soit  $Z_i, i = 1, \dots, k$ , une base de l'algèbre de Lie  $\mathcal{G}$  de  $G$ ,  $G$  étant supposé de dimension finie. Par hypothèse, il existe une famille de fonctions  $\mu_i \in C^\infty(X, \mathbf{R})$ , telles que  $\omega(Z_{iX}, \cdot) = -d\mu_i$ . Soit  $Z = \sum_i a_i Z_i$  un vecteur quelconque de  $\mathcal{G}$  :  $\omega(Z_X, \cdot) = \omega(\sum_i a_i Z_{iX}, \cdot) = \sum_i a_i \omega(Z_{iX}, \cdot) = -\sum_i a_i d\mu_i$  ; définissons  $\mu \cdot Z = \sum_i a_i \mu_i$  et il vient  $\omega(Z_X, \cdot) = -d\mu \cdot Z$ . ■*

Il faut noter que si  $X$  est connexe, deux moments ne diffèrent évidemment que d'une constante.

Pour des exemples d'utilisation de l'application moment dans la classification des variétés symplectiques, le lecteur pourra consulter les articles suivants [Igl1, Igl3], [Del1], [Del2] ou encore le livre de M. Audin [Aud2] qui intègre à peu près tout ce que l'on sait aujourd'hui sur cette question.

## VI.1 Conditions d'existence d'une application moment

Une des premières questions que l'on se pose lorsqu'on s'intéresse à l'application moment est celle des conditions de son existence. Soit  $Z$  un élément de l'algèbre de Lie  $\mathcal{G}$  de  $G$ , on a :

$$\mathcal{L}_{Z_X} \omega = 0 \quad \Leftrightarrow \quad d[\omega(Z_X, \cdot)] = 0. \quad (\text{VI.2})$$

Soit  $\Phi$  l'application définie sur  $\mathcal{G}$  à valeurs dans  $H_{\text{dR}}^1(X)$  qui à  $Z$  associe la classe de cohomologie de  $\omega(Z_X, \cdot)$  :

$$\Phi : \mathcal{G} \rightarrow H_{\text{dR}}^1(X) : \quad \Phi(Z) = [\omega(Z_X, \cdot)]. \quad (\text{VI.3})$$

Cette application est évidemment linéaire. Supposer l'action de  $G$  *hamiltonienne*, c'est-à-dire supposer qu'il existe un moment  $\mu$  pour cette action, se traduit par  $\Phi = 0$ . D'autre part, le groupe  $G$  agit sur l'algèbre de Lie par l'action adjointe. Grâce à la variance de la 1-forme  $\omega(Z_X, \cdot)$  par rapport à l'action du groupe  $G$ , que nous rappelons :

$$g^*(\omega(Z_X, \cdot)) = \omega(\text{ad}(g^{-1})(Z)_X, \cdot), \quad (\text{VI.4})$$

on déduit :

$$\Phi(\text{ad}(g)(Z)) = [g_*(\omega(Z_X, \cdot))], \quad (\text{VI.5})$$

mais comme le groupe est supposé connexe :

$$\Phi(\text{ad}(g)(Z)) = [g_*(\omega(Z_X, \cdot))] = [\omega(Z_X, \cdot)] = \Phi(Z),$$

c'est-à-dire :

$$\forall g \in G : \quad \Phi \circ \text{ad}(g) = \Phi. \quad (\text{VI.6})$$

En dérivant cette fonction par rapport à  $g$  on remarque qu'elle s'annule sur l'algèbre de Lie dérivée, autrement dit qu'elle se projette en une application  $\phi$  définie sur l'espace vectoriel quotient  $\mathcal{G}/[\mathcal{G}, \mathcal{G}] = H_1(\mathcal{G}, \mathbf{R})$ . En termes de cohomologie,  $\phi$  est une classe de cohomologie de  $\mathcal{G}$  à valeurs dans  $H_{\text{dR}}^1(X)$  :

$$\phi \in H^1(\mathcal{G}, H_{\text{dR}}^1(X)) = H^1(\mathcal{G}, \mathbf{R}) \otimes H_{\text{dR}}^1(X). \quad (\text{VI.7})$$

Cette classe représente donc l'obstruction à l'existence du moment pour l'action de  $G$ . On remarque en particulier deux conditions suffisantes pour l'existence du moment : d'une part, de façon évidente, si le premier groupe de cohomologie de Rham de  $X$  est nul, d'autre part si le premier groupe de cohomologie de  $G$  est nul, c'est-à-dire si  $\mathcal{G} = [\mathcal{G}, \mathcal{G}]$ , en particulier si  $G$  est semi-simple.

Dans ce paragraphe, nous avons situé l'habitat naturel de l'obstruction  $\phi$  à l'existence du moment. Mais cette obstruction peut être calculée simplement, dans tous les cas, comme un sous-groupe de  $\mathcal{G}^*$ , comme nous l'indiquons dans le paragraphe suivant.

## VI.2 Définition alternative du moment

Dans le paragraphe précédent nous avons introduit le lieu naturel de l'obstruction à l'existence du moment comme l'espace  $H^1(\mathcal{G}, \mathbf{R}) \otimes H_{\text{dR}}^1(X)$ . Voici une façon directe de présenter cette obstruction comme un sous-groupe de  $\mathcal{G}^*$ , et qui permet en même temps d'élargir la notion d'application moment.

Considérons l'action d'un groupe de Lie  $G$  sur une variété  $X$  préservant une 2-forme fermée  $\omega$ , la variété  $X$  étant supposée connexe. Choisissons un point base  $x_o \in X$  et soit  $\text{Arc}(X, x_o)$  l'espace des arcs de  $X$  basés en  $x_o$  :

$$\gamma \in \text{Arc}(X, x_o) \quad \Rightarrow \quad \gamma \in \text{Arc}(X) \text{ et } \gamma(0) = x_o. \quad (\text{VI.8})$$

Les arcs considérés, éléments de  $\text{Arc}(X)$ , sont les applications différentiables par morceaux de  $[0, 1]$  dans  $X$ . La différentiabilité par morceaux est requise pour autoriser les opérations de juxtaposition. Rappelons que le juxtaposé  $\gamma \vee \gamma'$  de deux arcs  $\gamma$  et  $\gamma'$ , basés en  $x_o$ , est l'arc, basé en  $x_o$ , défini par :

$$\gamma \vee \gamma' = \begin{cases} \gamma(2t) & 0 \leq t \leq \frac{1}{2} \\ \gamma(2t - 1) & \frac{1}{2} \leq t \leq 1 \end{cases} \quad (\text{VI.9})$$

Rappelons aussi que l'*opposé* d'un arc  $\gamma$  est l'arc noté  $\bar{\gamma}$ , défini par :

$$\bar{\gamma} = \gamma(1 - t). \quad (\text{VI.10})$$

Soit  $Z$  un élément de l'algèbre de Lie  $\mathcal{G}$ ,  $Z_X$  le champ de vecteurs fondamental sur  $X$ , associé à  $Z$ , et  $\omega(Z_X)$  le contracté de  $\omega$  par  $Z_X$ . Considérons l'intégrale suivante, dépendante de  $Z \in \mathcal{G}$  et de  $\gamma \in \text{Arc}(X, x_o)$  :

$$\int_{\gamma} \omega(Z_X) = \int_0^1 \omega \left( Z_X(\gamma(t)), \frac{d\gamma(t)}{dt} \right) dt. \quad (\text{VI.11})$$

La dépendance en  $Z$  est évidemment linéaire ; la fonction  $[Z \mapsto \int_{\gamma} \omega(Z_X)]$  est donc naturellement à valeurs dans  $\mathcal{G}^*$  ; nous définissons ainsi l'application :

$$\begin{aligned} \Psi : \text{Arc}(X, x_o) &\longrightarrow \mathcal{G}^* \\ \gamma &\longmapsto \left[ Z \mapsto \int_{\gamma} \omega(Z_X) \right] \end{aligned} \quad (\text{VI.12})$$

Cette fonction est additive pour la juxtaposition des arcs : on vérifie, par un simple changement de paramètre sous le signe somme, que

$$\Psi(\gamma \vee \gamma') = \Psi(\gamma) + \Psi(\gamma'). \quad (\text{VI.13})$$

Considérons maintenant deux arcs  $\gamma$  et  $\gamma'$  qui aboutissent tous les deux au point  $x = \gamma(1) = \gamma'(1)$  ; on a clairement, en notant que  $\Psi(\bar{\gamma}) = -\Psi(\gamma)$  :

$$\begin{aligned} \Psi(\gamma') &= \Psi(\gamma) + [\Psi(\gamma') - \Psi(\gamma)] \\ &= \Psi(\gamma) + \Psi(\ell), \text{ avec } \ell = \gamma' \vee \bar{\gamma} \end{aligned}$$

où  $\ell$  est un lacet pointé en  $x_o$ . Définissons alors  $\Gamma_\omega$  comme l'image par  $\Psi$  du sous-espace des lacets de  $X$ , pointés en  $x_o$

$$\Gamma_\omega = \Psi(\text{Lac}(X, x_o)) = \left\{ \int_\ell \omega(Z_X) \mid \ell \in \text{Lac}(X, x_o) \right\} \subset \mathcal{G}^*. \quad (\text{VI.14})$$

Puisque  $\Psi$  est additive par juxtaposition des arcs, son image  $\Gamma_\omega$  est un sous-groupe additif de  $\mathcal{G}^*$ . On peut alors définir l'application  $\psi$  par passage au quotient de  $\Psi$  sur  $\mathcal{G}^*/\Gamma_\omega$  par :

$$\begin{aligned} \psi : X &\longrightarrow \mathcal{G}^*/\Gamma_\omega \\ x &\longmapsto \left[ Z \mapsto \int_{x_o}^x \omega(Z_X) \right]_{\Gamma_\omega} \end{aligned} \quad (\text{VI.15})$$

où  $[\dots]_{\Gamma_\omega}$  représente la classe d'équivalence dans le groupe quotient  $\mathcal{G}^*/\Gamma_\omega$ . D'autre part, pour tout  $Z \in \mathcal{G}$ , la 1-forme  $\omega(Z_X)$  est fermée, et  $\Psi(\ell)$  ne dépend que de la classe d'homologie  $[\ell] \in H_1(X, \mathbf{Z})$  du lacet  $\ell$ . Le groupe  $\Gamma_\omega$  est donc l'image du groupe  $H_1(X, \mathbf{Z})$  dans  $\mathcal{G}^*$  par la projection  $h_\omega$  de  $\Psi|_{\text{Lac}(X, x_o)}$ , qui est évidemment un morphisme de groupe abélien :

$$\forall c \in H_1(X, \mathbf{Z}), \quad h_\omega(c) = \Psi(\ell) \text{ si } c = [\ell], \quad \Gamma_\omega = h_\omega(H_1(X, \mathbf{Z})). \quad (\text{VI.16})$$

Voici quelques propriétés immédiates de cette construction :

1. L'action de  $G$  est hamiltonienne si et seulement si  $\Gamma_\omega = \{0\}$ , auquel cas  $\psi$  est le moment.

2. Dans tous les cas, le quotient  $\mathcal{G}^*/\Gamma_\omega$  est un groupe abélien. Ce n'est un groupe de Lie que lorsque  $\Gamma_\omega$  est fermé ; dans ce cas  $\mathcal{G}^*/\Gamma_\omega$  est le produit d'un tore par un espace vectoriel réel. Sinon,  $\mathcal{G}^*/\Gamma_\omega$  doit être considéré comme un *groupe différentiable* au sens général des *espaces différentiables* (voir annexe D).
3. Dans tous les cas, l'application  $\psi$  est différentiable, même lorsque  $\Gamma_\omega$  n'est pas fermé dans  $\mathcal{G}^*$ . Cela signifie simplement que, pour tout  $x \in X$ , il existe un voisinage  $U$  de  $x$  et une application différentiable  $F : U \rightarrow \mathcal{G}^*$  telle que pour tout  $y \in U$  :  $\psi(y) = [F(y)]_{\Gamma_\omega}$ , ce qui peut s'interpréter aussi comme une simple conséquence de ce qu'une forme fermée est localement exacte.
4. L'application  $\psi$  est invariante sur les caractéristiques<sup>1</sup> de la 2-forme  $\omega$  : étant donnés deux points  $x$  et  $y$  sur une même caractéristique, il suffit de joindre  $x_o$  à  $x$  par un chemin quelconque, et de joindre  $x_o$  à  $y$  en juxtaposant au chemin précédent un arc joignant  $x$  à  $y$  tracé entièrement dans la caractéristique.

EXEMPLE. Considérons l'exemple suivant :  $X$  est le tore  $T^2 = \mathbf{R}^2/\mathbf{Z}^2$  ; la forme  $\omega$  est le volume ordinaire  $\omega = dx \wedge dy$  (*modulo* un abus de langage courant). Le groupe  $T^2$  agit sur lui-même par translation, en préservant le volume  $\omega$ . Considérons les deux générateurs canoniques de  $H_1(T^2, \mathbf{Z})$  :

$$a = [t \mapsto ([t], 1)] \quad b = [t \mapsto (1, [t])], \quad (\text{VI.17})$$

où  $t$  varie de 0 à 1. L'algèbre de Lie  $\mathcal{T}^2$  de  $T^2$  est évidemment  $\mathbf{R}^2$  et son dual  $\mathcal{T}^{2*}$  est aussi identifié à  $\mathbf{R}^2$  grâce au produit scalaire ordinaire. Les images des générateurs  $a$  et  $b$  de  $H_1(T^2, \mathbf{Z})$  par l'homomorphisme  $h_\omega$  sont données par :

$$h_\omega(a) = (0, -1) \quad h_\omega(b) = (1, 0). \quad (\text{VI.18})$$

Ainsi, l'image  $\Gamma_\omega = h_\omega(H_1(T^2, \mathbf{Z}))$  est le réseau ordinaire  $\mathbf{Z}^2 \subset \mathcal{T}^{2*}$ , engendré par  $h_\omega(a)$  et  $h_\omega(b)$ . L'application  $\psi$  est un difféomorphisme, et le lecteur peut vérifier

<sup>1</sup>J'appelle sous-variété *caractéristique* de la 2-forme  $\omega$  toute sous-variété de  $X$  connexe dont l'espace tangent est, en tout point, contenu dans le noyau de  $\omega$ . Un arc caractéristique est un arc différentiable tangent en tout point au noyau de  $\omega$ .

directement que :

$$\psi[x, y] = [xh_\omega(a) + yh_\omega(b)] \in T^2. \quad (\text{VI.19})$$

L'espace quotient  $\mathcal{J}^{2*}/\Gamma_\omega$  est donc isomorphe à  $T^2$ .

EXEMPLE. Dans l'exemple précédent, considérons l'action du groupe additif  $\mathbf{R}$  définie par restriction de l'action de  $T^2$  ci-dessus :

$$t : [x, y] \mapsto [(x, y) + t(\alpha, \beta)], \quad (\text{VI.20})$$

où  $\alpha$  et  $\beta$  sont des réels quelconques. Le dual de  $\mathbf{R}$  étant naturellement identifié à  $\mathbf{R}$  lui-même grâce au produit scalaire, le groupe  $\Gamma_\omega = h_\omega(H_1(T^2, \mathbf{Z}))$  est le sous-groupe :

$$\Gamma_\omega = \alpha\mathbf{Z} + \beta\mathbf{Z} \in \mathbf{R}. \quad (\text{VI.21})$$

Dans ce cas, le quotient  $\mathbf{R}/\Gamma_\omega$  est ou bien un cercle si  $\alpha$  et  $\beta$  sont commensurables ou bien un tore irrationnel.

EXEMPLE. On considère le tore  $\mathbf{T}^3 = \mathbf{R}^3/\mathbf{Z}^3$ . L'espace tangent en tout point  $X \in \mathbf{T}^3$  est naturellement identifié à  $\mathbf{R}^3$ . Soit  $\omega$  la 2-forme suivante :

$$\omega_X(u, v) = \langle \Omega, u \wedge v \rangle$$

où  $\Omega$  est un vecteur non nul de  $\mathbf{R}^3$ , et  $\wedge$  désigne le produit vectoriel de  $\mathbf{R}^3$ . Le groupe  $\mathbf{R}^3$  agit naturellement sur  $\mathbf{T}^3$  par translation en préservant la forme  $\omega$  :

$$\begin{cases} \forall \zeta = (r, s, t) \in \mathbf{R}^3, \forall X = [x, y, z] \in \mathbf{T}^3, \zeta_{\mathbf{T}^3}(X) = [x + r, y + s, z + t] \\ \zeta^*\omega = \omega. \end{cases}$$

L'algèbre de Lie de  $\mathbf{R}^3$  est naturellement identifiée à  $\mathbf{R}^3$ , ainsi que son dual grâce au produit scalaire ordinaire. Tout chemin basé en  $[0, 0, 0]$  dans  $\mathbf{T}^3$  est homotope à un chemin du type  $\gamma_X : t \mapsto [tX]$ , où  $X = (x, y, z) \in \mathbf{R}^3$ . La fonction  $\Psi$  est donnée par :

$$\Psi(\gamma_X) = [Z \mapsto -\langle \Omega \wedge X, Z \rangle] \quad i.e. \quad \Psi(\gamma_X) = -\Omega \wedge X.$$

Ce qui nous donne immédiatement le groupe  $\Gamma_\omega$  et le « moment généralisé »  $\psi$  :

$$\Gamma_\omega = \{\Omega \wedge \ell \mid \ell \in \mathbf{Z}^3\} \quad \text{et} \quad \psi[X] = [\Omega \wedge X]_{\Gamma_\omega}.$$

Soit  $\mathcal{M}$  l'image de  $\mathbf{T}^3$  par le « moment »  $\psi$ . D'après ce qui précède,  $\mathcal{M}$  est le quotient de l'espace vectoriel  $\Omega^\perp = j(\Omega)(\mathbf{R}^3)$  par le sous-groupe  $j(\Omega)(\mathbf{Z}^3)$  :

$$\mathcal{M} = \psi(\mathbf{T}^3) = j(\Omega)(\mathbf{R}^3)/j(\Omega)(\mathbf{Z}^3).$$

L'opérateur  $j$  désignant le produit vectoriel,  $j(\Omega)(X) = \Omega \wedge X$ . Précisons davantage la structure de  $\mathcal{M}$ . Soit  $\Omega = (a, b, c)$ ; puisque  $\Omega \neq 0$ , une de ses trois coordonnées  $a$ ,  $b$  ou  $c$  est non nulle; supposons que ce soit  $c$ . Nous pouvons alors identifier  $\Omega^\perp$  avec le plan  $z = 0$  par projection  $(x, y, z) \mapsto (x, y)$ . Le sous-groupe  $j(\Omega)(\mathbf{Z}^3)$  se projette alors sur le plan  $z = 0$  comme le sous-groupe :

$$j(\Omega)(\mathbf{Z}^3) \sim \left\{ \begin{pmatrix} bm - cl \\ ck - am \end{pmatrix} \in \mathbf{R}^2 \text{ tel que } \begin{pmatrix} k \\ l \\ m \end{pmatrix} \in \mathbf{Z}^3 \right\}.$$

Ainsi, l'image du moment est isomorphe au quotient de  $\mathbf{R}^2$  par la relation :

$$\begin{pmatrix} x \\ y \end{pmatrix} \sim \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} bm - cl \\ ck - am \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + c \begin{pmatrix} -l \\ k \end{pmatrix} + m \begin{pmatrix} b \\ -a \end{pmatrix}.$$

Le quotient s'obtient donc en deux étapes :

1) on prend d'abord le quotient de  $\mathbf{R}^2$  par le réseau  $c\mathbf{Z}^2$  et on obtient alors un tore ordinaire (de côté  $c$ )  $\mathbf{T}^2 = \mathbf{R}^2/c\mathbf{Z}^2$  ;

2) on prend ensuite le quotient de ce tore par  $\mathbf{Z}$ , agissant selon  $m : [x, y] \mapsto [x + mb, y - ma]$ ,  $m \in \mathbf{Z}$ . Ce quotient  $\mathbf{T}^2/\mathbf{Z}$  fibre sur le quotient  $\mathbf{T}^2/\mathbf{R}$ , où  $\mathbf{R}$  agit sur  $\mathbf{T}^2$  par  $t : [x, y] \mapsto [x + tb, y - ta]$ ,  $t \in \mathbf{R}$ , avec pour fibre  $S^1 = \mathbf{R}/\mathbf{Z}$ . Supposons alors que les nombres  $a$  et  $b$  soient rationnellement indépendants, et laissons l'autre cas comme exercice pour le lecteur. Soit  $\alpha = b/a$ , le quotient  $\mathbf{T}^2/\mathbf{R}$  est équivalent à  $T_\alpha = \mathbf{R}/(\mathbf{Z} + \alpha\mathbf{Z})$ ; en effet, il suffit de remonter au revêtement  $\mathbf{R}^2$  et de caractériser chaque orbite de  $\mathbf{R} \times \mathbf{Z}^2$  sur  $\mathbf{R}^2$  par sa trace sur l'axe des  $x$  (grâce à la projection  $(x, y) \mapsto x + \alpha y$ ). Ainsi, l'espace des moments de  $\mathbf{R}^3$ , pour la 2-forme fermée  $\omega$ , est un fibré en cercle sur un tore irrationnel  $\mathbf{T}_\alpha$ . La trivialité de ce fibré est liée à la commensurabilité des coefficients  $a, b, c$ . Nous laissons au lecteur le soin de préciser cela.

On vérifie d'autre part que l'application moment réalise le quotient symplectique de  $\mathbf{T}^3$  par le noyau de  $\omega$  ; autrement dit, l'espace des moments — ce fibré en cercle  $S^1$  sur un tore  $\mathbf{T}_\alpha$  — est muni de la structure « symplectique » naturelle d'espace de caractéristiques d'une 2-forme présymplectique.

REMARQUE. Cette construction nous encourage à étendre la notion de moment d'un groupe en définissant de façon générale le *Moment* d'un groupe de Lie  $G$  qui agit sur une variété  $X$  en préservant une 2-forme fermée  $\omega$ , comme l'application  $\psi$  à valeur dans  $\mathcal{G}^*/\Gamma_\omega$ , telle qu'elle est définie plus haut (VI.15). Cette définition n'est évidemment plus compatible avec la définition originale avec laquelle elle s'accorde seulement si  $\Gamma_\omega$  est nul. Mais la persistance du théorème de Noether rend cette élargissement de la définition du moment tout à fait raisonnable et même souhaitable. On pourrait alors conserver l'adjectif *hamiltonien* uniquement pour désigner les actions de groupes  $G$  sur  $(X, \omega)$  tels que  $\Gamma_\omega$  soit nul.

Cette construction a été utilisée dans le cas particulier de  $G = S^1$ , pour introduire une certaine notion de « moment à valeurs dans un groupe », mais cette notion (à part le cas précis de ce paragraphe) est loin d'être claire et nous ne l'évoquerons pas plus.

### VI.3 Le cas particulier des formes exactes

Un cas particulier important, nous l'avons vu avec les problèmes variationnels, est lorsque la 2-forme  $\omega$  est exacte :  $\omega = d\varpi$ . Pour tout élément  $g \in G$ , la 1-forme  $g^*\varpi - \varpi$  est fermée puisque  $g^*d\varpi = 0$  ; on a donc une application

$$\beta : G \rightarrow Z_{\text{dR}}^1(X) \quad \text{définie par} \quad \beta : g \mapsto g^*\varpi - \varpi, \quad (\text{VI.22})$$

telle que

$$\beta(gg') = g'^*\beta(g) + \beta(g'). \quad (\text{VI.23})$$

On dit que  $\beta$  est un un-cocycle de  $G$  dans  $Z_{\text{dR}}^1(X)$ . De cette définition de  $\beta$  on tire :

$$\mathcal{L}_Z \varpi = \left. \frac{\partial e^{tZ^*} \varpi}{\partial t} \right|_{t=0} = d\beta_1 \cdot Z \quad (\text{VI.24})$$

où  $d\beta_{\mathbf{1}}$  désigne la dérivée de  $\beta$  calculée en l'identité :  $D(\beta)(\mathbf{1})$ . On déduit donc que l'action de  $G$  est hamiltonienne si et seulement si  $d\beta_{\mathbf{1}} \cdot Z$  est exacte pour tout  $Z \in \mathcal{G}$ , mais on a :

PROPOSITION. *Soit  $X$  une variété différentiable munie d'une action d'un groupe de Lie connexe  $G$ . Soit  $\varpi$  une 1-forme sur  $X$  telle que sa dérivée extérieure soit invariante par  $G$ , soit  $\beta : g \mapsto g^*\varpi - \varpi \in Z_{\text{dR}}^1(X)$ . L'action de  $G$  est hamiltonienne pour  $d\varpi$  si et seulement si  $\beta(g)$  est exacte pour tout  $g \in G$ .*

DÉMONSTRATION. *Il est évident que si  $\beta(g)$  est exacte pour tout  $g$ ,  $d\beta_{\mathbf{1}} \cdot Z$  sera exacte pour tout  $Z \in \mathcal{G}$ . Réciproquement, supposons que  $d\beta_{\mathbf{1}} \cdot Z$  soit exacte pour tout  $Z \in \mathcal{G}$ . Par connexité de  $G$  la classe de cohomologie de  $g^*\beta(g)$  est égale à la classe de cohomologie de  $\beta(g)$ ; l'application  $g \mapsto [\beta(g)]$  étant donc un homomorphisme de  $G$  dans  $Z_{\text{dR}}^1(X)$ , il se projette donc sur l'abélianisé  $G/[G, G]$  de  $G$ . Soit  $[\beta]$  cette application :*

$$\begin{array}{ccc} G & \xrightarrow{\beta} & Z_{\text{dR}}^1(X) \\ \downarrow & & \downarrow \\ G/[G, G] & \xrightarrow{[\beta]} & H_{\text{dR}}^1(X) \end{array} \quad (\text{VI.25})$$

*Mais le groupe de Lie  $G$  étant connexe et de dimension finie,  $G/[G, G]$  est le produit direct d'un tore  $T^k$  par un espace vectoriel  $\mathbf{R}^l$ . L'homomorphisme  $\beta$  est évidemment nul sur le facteur  $T^k$ , et reste un homomorphisme de  $\mathbf{R}^l$  dans  $H_{\text{dR}}^1(X)$ , c'est-à-dire une application linéaire, puisque tout cela est différentiable. L'hypothèse revient à supposer  $d[\beta]_0(T_0\mathbf{R}^l) = \{0\}$ , mais  $[\beta]$  est linéaire donc  $[\beta](\mathbf{R}^l) = \{0\}$ , or  $[\beta](T^k) = \{0\}$ , d'où  $[\beta](G/[G, G]) = \{0\}$ ; autrement dit,  $\beta(g)$  est exacte pour tout  $g \in G$ . ■*

REMARQUE. Par définition de la cohomologie des groupes l'application  $[\beta]$  est un un-cocycle de  $G$  à valeurs dans  $H_{\text{dR}}^1(X)$  :

$$[\beta] \in H^1(G, H_{\text{dR}}^1(X)) = H^1(G, \mathbf{R}) \otimes H_{\text{dR}}^1(X). \quad (\text{VI.26})$$

Nous venons de dire que l'action est hamiltonienne si et seulement si ce cocycle est nul. On retrouve ainsi, d'une autre façon, l'obstruction précédente à l'existence du moment.

Nous pouvons appliquer maintenant tout ce qui a été dit au paragraphe V.2. L'action de  $G$  est hamiltonienne si et seulement si il existe une fonction  $F : G \mapsto C^\infty(G \times X, \mathbf{R})$  telle que :

$$g^*\varpi = \varpi + dF(g, \cdot), \quad (\text{VI.27})$$

Le défaut d'invariance de  $\varpi$  est encore défini par le deux-cocycle  $c$  de  $G$  :

$$c(g, g') = F(gg') - g'^*F(g) - F(g'), \quad (\text{VI.28})$$

et le moment  $\mu$  de l'action de  $G$  est encore donné par :

$$\mu \cdot Z = \varpi(Z_X) - f \cdot Z, \quad \text{avec } f : x \mapsto dF(\cdot, x)_1 \in \mathfrak{G}^*. \quad (\text{VI.29})$$

Les formules établies au paragraphe V.2 pour le défaut d'équivariance  $\Theta$  du moment  $\mu$  et sa relation avec  $c$  sont encore valables dans ce cadre.

Nous allons analyser plus en détail la nature du moment dans les trois différents cas qui se présentent :

1. La primitive  $\varpi$  est invariante par  $G$ .
2. La primitive  $\varpi$  n'est pas invariante par  $G$ , mais le cocycle  $c$  est trivial,
3. La primitive  $\varpi$  n'est pas invariante par  $G$  et  $c$  n'est pas trivial.

### Cas 1. La primitive $\varpi$ est invariante par $G$

Supposons que la primitive  $\varpi$  de  $\omega$  soit invariante par  $G$ , le moment  $\mu$  est alors donné par :

$$\mu : x \mapsto \mu(x) = [Z \mapsto \varpi_x(Z_X(x))]. \quad (\text{VI.30})$$

Nous pouvons aussi interpréter le moment  $\mu$  directement sur le groupe  $G$ . Considérons l'application  $\hat{x} : G \rightarrow X$  telle que  $\hat{x}(g) = g(x)$ ; soit  $\varpi^x = \hat{x}^*\varpi$ , c'est une 1-forme sur  $G$  invariante par l'action à gauche. En effet :  $L_g^*\varpi^x = L_g^*\hat{x}^*\varpi =$

$(\hat{x} \circ L_g)^* \varpi = (g \circ \hat{x})^* \varpi = \hat{x}^* g^* \varpi = \hat{x}^* \varpi = \varpi^x$ . C'est donc un élément du dual de l'algèbre de Lie de  $G$  par définition même, puisque l'algèbre de Lie de  $G$  est l'espace des champs de vecteurs sur  $G$  invariants à gauche et son dual l'espace des 1-formes invariantes à gauche<sup>2</sup>. La 1-forme  $\varpi^x$  est donc définie par sa valeur en l'identité, soit :  $\mu^x = \varpi_1^x$ , tel que  $\varpi^x = \mu^x \circ \theta$ ,  $\theta$  étant la forme de Maurer-Cartan. Il est clair que  $\mu^x = \mu(x)$ . Dans ce cas le moment est simplement l'application :

$$\mu : x \mapsto \mu(x) = \varpi^x = \hat{x}^* \varpi. \quad (\text{VI.31})$$

REMARQUE. Il est important d'insister sur ce cas, car il est l'essence même de l'application moment. Contrairement à ce qu'on peut penser et écrire (même ici), le moment n'est pas un objet de la géométrie des 2-formes différentielles fermées invariantes, encore moins un objet de la géométrie symplectique :

*Le Moment est un objet de la géométrie des 1-formes différentielles invariantes.*

C'est ce que nous nous efforcerons de montrer jusqu'à la fin de ces notes et nous définirons formellement :

DÉFINITION. (MOMENT DES 1-FORMES) *Soit  $\alpha$  une 1-forme différentielle définie sur une variété  $X$ . Soit  $G$  un groupe de Lie agissant différemment sur  $X$ . Si  $\alpha$  est invariante par  $G$ , nous appellerons moment de  $\alpha$  (sous l'action de  $G$ ) la fonction  $\mu : X \mapsto \mathcal{G}^*$  définie par  $\mu(x) = \hat{x}^* \alpha = [Z \mapsto \alpha_x(Z_X(x))]$ .*

Il faut noter que dans ce cas l'application moment est bien définie, et pas seulement à une constante près. La remarque précédente en appelle une nouvelle : il n'est pas nécessaire de se restreindre à un groupe de Lie préservant la forme  $\alpha$  ; on peut immédiatement considérer le moment de  $\alpha$  défini pour le groupe  $G_\alpha$  de tous les difféomorphismes de  $X$  préservant  $\alpha$ .

---

<sup>2</sup>On peut d'ailleurs définir le *dual-de-l'algèbre-de-Lie*, en un seul mot, directement comme l'espace des 1-formes de  $G$ , invariantes par l'action à gauche. Cela permet des généralisations au cas des groupes pour lesquels l'algèbre de Lie est mal définie. On peut préférer le nom d'*espace des moments*.

**Cas 2. La primitive  $\varpi$  n'est pas invariante par  $G$ , mais le cocycle  $c$  est trivial**

Supposons que la primitive  $\varpi$  ne soit pas invariante par  $G$ , la forme  $\varpi^x$  n'est alors, pas davantage, invariante par l'action à gauche de  $G$ . Notons  $F^x$  l'application réelle de  $G : F(\cdot, x)$ ; on a immédiatement  $L_g^* \varpi^x = \varpi^x + d[\hat{x}^* F(g, \cdot)]$ , et les identités sur  $F$  et  $c$  se traduisent par :

$$x^* F(g, \cdot) = L_g^* F^x - F^x - c(g, \cdot). \quad (\text{VI.32})$$

Si, comme c'est ici l'hypothèse,  $c \equiv 0$  alors  $L_g^* \varpi^x = \varpi^x + d[L_g^* F^x - F^x]$ . Autrement dit, la 1-forme différentielle  $\tilde{\varpi}^x = \varpi^x - dF^x$  est invariante par l'action à gauche de  $G$  : elle appartient au dual de l'algèbre de Lie de  $G$ , et sa valeur en l'identité nous donne exactement le moment  $\mu$  (tel qu'il est défini à la page 102) : pour tout  $Z \in \mathfrak{G}$ ,  $\tilde{\varpi}_1^x(Z) = \varpi(Z_X) - f \cdot Z$ . En d'autres termes

$$\mu : x \mapsto \tilde{\varpi}^x = \varpi^x - dF^x = x^* \varpi - dF(\cdot, x). \quad (\text{VI.33})$$

**Cas 3. La primitive  $\varpi$  n'est pas invariante par  $G$  et  $c$  n'est pas trivial**

Plaçons nous maintenant dans le cas général. La primitive  $\varpi$  n'est pas invariante et le cocycle  $c$  n'est pas trivial. Utilisons donc ce cocycle  $c$  qui nous est offert pour l'extension de  $G$  par  $\mathbf{R}$  suivante :

$$\tilde{G} = G \times \mathbf{R}, \quad (g, t) \cdot (g', t') = (gg', t + t' + c(g, g')). \quad (\text{VI.34})$$

La condition de cocycle sur  $c$  signifie exactement que  $\tilde{G}$  est un groupe. C'est une extension centrale de  $G$ . On voit tout de suite que le sous-groupe  $\{\mathbf{1}\} \times \mathbf{R}$  commute avec tous les éléments de  $G$ . Utilisons maintenant la fonction  $F$ , dont  $c$  est issu, pour définir une action de  $\tilde{G}$  sur le produit direct  $\tilde{X} = X \times \mathbf{R}$  :

$$\begin{pmatrix} g \\ t \end{pmatrix} \cdot \begin{pmatrix} x \\ s \end{pmatrix} = \begin{pmatrix} g(x) \\ s + t - F(g, x) \end{pmatrix}. \quad (\text{VI.35})$$

Là encore, les propriétés conjuguées de  $F$  et de  $c$  nous permettent de vérifier que l'on a bien défini une action de  $\tilde{G}$  sur  $\tilde{X}$ . L'image réciproque  $\tilde{\omega}$  de  $\omega$  sur  $\tilde{X}$  est évidemment exacte, mais choisissons comme primitive :

$$\tilde{\omega} = \varpi + ds. \quad (\text{VI.36})$$

Nous laissons en exercice le soin de vérifier que les signes ont été correctement choisis pour que  $\tilde{\omega}$  soit invariante par  $\tilde{G}$  :

$$\forall \tilde{g} \in \tilde{G} : \tilde{g}^* \tilde{\omega} = \tilde{\omega}. \quad (\text{VI.37})$$

Le moment de l'action de  $\tilde{G}$  sur  $\tilde{X}$  est donc donné, grâce à ce qui précède, par :

$$\tilde{\mu} : \tilde{x} \mapsto \tilde{x}^* \tilde{\omega}. \quad (\text{VI.38})$$

Mais l'action de  $\tilde{G}$  induite sur  $X$  est l'action de  $G$  elle-même, le moment de  $G$  coïncide alors avec celui de  $\tilde{G}$ , autrement dit ;

$$\mu(x) = \tilde{\mu}(\tilde{x}) \quad \text{pour tout } \tilde{x} \mapsto x. \quad (\text{VI.39})$$

On pourrait résumer cette analyse par la proposition suivante :

REMARQUE. Dans le cas des 2-formes  $\omega$  exactes  $\omega = d\varpi$ , la nature de l'application moment est toujours la même : c'est l'application qui à  $x \in X$  associe la 1-forme invariante à gauche  $\hat{x}^* \varpi$  sur le groupe  $G$ , ou si nécessaire sur une extension centrale bien choisie  $\tilde{G}$  de  $G$ . Nous allons voir comment nous pouvons adapter cette proposition au cas général.

#### VI.4 Classes de cohomologie associées au moment

Il est bon de préciser de quels objets dépendent vraiment ces cocycles que nous avons introduits jusqu'à présent, ou plutôt leur classe de cohomologie. Comme nous l'avons déjà remarqué, la classe de cohomologie de  $\Theta$ , le défaut d'équivariance du moment, ne dépend que de  $\omega$  et de l'action de  $G$ , mais elle est indépendante du choix de la primitive  $\varpi$ . Il n'en est pas de même pour la classe de cohomologie du deux-cocycle  $c$ . Elle dépend explicitement du choix de  $\varpi$ . En effet, considérons une autre primitive  $\varpi'$  de  $\omega$ , alors  $\varpi' = \varpi + b$  où  $b$  est une 1-forme fermée quelconque de  $X$ . Puisque  $b$  est fermée et que  $G$  est connexe, il existe une application  $\phi : G \rightarrow C^\infty(X, \mathbf{R})$  telle que  $g^* b = b + d\phi(g)$ , et donc :

$$F'(g) = F(g) + \phi(g), \quad \text{avec } g^* b = b + d\phi(g). \quad (\text{VI.40})$$

Par définition, la fonction  $\phi$  définit un deux-cocycle réel de  $G$ ,

$$e(g, g') = \phi(gg') - g'^*(\phi(g)) - \phi(g'). \quad (\text{VI.41})$$

Puisque  $X$  est connexe, la classe de cohomologie  $[e] \in H^2(G, \mathbf{R})$  ne dépend que de la classe de cohomologie  $[b] \in H_{\text{dR}}^1(X)$  et non du choix de  $\phi$ . Nous venons de définir un homomorphisme par :

$$\chi : H_{\text{dR}}^1(X) \rightarrow H^2(G, \mathbf{R}), \quad \text{avec} \quad \chi([b]) = [e]. \quad (\text{VI.42})$$

Il n'y a aucune raison *a priori* pour que cet homomorphisme soit nul. La classe de cohomologie de  $c$  dépend non seulement de  $\omega$  mais du choix de la primitive  $\varpi$ . Mais on peut remarquer que la classe de  $[c]$ , modulo le sous-espace vectoriel  $\chi(H_{\text{dR}}^1(X)) \subset H^2(G, \mathbf{R})$ , est justement caractérisée par la classe  $[\Theta] \in H^1(G, \mathfrak{G}^*)$ .

Pour comprendre davantage la nature de ces objets, introduisons un cadre un peu plus général et plus géométrique. Considérons toujours une variété différentiable connexe  $X$ , munie d'une action différentiable d'un groupe de Lie connexe  $G$ . Considérons un deux-cocycle  $c$  de  $G$  à valeurs dans un groupe abélien  $A$  et son extension centrale associée, définie par la multiplication

$$(g, a) \cdot (g', a') = (gg', a + a' + c(g, g')), \quad g, g' \in G, \quad a, a' \in A.$$

Nous allons réaliser ce groupe comme un groupe de difféomorphismes, comme un sous-groupe d'automorphismes du *fibré d'homologie* de  $X$ . Commençons par construire ce fibré.

Soit  $C_k(X, A)$  l'ensemble des  $k$ -chaînes simpliciales de  $X$  à valeurs dans le groupe abélien  $A$ , et  $\partial_k$  l'opérateur *bord* de  $C_k(X, A)$  dans  $C_{k-1}(X, A)$ . Rappelons qu'une  $k$ -chaîne simpliciale est une combinaison linéaire de simplexes différentiables formelle à coefficients dans  $A$ , et que les groupes d'homologie  $H_k(X, A)$  sont définis par :

$$H_k(X, A) = Z_k(X, A) / B_k(X, A), \quad (\text{VI.43})$$

avec

$$Z_k(X, A) = \ker \partial_k \quad \text{et} \quad B_k(X, A) = \text{im } \partial_{k+1}.$$

Les éléments de  $Z_k(X, A)$  sont appelés  $k$ -cycles simpliciaux de  $X$  à valeurs dans  $A$ , et ceux  $B_k(X, A)$  les  $k$ -bords. On a en particulier pour les 0-chaînes :

$$c = \sum_x n_x x, \quad n_x \in A, \quad x \in X, \quad \partial_0 c = 0, \quad (\text{VI.44})$$

où les  $n_x$  sont tous nuls à l'exception d'un nombre fini ; et pour les 1-chaînes :

$$c = \sum_\gamma n_\gamma \gamma \quad n_\gamma \in A, \quad \gamma \in \text{Arc}(X), \quad \partial_1 c = \sum_\gamma n_\gamma (\gamma(1) - \gamma(0)). \quad (\text{VI.45})$$

Puisque  $B_1(X, A) \subset Z_1(X, A)$ , le sous-groupe quotient  $C_1(X, A)/B_1(X, A)$  fibre sur  $C_1(X, A)/Z_1(X, A)$ , avec pour fibre  $Z_1(X, A)/B_1(X, A)$ , c'est-à-dire  $H_1(X, A)$  ; mais  $C_1(X, A)/Z_1(X, A) = C_1(X, A)/\ker \partial_1$ , c'est-à-dire

$$C_1(X, A)/Z_1(X, A) = \partial_1 C_1(X, A).$$

Par connexité de  $X$ , toute 0-chaîne de la forme  $\sum_{x,y} n_{x,y}(x - y)$  est un élément de  $\text{im } \partial_1$ , puisqu'on peut toujours trouver un arc  $\gamma$  tel que  $x = \gamma(1)$  et  $y = \gamma(0)$ . Ainsi, on montre que  $C_1(X, A)/Z_1(X, A)$  est le groupe abélien libre engendré par le produit direct  $X \times X$ , en associant à  $\sum_{x,y} n_{x,y}(x - y)$ , l'élément  $\sum_{x,y} n_{(x,y)}(x, y)$ . On a donc construit, au-dessus de  $C_0(X \times X, A)$ , un fibré principal de groupe structural  $A$ . Pointons  $X$  par une origine  $o$  et considérons l'injection de  $X$  dans  $C_0(X \times X, A)$  définie par  $\hat{o} : x \mapsto (o, x)$ . Soit  $\hat{X}$  l'image réciproque, par  $\hat{o}$ , de la fibration principale  $C_1(X, A)/B_1(X, A) \rightarrow C_1(X, A)/Z_1(X, A)$  :

$$\begin{array}{ccc} \hat{X} & \longrightarrow & C_1(X, A)/B_1(X, A) \\ H_1(X, A) \downarrow & & \downarrow H_1(X, A) \\ X & \longrightarrow & C_1(X, A)/Z_1(X, A) \end{array} \quad (\text{VI.46})$$

Nous ne démontrerons pas la proposition suivante :

**PROPOSITION.** *L'image réciproque, par l'injection  $\hat{o} : x \mapsto (o, x)$ , du fibré principal  $C_1(X, A)/B_1(X, A) \rightarrow C_1(X, A)/Z_1(X, A)$  est une variété différentiable connexe  $\hat{X}_A$  fibrée principalement sur  $X$ , de groupe structural  $H_1(X, A)$ . On l'appelle le revêtement d'homologie de  $X$ .*

La fonction de ce revêtement est de rendre exacte toute 1-forme fermée. Choisissons  $A = \mathbf{Z}$ , et notons simplement  $\hat{X} = \hat{X}_{\mathbf{Z}}$  et  $\pi$  sa projection sur  $X$  :

$$\pi : \hat{X} \rightarrow X, \quad \pi : x - o + c' \mapsto x, \quad \text{avec} \quad \partial c' = 0. \quad (\text{VI.47})$$

Soit  $\alpha$  une 1-forme fermée de  $X$ , considérons l'homomorphisme de  $h_\alpha : C_1(X, \mathbf{Z}) \rightarrow \mathbf{R}$  obtenu par intégration de  $\alpha$  le long des chaînes :

$$h_\alpha : C_1(X, \mathbf{Z}) \rightarrow \mathbf{R}, \quad h_\alpha \left( \sum_{\gamma} n_{\gamma} \gamma \right) = \sum_{\gamma} n_{\gamma} \int_{\gamma} \alpha. \quad (\text{VI.48})$$

Pour toute variation  $\delta c$  de la chaîne  $c$ , on a grâce, à la formule de Stokes, et parce que  $\alpha$  est fermée :

$$\delta h_\alpha(c) = \sum_{\gamma} n_{\gamma} [\alpha(\delta\gamma(1)) - \alpha(\delta\gamma(0))]. \quad (\text{VI.49})$$

Si la chaîne  $c$  représente  $\hat{x} \in \hat{X}$ , alors  $c \sim x - o + c'$  avec  $\partial_1 c' = 0$ ; on déduit de ce qui précède que :

$$\delta h_\alpha(\hat{x}) = \alpha(\delta x) \quad \Leftrightarrow \quad dh_\alpha = \pi^* \alpha. \quad (\text{VI.50})$$

On a intégré de cette façon  $\alpha$  sur le revêtement d'homologie  $\hat{X}$ . Le groupe  $H_1(X, \mathbf{Z})$  agit additivement sur  $H_1(X, \mathbf{R})$  par inclusion de  $\mathbf{Z}$  dans  $\mathbf{R}$ , les chaînes à coefficients dans  $\mathbf{Z}$  étant en particulier à coefficients dans  $\mathbf{R}$ . Cette action peut ne pas être libre, et son noyau est appelé le groupe de torsion de  $H_1(X, \mathbf{Z})$ . Mais on peut vérifier que :

$$H_1(X, \mathbf{R})/H_1(X, \mathbf{Z}) = H_1(X, S^1). \quad (\text{VI.51})$$

Considérons alors le produit fibré :

$$\hat{X} \times_{H_1(X, \mathbf{Z})} H_1(X, \mathbf{R}) \rightarrow X. \quad (\text{VI.52})$$

Puisque l'action de  $H_1(X, \mathbf{Z})$  est libre sur  $\hat{X}$  et que ces groupes sont commutatifs, cette fibration est principale de groupe  $H_1(X, \mathbf{R})$ . Mais le groupe  $H_1(X, \mathbf{R})$  étant en réalité un espace vectoriel sur  $\mathbf{R}$ , cette fibration principale est donc triviale. Il

existe une section de cette projection, ou, ce qui revient au même, une application  $T : \hat{X} \rightarrow H_1(X, \mathbf{R})$  équivariante sous l'action de  $H_1(X, \mathbf{Z})$  :

$$\hat{T}(\hat{x} + k) = \hat{T}(x) + k, \quad k \in H_1(X, \mathbf{Z}). \quad (\text{VI.53})$$

On en déduit ainsi l'existence d'une application

$$T : X \rightarrow H_1(X, S^1) = H_1(X, \mathbf{R})/H_1(X, \mathbf{Z})$$

telle que le diagramme suivant commute :

$$\begin{array}{ccc} \hat{X} & \xrightarrow{\hat{T}} & H_1(X, \mathbf{R}) \\ \pi \downarrow & & \downarrow \pi_1 \\ X & \xrightarrow{T} & H_1(X, S^1) \end{array} \quad (\text{VI.54})$$

Supposons maintenant que  $H_1(X, \mathbf{R})$  soit de dimension finie  $m$  ; alors  $H_1(X, S^1)$  est un tore  $\mathbf{T}^m$ , un groupe dont l'algèbre de Lie est justement  $H_1(X, \mathbf{R})$ . Soit  $\theta$  sa forme de Maurer-Cartan ; elle est fermée mais non exacte puisque c'est la projection de l'identité de  $H_1(X, \mathbf{R})$ . Son image réciproque sur  $X$  est une 1-forme fermée que nous noterons  $\Lambda$  :

$$\Lambda = T^*\theta, \quad \theta = \pi_{1*}\mathbf{1}, \quad \theta \in Z_{\text{dR}}^1(X, H_1(X, \mathbf{R})). \quad (\text{VI.55})$$

REMARQUE. Cette 1-forme n'est pas définie de façon unique à cause du choix de  $\hat{T}$ , mais elle a un caractère universel dans la mesure où elle représente, en un seul objet, les classes de cohomologie de toutes les 1-formes fermées de  $X$ . Considérons une 1-forme fermée  $\alpha$  et soit  $h_\alpha$  l'homomorphisme qu'elle induit sur  $H_1(X, \mathbf{R})$  par intégration, alors la 1-forme différentielle fermée :

$$\Lambda_\alpha = h_\alpha \circ \Lambda : \hat{x} \mapsto \Lambda_\alpha(\hat{x}) = \int_\alpha \hat{T}(\hat{x}) \quad (\text{VI.56})$$

lui est cohomologue. En effet, l'application  $h_\alpha \circ \hat{T} - h_\alpha$  est invariante sous l'action de  $H_1(X, \mathbf{Z})$  : soit  $\hat{x} \in \hat{X}$  et  $k \in H_1(X, \mathbf{Z})$ ,  $h_\alpha[\hat{T}(\hat{x} + k)] - h_\alpha(\hat{x} + k) = h_\alpha[\hat{T}(\hat{x}) + k] - h_\alpha(\hat{x}) - h_\alpha(k) = h_\alpha[\hat{T}(\hat{x})] + h_\alpha(k) - h_\alpha(\hat{x}) - h_\alpha(k) = h_\alpha[\hat{T}(\hat{x})] - h_\alpha(\hat{x})$ .

L'application  $h_\alpha \circ \hat{T} - h_\alpha$  est donc définie sur  $X$ . Soit  $f_\alpha(x) = h_\alpha[\hat{T}(\hat{x})] - h_\alpha(\hat{x})$ , pour tout  $\hat{x}$  au dessus de  $x$ , c'est-à-dire  $\int_{\hat{x}} \Lambda_\alpha = \int_{\hat{x}} \alpha + f_\alpha(x)$ ; dérivée, cette identité donne :  $\Lambda_\alpha = \alpha + df_\alpha$ . D'autre part, soit  $\chi : H_1(X, \mathbf{R}) \rightarrow \mathbf{R}$  un homomorphisme et  $\alpha = \chi \circ T$ , c'est une 1-forme fermée vérifiant  $h_\alpha = \chi$ . Cette 1-forme universelle  $\Lambda$  nous a permis de démontrer la proposition suivante :

**THÉORÈME. (DE RHAM)** *L'homomorphisme défini sur  $H_{\text{dR}}^1(X, \mathbf{R})$  par intégration des formes, à valeurs dans  $H^1(X, \mathbf{R}) = \text{Hom}(H_1(X, \mathbf{R}), \mathbf{R})$ , est un isomorphisme.*

C'est un cas particulier du véritable théorème de De Rham, concernant les formes fermées de tout degré.

Revenons maintenant au revêtement d'homologie  $\hat{X}$ , et au groupe  $G$  agissant sur  $X$ . Considérons le sous-groupe des difféomorphismes de  $\hat{X}$  commutant avec l'action de  $H_1(X, \mathbf{Z})$ ; c'est le groupe des automorphismes de  $\hat{X}$ , que l'on note :

$$\text{Aut}(\hat{X}) = \{\phi \in \text{Diff}(\hat{X}) \mid \phi(\hat{x} + k) = \phi(\hat{x}) + k, k \in H_1(X, \mathbf{Z})\}. \quad (\text{VI.57})$$

La projection de  $\text{Aut}(\hat{X})$  sur  $\text{Diff}(X)$  a pour noyau  $H_1(X, \mathbf{Z})$ ; en effet, la projection de  $\hat{X}$  sur  $X$  étant une fibration principale, l'injection  $(k, \hat{x}) \mapsto (\hat{x}, \hat{x} + k)$  est une immersion, donc si  $\phi$  est dans le noyau de la projection  $\text{Aut}(\hat{X}) \rightarrow \text{Diff}(X)$ , l'application  $\hat{x} \mapsto \kappa(\hat{x})$  telle que  $\phi(\hat{x}) = \hat{x} + \kappa(\hat{x})$  est différentiable; mais  $\kappa$  est à valeurs dans  $H_1(X, \mathbf{Z})$  qui est discret; puisque  $\hat{X}$  est connexe, cette application est constante. Puisque par définition de  $\text{Aut}(\hat{X})$ ,  $H_1(X, \mathbf{Z})$  commute avec tous ses éléments, on en déduit une extension centrale :

$$\mathbf{1} \rightarrow H_1(X, \mathbf{Z}) \rightarrow \text{Aut}(\hat{X}) \rightarrow \text{Diff}(X). \quad (\text{VI.58})$$

On en déduit alors, parce que  $G$  est connexe et parce que  $\hat{X}$  est un revêtement de  $X$ , une extension centrale de  $G$  par  $H_1(X, \mathbf{Z})$  :

$$\mathbf{1} \rightarrow H_1(X, \mathbf{Z}) \rightarrow \hat{G} \rightarrow G \rightarrow \mathbf{1}. \quad (\text{VI.59})$$

Nous pouvons lui associer par produit fibré une *extension centrale* de  $G$  par  $H^1(X, \mathbf{R})$  :  $\hat{G}_{\mathbf{R}} = \hat{G} \times_{H_1(X, \mathbf{Z})} H_1(X, \mathbf{R})$ . À cette nouvelle extension centrale

peut être associé un 2-cocycle de  $G$  à valeurs dans l'espace vectoriel  $H_1(X, \mathbf{R})$ . Nous pouvons d'ailleurs le deviner. Soit  $g \in G$ , nous savons, puisque  $G$  est connexe et que  $\Lambda$  est fermée, que  $g^*\Lambda - \Lambda$  est exacte. Nous pouvons définir ici, comme dans le cas des formes différentielles réelles avec lequel il n'y a aucune différence essentielle :

$$\Phi \in C^\infty(G \times X, H_1(X, \mathbf{R})), \quad g^*\Lambda = \Lambda + d\Phi(g, \cdot). \quad (\text{VI.60})$$

De façon identique au cas réel, cette fonction  $\Phi$  définit un 2-cocycle  $C$  :

$$C \in H^2(G, H_1(X, \mathbf{R})), \quad \Phi(gg') = g'^*\Phi(g) + \Phi(g') + C(g, g'). \quad (\text{VI.61})$$

EXERCICE. Vérifier que ce 2-cocycle  $C$  définit l'extension centrale  $\hat{G}$  de  $G$ .

Soit  $\alpha$  une 1-forme différentielle réelle fermée,  $F_\alpha$  la fonction exprimant le défaut d'invariance de  $\alpha$  et  $c_\alpha$  le deux-cocycle réel de  $G$  associé. Les relations entre, d'une part,  $\Phi$  et  $F_\alpha$  et, d'autre part,  $C$  et  $c_\alpha$  sont évidemment obtenues par intégration :

$$F_\alpha(g, x) = \int_{\Phi(g, x)} \alpha, \quad c_\alpha(g, g') = \int_{C(g, g')} \alpha. \quad (\text{VI.62})$$

Il suffit de se souvenir que  $\Lambda_\alpha = h_\alpha \circ \Lambda$  est cohomologue à  $\alpha$ . L'extension  $\hat{G}_\mathbf{R}^\alpha$  de  $G$ , associée à  $c_\alpha$ , est le quotient de  $\hat{G}_\mathbf{R}$  par le noyau de  $h_\alpha$  :

$$\begin{array}{ccc} \hat{G}_\mathbf{R} & \longrightarrow & \hat{G}_\mathbf{R}^\alpha = \hat{G}_\mathbf{R} / \ker h_\alpha \\ & \searrow H_1(X, \mathbf{R}) & \swarrow H_1(X, \mathbf{R}) / \ker h_\alpha \\ & & G \end{array} \quad (\text{VI.63})$$

Voilà donc l'interprétation géométrique de cette classe de cohomologie  $\sigma$  associée au choix d'une primitive particulière de la 2-forme fermée  $\omega$ .

## VI.5 Le cas général des 2-formes fermées quelconques

Nous avons vu que l'application moment est tout à fait naturelle lorsque l'on considère l'action d'un groupe de Lie  $G$  sur une variété  $X$  préservant une 1-forme

$\varpi$ . C'est l'application définie sur  $X$  à valeurs dans le dual  $\mathcal{G}^*$  de l'algèbre de Lie  $\mathcal{G}$  de  $G$  qui au point  $x \in X$  associe la 1-forme invariante à gauche  $\hat{x}^*\varpi$ , où  $\hat{x}$  est l'application orbite  $\hat{x} : g \mapsto g(x)$ . Nous avons pu encore remarquer que, si la 1-forme  $\varpi$  n'est pas invariante par  $G$  mais que sa dérivée extérieure  $\omega = d\varpi$  l'est, on pouvait définir aussi le moment de façon analogue, en considérant non plus la 1-forme  $\varpi$  mais la forme de connexion  $\tilde{\varpi} = \varpi + ds$ , définie sur le fibré principal trivial  $Y = X \times \mathbf{R}$ . Le groupe agissant sur  $Y$  étant l'extension centrale  $\tilde{G}$  de  $G$  par  $\mathbf{R}$  définie grâce au cocycle  $c$  et associée au défaut d'invariance  $F$ , défini par  $g^*\varpi = \varpi + dF(g)$ . Comme nous avons pu le constater, cette extension centrale est le groupe des automorphismes du fibré principal  $(Y, \tilde{\varpi})$ . Le moment est alors obtenu comme l'application qui à un point  $y = (x, t) \in Y$  associe la 1-forme invariante à gauche  $y^*\tilde{\varpi}$ , élément du dual  $\tilde{\mathcal{G}}^*$  de l'algèbre de Lie  $\tilde{\mathcal{G}}$  de  $\tilde{G}$ . Le fait que l'extension soit centrale indique que, *a priori*, cette fonction de  $y = (x, t)$  ne dépend que de  $x$ . D'autre part le moment de  $G$ , et non celui de  $\tilde{G}$ , se retrouve en projetant  $\tilde{\mathcal{G}}^*$  sur  $\mathcal{G}^*$  le long de la section  $t = 0$ .

Nous allons voir comment cette construction peut nous servir de guide et comment on peut l'adapter au cas général des 2-formes fermées non exactes. Pour cela nous allons montrer qu'il existe au moins un fibré principal  $Y$  au dessus de  $X$ , structuré par le tore des périodes  $T_\omega$  de la forme  $\omega$ , et muni d'une connexion de courbure  $\omega$ . De tels fibrés seront appelés, par la suite, *fibrés des périodes* de la forme  $\omega$ .

### VI.5.1 Intégration d'une 2-forme fermée

Considérons donc de façon générale une 2-forme fermée  $\omega$  sur une variété différentiable connexe  $X$ . Appelons *groupe des périodes* de la forme  $\omega$  le sous-groupe additif de  $\mathbf{R}$  défini par :

$$P_\omega = \left\{ \int_\sigma \omega \mid \sigma \in H_2(X, \mathbf{Z}) \right\}. \quad (\text{VI.64})$$

DÉFINITION. *Nous appellerons tore des périodes de la forme  $\omega$  le quotient  $T_\omega$  de  $\mathbf{R}$  par son groupe des périodes  $P_\omega$ . On notera  $\text{cl}_\omega$  la projection de  $\mathbf{R}$  sur le tore  $T_\omega$  :*

$$0 \longrightarrow P_\omega \longrightarrow \mathbf{R} \xrightarrow{\text{cl}_\omega} T_\omega \longrightarrow 1. \quad (\text{VI.65})$$

Le tore des périodes  $T_\omega$  est évidemment un groupe abélien, puisque c'est le quotient du groupe additif  $\mathbf{R}$  par le sous-groupe  $P_\omega$ . Nous noterons additivement sa loi de groupe, même si lorsqu'il est égal à  $S^1$ , l'usage veut qu'elle soit notée multiplicativement. Mais le tore des périodes est rarement un groupe de Lie :  $T_\omega$  est une variété uniquement dans l'un des deux cas suivants :

- 1) le groupe des périodes est trivial :  $P_\omega = \{0\}$ , alors  $T_\omega = \mathbf{R}$  et la forme  $\omega$  est exacte
- 2) le groupe des périodes est isomorphe à  $\mathbf{Z}$ , il existe un réel  $a \neq 0$  tel que  $P_\omega = a\mathbf{Z}$ , alors  $T_\omega \simeq S^1$ .

Dans ces deux cas nous dirons que la forme  $\omega$  est *entière*. Dans les autres cas, c'est-à-dire lorsque le groupe  $P_\omega$  est engendré par au moins deux éléments, le tore des périodes sera muni de sa structure de *groupe différentiable*<sup>3</sup>. En particulier, une application différentiable d'une variété  $V$  dans le tore  $T_\omega$  est la projection, au moins localement, d'une application différentiable de  $V$  dans  $\mathbf{R}$ . Autrement dit :  $\varphi : V \rightarrow T_\omega$  est dite différentiable si et seulement si, pour tout  $x \in V$ , il existe un voisinage  $U \subset V$  de  $x$  et  $\psi : U \rightarrow \mathbf{R}$ , différentiable, telle que  $\varphi|_U = \text{cl}_\omega \circ \psi$ .

Toute forme différentielle invariante sur  $T_\omega$  est proportionnelle à  $\theta$ , sa 1-forme de Maurer-Cartan, définie par :

$$\text{cl}_\omega^* \theta = dt. \quad (\text{VI.66})$$

Comme le tore des périodes  $T_\omega$ , les objets suivants doivent être compris au sens de la théorie des *espaces différentiables*.

Nous noterons  $\text{Arc}(X)$  l'espace des arcs d'une variété (ou d'un espace différentiable)  $X$ , c'est-à-dire l'espace des applications différentiables de  $\mathbf{R}$  dans  $X$ . Nous munirons l'espace des arcs de  $X$  de sa structure différentiable fonctionnelle. Ses *plaques* (ou *paramétrages différentiables*) sont les applications  $\varphi : U \rightarrow \text{Arc}(X)$ , où  $U$  est un ouvert de  $\mathbf{R}^p$ , telles que  $(r, t) \mapsto \varphi(r)(t)$  définie sur  $U \times \mathbf{R}$  dans  $X$  soit différentiable. Nous définirons les applications *source* et *but* :

$$\begin{aligned} \text{source} : \text{Arc}(X) &\rightarrow X & \text{avec} & \text{source}(\gamma) = \gamma(0) \\ \text{but} : \text{Arc}(X) &\rightarrow X & \text{avec} & \text{but}(\gamma) = \gamma(1). \end{aligned} \quad (\text{VI.67})$$

<sup>3</sup>Voir annexe D pour la définition précise de la structure d'espace et de groupe différentiable.

Ces applications sont évidemment différentiables. Pour toute plaque  $\varphi : U \rightarrow \text{Arc}(X)$  la composée  $\text{but} \circ \varphi : U \rightarrow X$  est :  $\text{but} \circ \varphi(r) = \varphi(r)(1)$  qui est la restriction à  $U \times \{1\}$  d'une application différentiable définie sur  $U \times \mathbf{R}$  (de même pour l'application source).

Nous noterons  $\text{Arc}(X, x_o)$  l'espace des arcs pointés en  $x_o$ , c'est-à-dire le sous-espace des arcs de  $X$  tels que  $\gamma(0) = x_o$  :

$$\text{Arc}(X, x_o) = \{\gamma \in \text{Arc}(X) \mid \text{source}(\gamma) = x_o\}. \quad (\text{VI.68})$$

Sa structure différentiable est évidemment sa structure de partie de  $\text{Arc}(X)$ . Mais il faut noter que cet espace est contractile, ce qui est une propriété utilisée de façon importante dans la suite. En effet, l'application :

$$\rho : \mathbf{R} \times \text{Arc}(X, x_o) \rightarrow \text{Arc}(X, x_o) \quad \rho(s)(\gamma) = [t \mapsto \gamma(st)], \quad (\text{VI.69})$$

est une rétraction de déformation de  $\text{Arc}(X, x_o)$  sur l'arc constant  $t \mapsto x_o$ .

Nous noterons aussi  $\text{Lac}(X, x_o)$  l'espace des lacets de  $X$  basés en  $x_o$ . De même que  $\text{Arc}(X, x_o)$ , il est muni de sa structure différentiable de partie de  $\text{Arc}(X)$ .

Nous dirons que deux arcs  $\gamma$  et  $\gamma'$  sont *homologues* si leur différence borde une 2-chaîne singulière ; nous noterons :

$$\gamma \sim \gamma' \Leftrightarrow \exists \sigma \quad \partial \sigma = \gamma' - \gamma. \quad (\text{VI.70})$$

Le quotient des arcs pointés par la relation d'homologie sera noté  $\hat{X}$  : c'est le *revêtement d'homologie* de  $X$ . Son groupe structural, quotient de  $\text{Lac}(X, x_o)$  par la relation d'homologie, est isomorphe au groupe  $H_1(X, \mathbf{Z})$ . En particulier si  $H_1(X, \mathbf{Z}) = \{0\}$  alors deux arcs pointés en  $x_o$  ont même but si et seulement si ils bordent une 2-chaîne.

Nous avons introduit les objets nécessaires à la démonstration de la proposition suivante.

**PROPOSITION.** *Soit  $\omega$  une 2-forme fermée sur une variété différentiable connexe  $X$ , dont le premier groupe d'homologie entière est nul :  $H_1(X, \mathbf{Z}) = 0$ . Il existe alors un fibré principal  $\pi : Y \rightarrow X$ , de base  $X$ , de groupe structural le tore des périodes  $T_\omega$ , muni d'une forme de connexion  $\lambda$  de courbure  $\omega$ . La structure  $(Y, \lambda)$  est unique à équivalence près.*

DÉMONSTRATION. *Démontrons d'abord l'existence. Relevons la relation d'équivalence, définie plus haut sur les arcs, au produit  $\text{Arc}(X, x_o) \times T_\omega$  :*

$$(\gamma, z) \sim (\gamma', z') \Leftrightarrow \gamma \sim \gamma' \text{ et } z' = z + \text{cl}_\omega \int_\sigma \omega \quad \text{où } \partial\sigma = \gamma' - \gamma. \quad (\text{VI.71})$$

*Il est facile de vérifier que c'est bien une relation d'équivalence. Soit  $Y$  le quotient de  $\text{Arc}(X, x_o) \times T_\omega$  par cette relation d'équivalence. Puisque  $H_1(X, \mathbf{Z}) = 0$  alors  $\hat{X} = X$  et  $Y$  est fibré principalement sur  $X$ , de groupe structural  $T_\omega$ . Soit  $K$  l'opérateur de chaîne-homotopie tel qu'il est défini dans l'annexe E, et  $\alpha$  la 1-forme définie sur le produit  $\text{Arc}(X, x_o) \times T_\omega$  par :*

$$\alpha = K\omega \oplus \theta. \quad (\text{VI.72})$$

*C'est évidemment une forme de connexion pour l'action naturelle de  $T_\omega$ , de courbure  $\text{but}^*\omega$ , puisque  $K\omega$  est une primitive de  $\text{but}^*\omega$ , où  $\text{but}$  désigne l'application  $\text{but}$  de  $\text{Arc}(X, x_o)$  sur  $X$  qui à  $\gamma$  associe  $\gamma(1)$ . En appliquant la proposition sur les quotients de formes différentiables (voir page 182), le passage de la forme  $\alpha$  au quotient  $Y$  est assuré par le lemme suivant :*

LEMME. *Soit  $p$  la projection de  $\text{Arc}(X, x_o) \times T_\omega$  sur son quotient  $Y$ . Si deux paramétrages différentiables (plaques)  $P$  et  $P'$  de  $\text{Arc}(X, x_o) \times T_\omega$  vérifient  $p \circ P = p \circ P'$ , alors  $P^*\alpha = P'^*\alpha$ .*

DÉMONSTRATION. *Soit  $U$  le domaine des plaques  $P$  et  $P'$  et  $r \in U$ , notons  $P(r) = (P_X(r), P_T(r))$  où  $P_X$  est une plaque de  $\text{Arc}(X, x_o)$  et  $P_T$  une plaque de  $T_\omega$ . On ne perd rien en généralité en supposant que  $P_T$  se relève globalement sur  $\mathbf{R}$  en une plaque  $Q \in C^\infty(U, \mathbf{R})$  (i.e.  $P_T = \text{cl}_\omega \circ Q$ ). On a donc  $P^*(K\omega \oplus \theta) = P_X^*(K\omega) + P_T^*\theta$ , c'est-à-dire  $P_X^*(K\omega) + dQ$ . Notons  $\gamma = P(r)$  et  $\delta\gamma = D(P)(r)(\delta r)$  où  $\delta r$  est un vecteur tangent à  $r \in U$  ; par définition (cf. annexe E) :*

$$P_X^*(K\omega)_r(\delta r) = \int_0^1 \omega_{\gamma(t)}(\dot{\gamma}(t), \delta\gamma(t)) dt.$$

*D'autre part,  $p \circ P = p \circ P'$  implique l'existence, pour tout  $r \in U$ , d'une 2-chaîne  $\sigma_r$  telle que  $\partial\sigma_r = P'(r) - P(r)$ . On peut vérifier, en restreignant si*

nécessaire le domaine  $U$ , que  $\sigma_r$  peut être choisie différenciablement. On a alors

$$\text{cl}_\omega(Q'(r)) = \text{cl}_\omega(Q(r)) + \text{cl}_\omega \int_{\sigma_r} \omega,$$

c'est-à-dire  $Q'(r) = Q(r) + \int_{\sigma_r} \omega + l(r)$ , où  $l : U \rightarrow P_\omega$  est différentiable. Comme  $P_\omega$  est différenciablement discret toute application différentiable de  $U$  dans  $\mathbf{R}$  à valeur dans  $P_\omega$  est localement constante, et on en déduit que  $l$  est constante et donc que :

$$dQ'(r) = dQ(r) + d \left[ \int_{\sigma_r} \omega \right].$$

En utilisant le théorème de Stokes (voir annexe F), qui peut s'écrire aussi

$$d \left[ \int_{\sigma} \omega \right] = \int_{\sigma} d\omega(\cdot) + \int_{\partial\sigma} \omega(\cdot),$$

et l'expression précédente de  $P_X^*(K\omega)$ , on déduit finalement :

$$dQ' - dQ = -P_X^*(K\omega) + P_X'^*(K\omega) \Rightarrow P^*[K\omega \oplus \theta] = P'^*[K\omega \oplus \theta].$$

C'est ce qu'il fallait démontrer. ■

Démontrons l'unicité de cette construction. Soit  $(Y', \lambda')$  une autre structure répondant aux hypothèses. L'image réciproque de  $Y$  par  $\text{but} : \text{Arc}(X, x_o) \rightarrow X$  est triviale, car  $\text{Arc}(X, x_o)$  est contractile et muni d'une connexion (cette propriété remplace la paracompacité dans le cas des variétés). L'image de la section nulle au-dessus de  $\text{Arc}(X, x_o)$  définit une application  $\psi$  de  $\text{Arc}(X, x_o)$  dans  $Y$ . On vérifie alors que la projection  $(\gamma, z) \mapsto z\psi(\gamma)$  réalise le quotient de  $\text{Arc}(X, x_o) \times T_\omega$  par la relation d'équivalence définie plus haut et donc que  $Y$  et  $Y'$  sont équivalents en tant que fibrés principaux. Il est toujours possible, d'autre part, de choisir un isomorphisme entre  $\text{but}^*(Y')$  et le produit  $\text{Arc}(X, x_o) \times T_\omega$  de telle sorte que l'image réciproque de  $\lambda'$  coïncide avec la forme  $\alpha$ . ■

REMARQUE. Le relevé à  $\text{Arc}(X, x_o)$  de la relation d'équivalence sur les arcs définit un cocycle qu'on pourrait appeler le cocycle d'arcs associé à  $\omega$ . C'est la fonction  $f_\omega$  définie sur les couples d'arcs homologues par :

$$f_\omega(\gamma, \gamma') = \text{cl}_\omega \int_{\sigma} \omega \quad \text{où} \quad \partial\sigma = \gamma' - \gamma. \quad (\text{VI.73})$$

L'intégrale dépend de la chaîne bordant  $\gamma' - \gamma$ , mais pas sa classe dans  $T_\omega$ . Le choix de cette terminologie résulte de l'identité suivante, vérifiée pour tout triplet d'arcs homologues :

$$f_\omega(\gamma, \gamma') + f_\omega(\gamma', \gamma'') + f_\omega(\gamma'', \gamma) = 0. \quad (\text{VI.74})$$

La fonction  $f_\omega$  est un *cobord* s'il existe  $\mu : \text{Arc}(X, x_\circ) \rightarrow T_\omega$  différentiable, telle que  $f_\omega(\gamma, \gamma') = \mu(\gamma') - \mu(\gamma)$ , ce qui se produit lorsque  $\omega = d\alpha$ , on a alors  $\mu(\gamma) = \int_\gamma \alpha$ .

Nous utiliserons le lemme précédent pour démontrer le théorème plus général suivant :

**THÉORÈME.** *Pour toute 2-forme fermée  $\omega$  définie sur une variété différentiable connexe  $X$ , il existe un fibré principal  $\pi : Y \rightarrow X$  de groupe structural  $T_\omega$  muni d'une forme de connexion  $\lambda$  de courbure  $\omega$ . Un tel fibré sera appelé fibré des périodes de la 2-forme  $\omega$ . Les fibrés des périodes de la 2-forme  $\omega$  sont classés, à équivalence de fibré principal près, par le premier groupe d'extension  $\text{Ext}(H_1(X, \mathbf{Z}), P_\omega)$ .*

**DÉMONSTRATION.** *Comme dans la démonstration précédente, nous considérons la relation d'équivalence définie par (VI.71) sur le produit  $\text{Arc}(X, x_\circ) \times T_\omega$ , ainsi que la 1-forme  $\alpha$  (définition VI.72). L'espace quotient  $\hat{Y}$  est alors fibré sur le revêtement d'homologie  $\hat{X}$  de  $X$ , de groupe  $T_\omega$ . Il est muni d'une connexion  $\hat{\lambda}$  de courbure  $\hat{\omega}$ , image réciproque de  $\omega$  par la projection naturelle  $\hat{X} \rightarrow X$ . En vertu de la proposition précédente, le fibré des périodes  $(\hat{Y}, \hat{\lambda})$  est unique à équivalence près.*

*Considérons une section  $s$  de  $H_1(X, \mathbf{Z})$  dans l'espace des lacets  $\text{Lac}(X, x_\circ)$ . Soit  $\phi$  le 2-cocycle de groupe défini sur  $H_1(X, \mathbf{Z})$ , à valeurs dans  $T_\omega$  :*

$$\phi(h, h') = f_\omega(s(h + h'), s(h) + s(h')). \quad (\text{VI.75})$$

*Ce cocycle étant symétrique, il définit une extension abélienne  $\Gamma$  de  $H_1(X, \mathbf{Z})$  par  $T_\omega$ . Cette extension agit naturellement sur  $\hat{Y}$  par :*

$$(h, \tau)[\gamma, z] = [s(h) + \gamma, z + \tau], \quad (\text{VI.76})$$

où  $l+\gamma$  désigne la juxtaposition du lacet  $l$  avec l'arc  $\gamma$ , et les crochets indiquent les classes d'équivalences. Mais le groupe  $T_\omega$  étant divisible, l'extension est triviale (voir par exemple [KM]) et le groupe  $\Gamma$  est isomorphe au produit direct de  $H_1(X, \mathbf{Z})$  par  $T_\omega$ . Tout isomorphisme permet de définir une action de  $H_1(X, \mathbf{Z})$  sur  $\hat{Y}$ , dont le quotient  $Y = \hat{Y}/H_1(X, \mathbf{Z})$  est un fibré principal de groupe structural  $T_\omega$  sur  $X$ . On vérifie alors que la forme de connexion  $\hat{\lambda}$ , puisqu'elle est invariante par ces actions de  $H_1(X, \mathbf{Z})$ , passe sur  $Y$  en une forme de connexion  $\lambda$  de courbure  $\omega$ . Le fibré  $Y \rightarrow X$ , muni de la forme de connexion  $\lambda$  est construit par quotients successifs :

$$\begin{array}{ccccc} A(X, x) \times T_\omega & \longrightarrow & \hat{Y} & \longrightarrow & Y \\ \downarrow & & \downarrow & & \downarrow \\ A(X, x) & \longrightarrow & \hat{X} & \longrightarrow & X \end{array}$$

Deux tels isomorphismes entre  $\Gamma$  et  $H^1(X, \mathbf{Z})$  ne diffèrent que d'un élément de  $\text{Hom}(H_1(X, \mathbf{Z}), T_\omega)$ , c'est-à-dire d'un élément de  $H^1(X, T_\omega)$ . Grâce à l'unicité du fibré  $(\hat{Y}, \hat{\lambda})$ , nous avons trouvé de cette façon tous les fibrés des périodes de  $(X, \omega)$ .

Soit alors  $\pi : Y \rightarrow X$  et  $\pi' : Y' \rightarrow X$  deux de ces fibrés obtenus par quotient de  $\hat{Y}$  à partir des actions  $\rho$  et  $\rho'$  de  $H_1(X, \mathbf{Z})$ . Ces actions diffèrent d'un élément  $r \in \text{Hom}(H_1(X, \mathbf{Z}), T_\omega)$ . Il y a donc une surjection naturelle  $\sigma$  de l'espace des classes de fibrés des périodes de  $\omega$  sur  $\text{Hom}(H_1(X, \mathbf{Z}), T_\omega)$ . On peut facilement vérifier, en prenant les images réciproques de  $Y$  et  $Y'$  par la projection naturelle de  $\hat{X}$  sur  $X$ , qu'ils sont équivalents si et seulement si il existe une application  $\zeta : \hat{X} \rightarrow T_\omega$  telle que  $r(k) = \zeta(k\hat{x}) - \zeta(\hat{x})$ , pour tout  $k \in H_1(X, \mathbf{Z})$ . La fonction  $\zeta$  définit alors une 1-forme fermée  $\varepsilon$  sur  $X$ , par  $\zeta^*\theta = \pi^*\varepsilon$ , qui définit à son tour (par intégration sur le revêtement d'homologie) une fonction  $F$  de  $\hat{X}$  dans  $\mathbf{R}$ , telle que  $\zeta^*\theta = dF$ ; on en déduit  $\zeta = \text{cl}_\omega \circ F$ . Autrement dit : le noyau de la surjection  $\sigma$  est le sous groupe des homomorphismes de  $H_1(X, \mathbf{Z})$  dans  $T_\omega$  provenant d'un homomorphisme de  $H_1(X, \mathbf{Z})$  dans  $\mathbf{R}$ . Comme  $\mathbf{R}$  est divisible, on conclut, en utilisant la suite exacte du foncteur  $\text{Ext}$  [McL], que le co-noyau de la flèche naturelle  $\text{Hom}(H_1(X, \mathbf{Z}), \mathbf{R}) \rightarrow \text{Hom}(H_1(X, \mathbf{Z}), T_\omega)$  est exactement le

premier groupe d'extension de  $H_1(X, \mathbf{Z})$  dans le groupe des périodes  $P_\omega$ , c'est-à-dire :  $\text{Ext}(H_1(X, \mathbf{Z}), P_\omega)$ . ■

REMARQUE. Si le groupe d'homologie  $H_1(X, \mathbf{Z})$  est sans torsion et si  $X$  est une variété compacte, il y a unicité du fibré des périodes de la 2-forme  $\omega$ . En effet, tout homomorphisme de  $H_1(X, \mathbf{Z})$  dans  $T_\omega$  est alors la projection d'un homomorphisme à valeurs réelles.

REMARQUE. Étant donné un fibré des périodes  $\pi : Y \rightarrow X$ , les formes de connexions inéquivalentes sur  $Y$  de courbure  $\omega$  sont classées par  $H^1(X, \mathbf{R})$ . L'ensemble des couples  $(Y, \lambda)$  qui « intègrent »  $(X, \omega)$  est donc classé, à isomorphisme près, par  $H^1(X, T_\omega)$ . Ce que nous dit le théorème précédent est que cette classification se scinde grâce à la suite exacte :

$$\begin{aligned} 0 &\rightarrow \text{Hom}(H_1(X, \mathbf{Z}), P_\omega) \rightarrow \text{Hom}(H_1(X, \mathbf{Z}), \mathbf{R}) \rightarrow \\ &\rightarrow \text{Hom}(H_1(X, \mathbf{Z}), T_\omega) \rightarrow \text{Ext}(H_1(X, \mathbf{Z}), P_\omega) \rightarrow 0 \end{aligned} \quad (\text{VI.77})$$

Le groupe  $\text{Ext}(H_1(X, \mathbf{Z}), P_\omega)$  classe les fibrés des périodes, et le groupe dual  $\text{Hom}(H_1(X, \mathbf{Z}), \mathbf{R})$ , c'est-à-dire  $H^1(X, \mathbf{R})$ , classe les formes de connexion.

REMARQUE. Parce que les cocycles d'arcs de deux 2-formes fermées cohomologues sont cohomologues, leurs fibrés des périodes sont équivalents. Autrement dit, la classe de cohomologie de la forme  $\omega$  est, en un certain sens, la première « classe de Chern » de ces fibrés des périodes.

REMARQUE. On peut être surpris que des fibrés des périodes existent pour toute forme fermée, même multiples les unes des autres par un réel non entier. Mais, c'est un exercice de montrer que pour  $\omega = s\omega_0$ , où  $s \in \mathbf{R}$ , le fibré des périodes de  $\omega_0$  est équivalent comme fibré à celui de  $\omega$  ; mais attention : les groupes  $T_\omega$  et  $T_{\omega_0}$  ne sont pas identiques, seulement isomorphes. La différence avec la situation classique des 2-formes entières est qu'il n'y a pas d'identification possible, *a priori*, des tores des périodes pour les 2-formes fermées quelconques. Une telle identification, dans le cas entier, consiste à fixer à l'avance la longueur de la période.

### VI.5.2 Les difféomorphismes hamiltoniens revisités

Considérons une variété différentielle  $X$  munie d'une 2-forme fermée  $\omega$  et  $(Y, \lambda)$  un de ses fibrés des périodes ; nous noterons  $\pi : X \rightarrow Y$  la projection. Considérons le groupe des automorphismes  $\text{Aut}(Y, \lambda)$  des difféomorphismes de  $Y$  qui commutent avec l'action du tore des périodes et qui préservent  $\lambda$  :

$$\text{Aut}(Y, \lambda) = \{\varphi \in \text{Diff}(Y) \mid \forall \tau \in T_\omega : \varphi \circ \tau_Y = \tau_Y \circ \varphi, \varphi^* \lambda = \lambda\}. \quad (\text{VI.78})$$

Soit  $\varphi$  un automorphisme de  $(Y, \lambda)$  ; grâce à la commutativité avec l'action de  $T_\omega$ , il existe un difféomorphisme  $\phi$  de  $X$  défini par  $\pi \circ \varphi = \phi \circ \pi$ . Nous dirons que  $\phi$  est la *projection* de  $\varphi$  et noterons  $\phi = \pi_* \varphi$ . Le difféomorphisme  $\phi$  préserve évidemment la courbure  $\omega$  de  $\lambda$  :  $\phi^* \omega = \omega$ .

$$\begin{array}{ccc} Y & \xrightarrow{\varphi} & Y \\ \pi \downarrow & & \downarrow \pi \\ X & \xrightarrow{\phi = \pi_* \varphi} & X \end{array}$$

Le sous-groupe des difféomorphismes de  $X$  préservant la 2-forme  $\omega$  sera notée  $\text{Diff}(X, \omega)$  :

$$\text{Diff}(X, \omega) = \{\phi \in \text{Diff}(X) \mid \phi^* \omega = \omega\}. \quad (\text{VI.79})$$

Le groupe des périodes  $T_\omega$  s'injecte, par son action, dans le groupe des automorphismes de  $(Y, \lambda)$ . Il est clair qu'ainsi  $T_\omega$  est dans le noyau de la projection  $\pi_*$ . Nous avons plus précisément :

PROPOSITION. *Le noyau de la projection  $\pi_* : \text{Aut}(Y, \lambda) \rightarrow \text{Diff}(X, \omega)$  est le tore des périodes  $T_\omega$ . On a donc la suite exacte d'homomorphismes :*

$$0 \longrightarrow T_\omega \longrightarrow \text{Aut}(Y, \lambda) \xrightarrow{\pi_*} \text{Diff}(X, \omega). \quad (\text{VI.80})$$

DÉMONSTRATION. *Soit  $\varphi$  un automorphisme de  $(Y, \lambda)$  se projetant sur l'identité de  $X$ . Puisque  $\pi : Y \rightarrow X$  est une fibration principale, il existe une application*

différentiable  $\tau : Y \rightarrow T_\omega$  telle que  $\varphi(y) = \tau(y).y$ . L'identité  $\varphi^*\lambda = \lambda$  signifie que pour toute plaque  $P$  de  $Y$  on a  $P^*(\varphi^*\lambda) = P^*(\lambda)$ . Considérons alors l'application  $y \mapsto \tau(y).y$  comme la composition des applications :

$$\begin{array}{ccccc} Y & \rightarrow & T_\omega \times Y & \rightarrow & Y \\ y & \mapsto & \begin{pmatrix} \tau(y) \\ y \end{pmatrix} & \mapsto & \tau(y).y \end{array} \quad (\text{VI.81})$$

Nous pouvons écrire :

$$\varphi^*\lambda = \left[ y \mapsto \begin{pmatrix} \tau(y) \\ y \end{pmatrix} \right]^* \circ \left[ \begin{pmatrix} \tau \\ y \end{pmatrix} \mapsto \tau.y \right]^* \lambda. \quad (\text{VI.82})$$

Soit  $\hat{y} : T_\omega \rightarrow Y$  l'application orbite  $\hat{y} : \tau \mapsto \tau.y$ ,  $\tau_Y$  désignant toujours l'action de  $\tau \in T_\omega$  sur  $Y$ , on a :

$$\left( \left[ \begin{pmatrix} \tau \\ y \end{pmatrix} \mapsto \tau.y \right]^* \lambda \right)_{(\tau,y)} = (\hat{y}^*\lambda)_\tau \oplus (\tau_Y^*\lambda)_y = (\hat{y}^*\lambda)_\tau \oplus \lambda_y, \quad (\text{VI.83})$$

On déduit ainsi :

$$\varphi^*\lambda_y = \left[ y \mapsto \begin{pmatrix} \tau(y) \\ y \end{pmatrix} \right]^* ((\hat{y}^*\lambda)_{\tau(y)} \oplus \lambda_y) = (\tau^* \circ \hat{y}^*\lambda)_y + \lambda_y. \quad (\text{VI.84})$$

On déduit ainsi de l'identité  $\varphi^*\lambda = \lambda$  que  $\tau^*(\hat{y}^*\lambda) = 0$  ; mais puisque  $\lambda$  est une forme de connexion sur  $Y : y^*\lambda = \theta$  — où, rappelons-le,  $\theta$  est la forme de Maurer-Cartan sur  $T_\omega$  —, alors la seule solution pour que  $\tau^*\theta = 0$  est que  $\tau = \text{cste}$ . Autrement dit, il existe  $\tau \in T_\omega$  tel que  $\varphi(y) = \tau.y$ . Le noyau de  $\pi_* : \text{Aut}(Y, \lambda) \rightarrow \text{Diff}(X, \omega)$  est donc bien le tore des périodes  $T_\omega$ . ■

On dit aussi, en utilisant le langage des physiciens, que les seules « transformations de jauge » sont les transformations de « première espèce ».

Soit  $\mu_Y$  le moment de  $\lambda$  par rapport à  $\text{Aut}(Y, \lambda)$ , au sens de la définition du moment des 1-formes (page 112). Il est défini sur  $Y$ , à valeurs dans l'espace vectoriel des 1-formes différentielles invariantes par multiplication à gauche<sup>4</sup> du groupe  $\text{Aut}(Y, \lambda)$ . Nous noterons  $\text{aut}^*(Y, \lambda)$  cet espace vectoriel.

$$\mu_Y : Y \rightarrow \text{aut}^*(Y, \lambda) \quad \mu_Y : y \mapsto \hat{y}^*\lambda. \quad (\text{VI.85})$$

<sup>4</sup>La multiplication à gauche dans  $\text{Aut}(Y, \lambda)$  par un élément donné  $f : L_f(\varphi) = \varphi \circ f^{-1}$

DÉFINITION. Nous dirons qu'un difféomorphisme  $\phi$  de  $X$ , préservant  $\omega$ , est un difféomorphisme hamiltonien s'il est l'image d'un automorphisme  $\varphi$  du fibré des périodes  $(Y, \lambda)$ ; nous noterons  $\text{Ham}_Y(X, \omega) \subset \text{Diff}(X, \omega)$  le groupe des difféomorphismes hamiltoniens.

$$\{1\} \longrightarrow T_\omega \longrightarrow \text{Aut}(Y, \lambda) \xrightarrow{\pi^*} \text{Ham}_Y(X, \omega) \longrightarrow \{1\} \quad (\text{VI.86})$$

L'action d'un groupe de Lie  $G$  sur  $X$  préservant  $\omega$  sera dite hamiltonienne si l'image de  $G$  dans  $\text{Diff}(X, \omega)$  est toute entière contenue dans  $\text{Ham}_Y(X, \omega)$ .

Cette définition appelle immédiatement la remarque suivante :

REMARQUE. L'extension décrite ci-dessus (suite exacte VI.86) est centrale<sup>5</sup>, autrement dit, le groupe  $\text{Aut}(Y, \lambda)$  est une extension centrale du groupe  $\text{Ham}_Y(X, \omega)$  par le tore des périodes  $T_\alpha$ .

On en déduit immédiatement que le moment  $\mu_Y$  de  $\text{Aut}(Y, \lambda)$ , défini plus haut, est constant sur les fibres de la projection  $\pi$ . C'est une conséquence immédiate de la commutativité du tore des périodes  $T_\omega \subset \text{Aut}(Y, \lambda)$ . Autrement dit, il existe une application  $\mu_X$  définie sur  $X$  à valeurs dans  $\text{aut}^*(Y, \lambda)$  définie par  $\mu_X(x) = \mu_Y(y)$ , où  $\pi(y) = x$ .

$$\begin{array}{ccc} Y & & \\ \downarrow \pi & \searrow \mu_Y & \\ & \text{aut}^*(Y, \lambda) & \\ & \nearrow \mu_X & \\ X & & \end{array}$$

est une application du groupe  $\text{Aut}(Y, \lambda)$  dans lui-même. À ce titre, on peut considérer l'image réciproque, par  $L_f$ , de toute forme différentielle  $\alpha$  définie sur  $\text{Aut}(Y, \lambda)$ . Ici, les termes sont à considérer au sens des espaces différentiables. Il est clair que les 1-formes différentielles de  $\text{Aut}(Y, \lambda)$  invariantes par multiplication à gauche forment un sous-espace vectoriel de l'espace des 1-formes de  $\text{Aut}(Y, \lambda)$ . On peut l'appeler par commodité le *dual-de-l'algèbre-de-Lie* sans qu'il soit nécessaire de définir *a priori* l'algèbre de Lie de  $\text{Aut}(Y, \lambda)$ . Nous lui préférons le terme d'*espace des moments*.

<sup>5</sup>Rappelons qu'une extension  $1 \rightarrow A \rightarrow \Gamma \rightarrow G \rightarrow 1$  de  $G$  par  $A$  est dite centrale si le groupe  $A$  commute avec  $\Gamma$ , c'est-à-dire si tout élément de  $A$  commute avec tous les éléments de  $\Gamma$ ; ce qui implique, en particulier, que  $A$  est abélien.

Mais cela ne suffit pas encore à définir le moment de l'action de  $\text{Ham}_Y(X, \omega)$  ; en effet  $\mu_X$  est à valeurs dans  $\text{aut}^*(Y, \lambda)$  ; or le moment de  $\text{Ham}_Y(X, \omega)$  doit bien être défini sur  $X$ , mais à valeurs dans  $\text{ham}^*(X, \omega)$ , espace vectoriel des 1-formes invariantes de  $\text{Ham}_Y(X, \omega)$ . Il nous faut alors utiliser la remarque suivante.

REMARQUE. La suite exacte de groupes (VI.86) se renverse pour donner une suite exacte de « duaux d'algèbre de Lie<sup>6</sup> » :

$$0 \longrightarrow \text{ham}_Y^*(X, \omega) \longrightarrow \text{aut}^*(Y, \lambda) \xrightarrow{i^*} \tau_\omega^* \longrightarrow 0 \quad (\text{VI.87})$$

En effet, l'injection  $i$  de  $T_\omega$  dans  $\text{Aut}(Y, \lambda)$  induit naturellement, par restriction, une projection de  $\text{aut}^*(Y, \lambda)$  sur  $\tau_\omega^*$ . Le noyau de cette projection est bien constitué des formes différentielles qui « s'annulent » sur les fibres. Étant invariantes et nulles le long des fibres, elles sont donc issues de la base  $\text{Ham}_Y(X, \omega)$  (voir annexe D).

Cette dernière remarque va nous permettre de définir un moment pour l'action de  $\text{Ham}_Y(X, \omega)$  sur  $X$ . En effet,  $\mu_X(x)$  n'est jamais dans  $\text{ham}_Y^*(X, \omega)$  puisque justement  $i^*(\mu_X(x)) = i^*(\hat{y}^*\lambda) = (\hat{y} \circ i)^*\lambda = \theta$ , où  $\theta$  désigne la forme de Maurer-Cartan du tore  $T_\omega$ . Cela est une conséquence directe<sup>7</sup> de ce que  $\lambda$  est une forme de connexion sur  $Y$ . Ainsi le moment  $\mu_X$  est à valeurs dans un sous-espace affine de  $\text{aut}^*(Y, \lambda)$ , parallèle à  $\text{ham}_Y^*(X, \omega)$ . Il suffit donc de retrancher la valeur du moment  $\mu_X$  en un point quelconque  $o \in X$  pour être à valeurs dans  $\text{ham}_Y^*(X, \omega)$ . Nous définirons donc un moment de l'action de  $\text{Ham}_Y(X, \omega)$  sur  $X$  de la façon suivante.

DÉFINITION. Soit  $o \in X$  un point de  $X$ , nous appellerons moment attaché à  $o$ , du groupe  $\text{Ham}_Y(X, \omega)$ , l'application :  $\mu_o : X \rightarrow \text{ham}_Y^*(X, \omega)$  définie par :

$$\forall x \in X \quad \mu_o(x) = \mu_X(x) - \mu_X(o). \quad (\text{VI.88})$$

Autrement dit : soit  $y_o$  un point de  $Y$  au-dessus de  $o \in X$ , et  $y$  un point au-dessus de  $x$ , le moment  $\mu_o$  a la valeur :

$$\mu_o(x) = \hat{y}^*\lambda - \hat{y}_o^*\lambda. \quad (\text{VI.89})$$

<sup>6</sup>Là encore il faut entendre duaux-d'algèbre-de-Lie, en un seul mot.

<sup>7</sup>Par définition même.

Il faut maintenant se convaincre que cette définition est bien conforme à ce que l'on connaît, par ailleurs, de l'application moment. Nous l'établirons donc dans le cadre de l'action hamiltonienne (au sens défini plus haut) d'un groupe de Lie  $G$  sur  $X$ . Nous avons alors, par définition, une extension centrale :

$$\{1\} \longrightarrow T_\omega \longrightarrow \tilde{G} \xrightarrow{\pi_*} G \longrightarrow \{1\}. \quad (\text{VI.90})$$

Considérons alors un homomorphisme  $\tilde{h} : \mathbf{R} \rightarrow \tilde{G}$  et appliquons la formule de Cartan (voir D.4) à  $\tilde{h}$  et  $\lambda$  :

$$\mathcal{L}_{\tilde{h}} \lambda = d[i_{\tilde{h}}(\lambda)] + i_{\tilde{h}}[d\lambda], \quad (\text{VI.91})$$

où, rappelons-le :

$$\mathcal{L}_h \lambda = \left. \frac{\partial h(t)^* \lambda}{\partial t} \right|_{t=0}.$$

Mais comme  $\tilde{h}$  est justement, pour tout  $t$ , un automorphisme de  $(Y, \lambda) : \mathcal{L}_h \lambda = 0$ , il reste l'identité :

$$i_{\tilde{h}}[d\lambda] = -d[i_{\tilde{h}}(\lambda)]. \quad (\text{VI.92})$$

Soit alors  $h = \pi_* \circ \tilde{h}$  l'homomorphisme de  $\mathbf{R}$  dans  $G$ , projection de  $\tilde{h}$ . Utilisons ensuite le fait que  $\omega$  est la courbure de la connexion  $\lambda : d\lambda = \pi^* \omega$ , nous obtenons :  $i_{\tilde{h}}[d\lambda] = i_h(\omega)$ . D'autre part,  $i_{\tilde{h}}(\lambda)$  est une fonction de  $Y$ , dont la valeur en  $y \in Y$  est donnée par :  $i_{\tilde{h}}(\lambda)(y) = (\hat{y}_{\tilde{h}}^* \lambda)_0(1)$  (voir formule (D.5)) ; il suffit de vérifier que  $\hat{y}_{\tilde{h}} = \hat{y} \circ \tilde{h}$  et on déduit immédiatement  $i_{\tilde{h}}(\lambda)(y) = \tilde{h}^*(\hat{y}^* \lambda)_0(1) = i_{\tilde{h}}(\mu_Y(y))$ . Ainsi, en utilisant la suite d'identités :  $d[i_{\tilde{h}}(\mu_Y(y))] = d[i_{\tilde{h}}(\mu_Y(y)) - i_{\tilde{h}}(\mu_Y(y_o))] = d[i_{\tilde{h}}(\mu_o(x))] = d[i_h(\mu_o(x))]$ , on obtient l'expression finale :

$$i_h[\omega] = -d[i_h(\mu_o(x))]. \quad (\text{VI.93})$$

On peut reconnaître sur cette dernière formule l'expression originale du moment tel qu'il a été défini par J.-M. Souriau [Sou2]. Plus précisément en notant  $Z_X$  le champ de vecteur associé à  $h$  sur  $X : Z_X(x) = (\partial h(t)(x)/\partial t)_{t=0}$  l'expression précédente devient :

$$\omega(Z_X, \cdot) = -d[\mu_o \cdot Z]. \quad (\text{VI.94})$$

Il ne nous reste plus, avant de clore ce chapitre, qu'à retrouver grâce à cette définition de l'application moment ses différentes variances par rapport à l'action

de  $G$ . Considérons donc un élément  $g \in G$  (on peut supposer si on désire que  $G = \text{Diff}(X, \omega)$ ). Soit  $\tilde{g} \in \tilde{G}$  se projetant sur  $G$ . On a<sup>8</sup> :

$$\begin{aligned} \mu_o(g(x)) &= \widehat{\tilde{g}(y)}^* \lambda - \hat{y}_o^* \lambda \\ &= (\hat{y} \circ R_{\tilde{g}})^* \lambda - \hat{y}_o^* \lambda \\ &= R_{\tilde{g}}^*(\hat{y}^* \lambda) - \hat{y}_o^* \lambda \\ &= R_{\tilde{g}}^*(\hat{y}^* \lambda - \hat{y}_o^* \lambda) + R_{\tilde{g}}^*(\hat{y}_o^* \lambda) - \hat{y}_o^* \lambda. \end{aligned}$$

Remarquons alors que :  $R_{\tilde{g}}^*(\hat{y}^* \lambda - \hat{y}_o^* \lambda) = R_{\tilde{g}}^*(\mu_o(x)) = R_g^*(\mu_o(x))$ . D'autre part,  $R_g^*(\mu_o(x)) = \text{Ad}^*(g)(\mu_o(x))$ , par définition même de l'action coadjointe de  $G$  sur le dual de son algèbre de Lie  $\mathfrak{G}^*$ . Notons encore que  $R_{\tilde{g}}^*(\hat{y}_o^* \lambda) - \hat{y}_o^* \lambda$  est un élément de  $\mathfrak{G}^*$  ne dépendant que de  $g$  ; nous le noterons :

$$\Theta(g) = R_{\tilde{g}}^*(\hat{y}_o^* \lambda) - \hat{y}_o^* \lambda. \quad (\text{VI.95})$$

Il vient alors l'expression de la variance de  $\mu_o$  par rapport au groupe  $G$  :

$$\mu_o(g(x)) = \text{ad}^*(g)(\mu_o(x)) + \Theta(g). \quad (\text{VI.96})$$

Il est inutile de préciser que, par construction même, la fonction  $\Theta$ , de  $G$  dans  $\mathfrak{G}^*$  est un 1-cocycle, que l'on appelle le « défaut d'équivariance » du moment :

$$\theta \in H^1(G, \mathfrak{G}^*), \quad \Theta(gg') = \text{Ad}^*(g')(\Theta(g)) + \Theta(g'). \quad (\text{VI.97})$$

Ce cocycle n'est pas nécessairement trivial, même si on peut se laisser tromper par sa définition. Il le serait (trivial) s'il existait une 1-forme  $\alpha \in \mathfrak{G}^*$  telle que  $\Theta(g) = \text{Ad}^*(g)(\alpha) - \alpha$  ; or dans la définition de  $\Theta$ , c'est une 1-forme de  $\tilde{\mathfrak{G}}^*$  qui intervient et non de  $\mathfrak{G}^*$ . Si par hasard le cocycle  $\Theta$  est trivial, on peut alors ajouter une constante au moment  $\mu$  de façon à le rendre équivariant<sup>9</sup>.

REMARQUE. Une condition évidente pour que le défaut d'équivariance  $\Theta$  de l'action hamiltonienne d'un groupe de Lie  $G$  soit trivial est que ce groupe ait un point fixe. En effet, supposons que  $a \in X$  soit un point fixe :  $g(a) = a$  pour

<sup>8</sup>Rappelons que  $R_k$  désigne, sur un groupe, la multiplication à droite par l'élément  $k$ .

<sup>9</sup>Pour une discussion générale sur la variance du moment et des exemples fameux de défaut d'équivariance (par exemple la masse totale d'un système dynamique), voir [Sou2].

tout  $g \in G$ . On a alors d'une part :  $\mu_o(g(a)) = \text{Ad}^*(g)(\mu_o(a)) + \Theta(g)$  et d'autre part :  $\mu_o(g(a)) = \mu_o(a) \Rightarrow \Theta(g) = \mu_o(a) - \text{Ad}^*(g)(\mu_o(a))$ ; en posant ensuite  $\nu = -\mu_o(a) \in \mathfrak{G}^*$ , il vient :  $\Theta(g) = \text{Ad}^*(g)(\nu) - \nu$ . Ce qui prouve que le moment peut être choisi équivariant.

**CONCLUSION.** Dans tous les cas où un groupe de Lie  $G$  agit sur une variété  $X$  en préservant une 2-forme fermée  $\omega$ , nous avons ramené la construction de l'*application moment* au cas le plus simple possible, celui d'une 1-forme invariante par un groupe de Lie, cas que nous avons largement commenté précédemment. Mais, pour cela nous avons dû :

- 1) *intégrer* la 2-forme fermée invariante  $\omega$  en construisant le fibré des périodes  $Y$  associé pour l'y relever et l'y intégrer ;
- 2) sélectionner parmi les difféomorphismes  $g \in G$  préservant  $\omega$  les *difféomorphismes hamiltoniens*, ceux qui se relèvent au fibré des périodes  $Y$ .

Une fois le moment défini sur  $Y$ , les problèmes de variance et de cohomologie associée se réduisent alors à des choix différents de réduction. Nous allons illustrer ces constructions générales dans les quelques exemples suivants.

### VI.5.3 Quelques exemples

Illustrons maintenant les constructions précédentes par quelques exemples, en laissant au lecteur intéressé et curieux le plaisir d'en construire d'autres...

#### EXEMPLE 1 : LE TORE $\mathbf{T}^3$ .

Pour notre premier exemple, considérons le tore  $\mathbf{T}^3$ , quotient de  $\mathbf{R}^3$  par le réseau  $\mathbf{Z}^3$ . Nous noterons  $X = (x, y, z)$  un point de  $\mathbf{R}^3$  et  $[x, y, z]$  sa classe dans  $\mathbf{T}^3$ . Considérons la forme  $\alpha$  sur  $\mathbf{R}^3$  définie par

$$\alpha = aydz + bzdx + cxdy, \quad a, b, c \in \mathbf{R}^3. \quad (\text{VI.98})$$

Cette forme ne passe pas sur  $\mathbf{T}^3$ , mais sa dérivée extérieure  $d\alpha$  oui. Soit  $\omega$  l'image de  $d\alpha$  sur  $\mathbf{T}^3$ , que nous noterons de la manière suivante :

$$\omega = [d\alpha] = [ady \wedge dz + bdz \wedge dx + cdx \wedge dy]. \quad (\text{VI.99})$$

Le groupe des périodes de la forme  $\omega$  est le sous-groupe de  $\mathbf{R}$  engendré par  $a$ ,  $b$  et  $c$ . On a donc :

$$P_\omega = a\mathbf{Z} + b\mathbf{Z} + c\mathbf{Z} \subset \mathbf{R}. \quad (\text{VI.100})$$

Plusieurs possibilités s'offrent alors : le groupe des périodes peut être isomorphe à  $\mathbf{Z}$ ,  $\mathbf{Z}^2$  ou  $\mathbf{Z}^3$  suivant les propriétés de commensurabilité des coefficients  $a$ ,  $b$  et  $c$ . Supposons, par exemple, que pour tout triplet  $l, m, n$  d'entiers relatifs  $la + mb + nc = 0 \Rightarrow l = m = n = 0$  ; alors  $a$ ,  $b$  et  $c$  sont incommensurables et le groupe des périodes est égal à  $a\mathbf{Z} \oplus b\mathbf{Z} \oplus c\mathbf{Z} \subset \mathbf{R}$ . Dans tous les cas, ce qui suit est valable.

Dans le cas du tore  $\mathbf{T}^3$ , le fibré des périodes  $Y$  est unique à isomorphisme près. D'autre part, l'image réciproque du fibré des périodes par la projection canonique  $\mathbf{R}^3 \rightarrow \mathbf{T}^3$  est trivial, et donc isomorphe à  $\mathbf{R}^3 \times T_\omega$ . Il peut donc se reconstruire par quotient en relevant l'action de  $\mathbf{Z}^3$  sur  $\mathbf{R}^3 \times T_\omega$ . Cette action est du type :

$$L \begin{pmatrix} X \\ \tau \end{pmatrix} = \begin{pmatrix} X + L \\ \tau + \phi(X, L) \end{pmatrix}, \quad L = \begin{pmatrix} l \\ m \\ n \end{pmatrix} \in \mathbf{Z}^3, \quad \tau = [t] \in T_\omega, \quad (\text{VI.101})$$

où l'action sur  $T_\omega$  est notée aditivement et où  $\phi$  est un cocycle, c'est-à-dire qu'il vérifie la propriété suivante :

$$\phi(X, L + L') = \phi(X + L, L') + \phi(X, L'). \quad (\text{VI.102})$$

Pour expliciter  $\phi$ , introduisons la forme de connexion  $\Lambda$ , définie sur  $\mathbf{R}^3 \times T_\omega$ , image réciproque de la « forme d'intégration »  $\lambda$  de  $\omega$ , définie sur  $Y$  :

$$\Lambda = \alpha + d\tau = aydz + bzdx + cxdy + d\tau. \quad (\text{VI.103})$$

La condition d'invariance  $L^*\Lambda = \Lambda$  impose tout de suite :

$$amdz + bndx + cldy + d\phi(X, L) = 0. \quad (\text{VI.104})$$

On en déduit alors l'expression du cocycle<sup>10</sup>  $\phi$  :

$$\phi(X, L) = -[amz + bnx + cly] \in T_\omega. \quad (\text{VI.105})$$

Ainsi le fibré des périodes de la 2-forme  $\omega$  au dessus de  $\mathbf{T}^3$  est le quotient :

$$Y = \mathbf{R}^3 \times T_\omega / \mathbf{Z}^3, \quad \begin{pmatrix} X \\ \tau \end{pmatrix} \sim L \begin{pmatrix} X \\ \tau \end{pmatrix} = \begin{pmatrix} X + L \\ \tau - [\bar{L}WX] \end{pmatrix}, \quad (\text{VI.106})$$

où la barre désigne la transposée et où on a posé :

$$W = \begin{pmatrix} 0 & c & 0 \\ 0 & 0 & a \\ b & 0 & 0 \end{pmatrix}. \quad (\text{VI.107})$$

Avec ces notations, la forme de connexion  $\Lambda$  s'écrit aussi :

$$\Lambda = \bar{X}WdX + d\tau. \quad (\text{VI.108})$$

Les translations du tore  $\mathbf{T}^3$  ( $\sim \mathbf{T}^3$ ) préservent la 2-forme  $\omega$ , la vérification étant immédiate. Nous allons montrer que cette action n'est pas hamiltonienne, autrement dit qu'elle ne se relève pas en un automorphisme de  $Y$ , sauf pour le seul sous-groupe engendré par le vecteur  $(a, b, c)$ . Considérons donc la translation associée à un vecteur  $V \in \mathbf{R}^3$ , et supposons que cette action soit hamiltonienne. Commençons par relever cette translation de  $\mathbf{T}^3$  sur  $\mathbf{R}^3$ , soit  $\phi(X) = X + V$ . Le relevé de  $\phi$  sur  $Y$  se relève sur  $\mathbf{R}^3 \times T_\omega$  en :

$$\psi \begin{pmatrix} X \\ \tau \end{pmatrix} = \begin{pmatrix} X + V \\ \tau + f(X, V) \end{pmatrix}, \quad (\text{VI.109})$$

mais  $\psi$  préserve la forme de connexion :  $\psi^*\Lambda = \Lambda$ . Un calcul analogue aux précédents donne la valeur de  $f$  :

$$f(X, V) = -\bar{V}WX + \tau_o \quad \Rightarrow \quad \psi \begin{pmatrix} X \\ \tau \end{pmatrix} = \begin{pmatrix} X + V \\ \tau - \bar{V}WX + \tau_o \end{pmatrix}. \quad (\text{VI.110})$$

<sup>10</sup>Vérifier que la propriété (VI.102) est bien satisfaite.

Il nous reste à vérifier que cet automorphisme de  $(\mathbf{R}^3 \times T_\omega, \Lambda)$  passe bien au quotient en un automorphisme de  $(Y, \lambda)$ , c'est-à-dire qu'il commute avec l'action de  $\mathbf{Z}^3$ ; il est immédiat de vérifier que :

$$\psi \circ L \begin{pmatrix} X \\ \tau \end{pmatrix} = L \circ \psi \begin{pmatrix} X \\ \tau \end{pmatrix} \Rightarrow \bar{L}WV = \bar{V}WL \quad \forall L \in \mathbf{Z}^3. \quad (\text{VI.111})$$

Autrement dit,  $\phi$  se relève sur  $\mathbf{R}^3 \times T_\omega$  en un automorphisme équivariant si et seulement si :

$$\bar{V}[\bar{W} - W]L = 0, \forall L \in \mathbf{Z}^3. \quad (\text{VI.112})$$

Mais  $\bar{W} - W$  est une matrice antisymétrique de  $\mathbf{R}^3$  et vaut :

$$\bar{W} - W = j(\Omega), \quad \Omega = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad (\text{VI.113})$$

où  $j$  désigne le produit vectoriel de  $\mathbf{R}^3$ . Ainsi, la translation par le vecteur  $V$  n'est hamiltonienne que si :

$$\forall L \in \mathbf{Z}^3 : \langle V, \Omega \wedge L \rangle = (V \wedge \Omega, L) = 0 \quad \Rightarrow \quad V \propto \Omega. \quad (\text{VI.114})$$

Remarquons toutefois que le moment associé au groupe des translations engendré par  $\Omega$  est nul. En effet, soit  $y = (X, \tau) \in \mathbf{R}^3 \times T_\omega$ , on a :

$$\hat{y} : s \mapsto s\Omega \begin{pmatrix} X \\ \tau \end{pmatrix} = \begin{pmatrix} X + s\Omega \\ \tau - s\bar{\Omega}WX \end{pmatrix}, \quad s \in \mathbf{R}, \quad (\text{VI.115})$$

où l'on a choisi la constante additive nulle. Le calcul de  $\hat{y}^*\Lambda$  est immédiat :

$$\hat{y}^*\Lambda_{s=0}(1) = \bar{X}[W - \bar{W}]\Omega = \langle X, \Omega \wedge \Omega \rangle = 0. \quad (\text{VI.116})$$

La raison de cela est simplement que le groupe des translations engendré par  $\omega$  est feuilletage caractéristique de la 2-forme  $\omega$ , ce qu'il est facile de vérifier directement.

#### EXEMPLE 2 : LE GROUPE DE HEISENBERG.

L'exemple suivant est fameux car il définit un groupe important en physique théorique : le *groupe de Heisenberg*. Considérons le plan vectoriel  $X = \mathbf{R}^n \times \mathbf{R}^n$  muni

de la forme symplectique canonique  $\omega = \sum_{i=1}^n dy^i \wedge dx^i$ . Considérons  $\mathbf{R}^n \times \mathbf{R}^n$  comme agissant par translation sur lui-même, notons :

$$T_U(X) = X + U, \text{ avec } X = \begin{pmatrix} x \\ y \end{pmatrix} \text{ et } U = \begin{pmatrix} u \\ v \end{pmatrix}. \quad (\text{VI.117})$$

Cette action est évidemment symplectique :  $T_U^* \omega = \omega$  ; elle est hamiltonienne et nous allons calculer son moment par la méthode exposée précédemment. La forme  $\omega$  est exacte ; choisissons comme primitive<sup>11</sup> :

$$\alpha = \frac{1}{2}[ydx - xdy]. \quad (\text{VI.118})$$

Le fibré des périodes  $Y$  est le produit direct  $\mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}$ , notons :

$$y = \begin{pmatrix} x \\ y \\ t \end{pmatrix} \in Y = \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}. \quad (\text{VI.119})$$

La forme de connexion  $\lambda$  de courbure  $\omega$  est évidemment donnée par :

$$\lambda = \alpha + dt = \frac{1}{2}[ydx - xdy] + dt. \quad (\text{VI.120})$$

Soit  $\varphi$  un relevé d'une translation, on a nécessairement :

$$\varphi \begin{pmatrix} X \\ t \end{pmatrix} = \begin{pmatrix} X + U \\ t + F(X, U) \end{pmatrix},$$

la fonction  $F$  étant déterminée par la condition d'invariance :  $\varphi^* \lambda = \lambda$ . On obtient après un petit calcul :

$$\varphi^* \lambda = \lambda \quad \Rightarrow \quad F(X, U) = \frac{1}{2} \omega(X, U) + c. \quad (\text{VI.121})$$

Le difféomorphisme  $\varphi$  est donc caractérisé par le couple  $(U, c)$ , ce que nous écrirons de la manière suivante :

$$\begin{pmatrix} U \\ c \end{pmatrix} \begin{pmatrix} X \\ t \end{pmatrix} = \begin{pmatrix} X + U \\ t + c + \frac{1}{2} \omega(X, U) \end{pmatrix}. \quad (\text{VI.122})$$

<sup>11</sup>On omet la somme sur l'indice  $i$  sans risque de confusion.

L'extension de  $\mathbf{R}^{2n}$  par  $\mathbf{R}$  que nous venons de construire ainsi s'appelle le *groupe de Heisenberg* :

$$0 \longrightarrow \mathbf{R} \longrightarrow \text{Heisenberg} \longrightarrow \mathbf{R}^{2n} \longrightarrow 0. \quad (\text{VI.123})$$

La loi de groupe s'explique en composant deux relevés ou, si l'on préfère, en réinterprétant la formule précédente, car dans cet exemple l'espace et le groupe se confondent :

$$\begin{pmatrix} U \\ c \end{pmatrix} \begin{pmatrix} U' \\ c' \end{pmatrix} = \begin{pmatrix} U + U' \\ c + c' + \frac{1}{2}\omega(U, U') \end{pmatrix}. \quad (\text{VI.124})$$

Le calcul explicite du moment est une formalité. Pour que les choses soient claires, notons deux ou trois formules nécessaires à ce calcul. Soit  $(Z, \epsilon)$  un élément de l'algèbre de Lie du groupe de Heisenberg (soit  $\mathcal{H}$ ), son action infinitésimale sur  $\mathbf{R}^{2n}$  est donnée par :

$$\begin{pmatrix} Z \\ \epsilon \end{pmatrix}_Y \begin{pmatrix} X \\ t \end{pmatrix} = \begin{pmatrix} Z \\ \epsilon + \frac{1}{2}\omega(X, Z) \end{pmatrix}, \quad \begin{pmatrix} Z \\ \epsilon \end{pmatrix} \in \mathcal{H}. \quad (\text{VI.125})$$

Le calcul des moments respectifs donne alors :

$$\hat{y}^* \lambda_{\substack{U=0 \\ c=0}} \begin{pmatrix} Z \\ \epsilon \end{pmatrix} = \omega(X, Z) + \epsilon \quad \text{et} \quad \hat{y}_o^* \lambda_{\substack{U=0 \\ c=0}} \begin{pmatrix} Z \\ \epsilon \end{pmatrix} = \epsilon, \quad (\text{VI.126})$$

d'où le résultat :

$$\mu_o(X) = \hat{y}^* \lambda - \hat{y}_o^* \lambda = \omega(X, \cdot) = [Z \mapsto \omega(X, Z)]. \quad (\text{VI.127})$$

### EXEMPLE 3 : LE COCYCLE DE BOTT-THURSTON.

Dans l'exemple qui suit, la 2-forme est encore exacte. Nous nous trouvons dans un cas analogue au précédent, mais l'espace sur lequel est définie cette forme différentielle est de dimension infinie. Nous laissons au lecteur, comme exercice de difféologie, le soin de vérifier que les affirmations qui suivent sont fondées. L'intérêt de cet exemple est qu'il fait apparaître une version du groupe de Virasoro, groupe qui a récemment connu une certaine vogue chez les physiciens. On voit apparaître le *cocycle de Bott-Thurston* et la *dérivée Schwartzienne* ainsi que le *cocycle de Gelfand-Fucks* comme différents produits de cet exemple.

Considérons l'espace  $\text{Imm}(S^1, \mathbf{R}^2)$  des immersions du cercle  $S^1$  dans  $\mathbf{R}^2$ . Il est pointé par le plongement ordinaire  $\iota : S^1 \hookrightarrow \mathbf{R}^2$ . Nous le munissons de la 1-forme suivante<sup>12</sup> :

$$\alpha(\delta\gamma) = \int \frac{1}{\|\gamma'\|^2} \langle \gamma'', \delta\gamma' \rangle, \quad (\text{VI.128})$$

où  $\delta\gamma$  est un « vecteur tangent » en  $\gamma$  à l'espace  $\text{Imm}(S^1, \mathbf{R}^2)$ , c'est-à-dire un relèvement dans  $T\mathbf{R}^2$  de l'immersion  $\gamma : \delta\gamma(z) \in T_{\gamma(z)}\mathbf{R}^2$  ; d'autre part  $\gamma'$  désigne la dérivée par rapport au paramètre de l'immersion, et l'intégration se fait sur le cercle  $S^1$ . Soit  $\omega = d\alpha$ . Considérons l'(anti)-action naturelle des difféomorphismes du cercle  $S^1$  sur  $\text{Imm}(S^1)$  par composition :

$$\forall \phi \in \text{Diff}(S^1), \quad \phi(\gamma) = \gamma \circ \phi. \quad (\text{VI.129})$$

On vérifie que

$$\phi^*\alpha = \alpha + dF(\phi)$$

pour tout difféomorphisme  $\phi \in \text{Diff}^\circ(S^1)$ , avec :

$$F(\phi) : \gamma \mapsto \int \log \|(\gamma \circ \phi)'\| \, d\log \phi'. \quad (\text{VI.130})$$

La 2-forme  $\omega$  est donc invariante par  $\text{Diff}^\circ(S^1)$ , et en voici une expression :

$$\begin{aligned} \omega(\delta\gamma, \delta'\gamma) = \int \frac{1}{\|\gamma'\|^2} \left\{ \langle \delta\gamma'' - \delta[\log \|\gamma'\|^2]\gamma'', \delta'\gamma' \rangle \right. \\ \left. - \langle \delta'\gamma'' - \delta'[\log \|\gamma'\|^2]\gamma'', \delta\gamma' \rangle \right\}. \end{aligned} \quad (\text{VI.131})$$

L'action de  $\text{Diff}(S^1)$  est donc évidemment hamiltonienne. Son extension centrale par  $\mathbf{R}$  est définie par :

$$\begin{pmatrix} \phi \\ c \end{pmatrix} \begin{pmatrix} \gamma \\ t \end{pmatrix} = \begin{pmatrix} \gamma \circ \phi \\ t + c - \int \log \|(\gamma \circ \phi)'\| \, d\log \phi' \end{pmatrix}. \quad (\text{VI.132})$$

La loi de groupe, obtenue par composition, est donnée par :

$$\begin{pmatrix} \phi \\ c \end{pmatrix} \begin{pmatrix} \psi \\ e \end{pmatrix} = \begin{pmatrix} \phi \cdot \psi \\ c + e + K_\omega(\phi, \psi) \end{pmatrix}, \quad (\text{VI.133})$$

<sup>12</sup>Exercice : donner la définition de  $\alpha$  au sens des espaces difféologiques.

où  $\phi \cdot \psi = \psi \circ \phi$ , et  $K_\omega$  étant le 2-cocycle suivant, de  $\text{Diff}(S^1)$  dans  $\mathbf{R}$  :

$$K_\omega(\phi, \psi) = \int \log(\psi \circ \phi)' d \log \phi'. \quad (\text{VI.134})$$

On reconnaît en  $K_\omega$  l'expression du *cocycle de Bott-Thurston* [Bot] qui décrit la seule extension centrale non triviale du groupe  $\text{Diff}^\circ(S^1)$  par  $\mathbf{R}$ . Un calcul ordinaire donne ensuite l'expression du moment associé à cette action de  $\text{Diff}(S^1)$  :

$$J(\gamma) : \eta \mapsto \int \left[ \frac{\|\gamma''\|^2}{\|\gamma'\|^2} - \left( \log \|\gamma'\|^2 \right)'' \right] \eta. \quad (\text{VI.135})$$

Nous laissons en exercice le calcul du défaut d'équivariance  $\Theta$ . Il faut toutefois faire attention que  $\int \phi_* \eta = \int \eta(\phi')^2$  ; il vient immédiatement :

$$\Theta(\phi) : \eta \mapsto \int \frac{3\phi''^2 - 2\phi'''\phi'}{\phi'^2} \eta \quad (\text{VI.136})$$

On reconnaît l'expression de la *dérivée Schwarzienne* dont la classe de cohomologie représente le défaut d'équivariance du moment  $J$  de cette 2-forme  $\omega$ . Il est facile ensuite de calculer le cocycle infinitésimal, que l'on reconnaît évidemment comme le *cocycle de Gelfand-Fuchs* [GF] :

$$k_\omega(\xi, \eta) = \int \xi''\eta' - \eta''\xi'. \quad (\text{VI.137})$$

#### EXEMPLE 4 : LES COURBES À COURBURE PRESCRITE.

Nous allons achever cette série d'exemples par un retour au disque de Poincaré, que nous avons introduit au chapitre IV.2. Nous en reprenons les notations. Rappelons que nous identifions le disque de Poincaré  $H^2$  avec la nappe supérieure de l'hyperboloïde à deux nappes, dans l'algèbre de Lie  $\mathfrak{sl}(2, \mathbf{R}) \sim \mathbf{R}^3$  du groupe  $\text{SL}(2, \mathbf{R})$ , d'équation  $x^2 + y^2 - z^2 = -1$  ; rappelons aussi que cette nappe est l'orbite adjointe du point :

$$X_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Considérons le fibré des directions tangentes à  $H^2$ , c'est-à-dire le fibré des demi-droites tangentes au disque. C'est une variété sur laquelle agit naturellement le

groupe  $SL(2, \mathbf{R})$ . Grâce à la métrique de Poincaré, nous l'identifions au fibré unitaire tangent  $UH^2$ , c'est-à-dire à l'ensemble des couples  $(X, U)$  où  $U$  est un vecteur tangent à  $H^2$  de carré 1, *i.e.*  $\|U\|^2 = 1$  (voir la définition IV.16 de la métrique de Poincaré). Dans le paragraphe IV.2, nous avons étudié le feuilletage géodésique du disque de Poincaré, c'est-à-dire le système dynamique défini par le feuilletage caractéristique de la 2-forme fermée  $\omega = d\alpha$  avec  $\alpha = \bar{U}dX$ , définie sur  $UH^2$ ; nous avons remarqué que cette 2-forme est invariante par l'action adjointe de  $SL(2, \mathbf{R})$ . Nous allons étudier maintenant le système dynamique défini par le feuilletage caractéristique de la 2-forme fermée la plus générale invariante par  $SL(2, \mathbf{R})$  définie sur le fibré des directions de  $H^2$  (c'est-à-dire sur  $UH^2$ ).

Notre première remarque sera que le fibré unitaire tangent est en bijection avec le groupe  $SL(2, \mathbf{R})$  lui-même. En effet, nous savons que l'action adjointe de  $SL(2, \mathbf{R})$  est transitive sur  $H^2$ . Pour tout  $X \in H^2$ , il existe  $A \in SL(2, \mathbf{R})$  tel que  $AX_0A^{-1} = X$ . Soit  $A$  et  $B$  deux éléments de  $SL(2, \mathbf{R})$  tels que :  $X = AX_0A^{-1} = BX_0B^{-1}$ , alors  $B = AM$  et  $M \in SO(2, \mathbf{R}) \subset SL(2, \mathbf{R})$ . En effet, par hypothèse  $MX_0M^{-1} = X_0$ , c'est-à-dire  $MX_0 = X_0M$ . On vérifie immédiatement que cette condition équivaut à :

$$M = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad \text{avec } a^2 + b^2 = 1 \quad \Rightarrow \quad M \in SO(2, \mathbf{R}).$$

Introduisons alors le vecteur  $U_0$  tangent à  $H^2$  au point  $X_0$  :

$$U_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad U_0 \in U_{X_0}H^2. \quad (\text{VI.138})$$

Considérons ensuite un couple  $(X, U) \in UH^2$  et soit  $A \in SL(2, \mathbf{R})$  tel que :  $AX_0A^{-1} = X$ . Soit  $U'_0$  le vecteur tel que  $AU'_0A^{-1} = U$ . Le vecteur  $U'_0$  est un vecteur unitaire dans le plan tangent à  $H^2$  au point  $X_0$ ; il existe un élément unique  $M \in SO(2)$  tel que  $MU'_0M^{-1} = U'_0$  (le vérifier). On pose alors  $B = AM$  et l'on a  $BX_0B^{-1} = X$  et  $BU_0B^{-1} = U$ ,  $B$  étant unique. L'équivalence :

$$UH^2 = \{(X, U) \mid X \in H^2, \|U\|^2 = 1, \langle X, U \rangle = 0\} \sim SL(2, \mathbf{R}) \quad (\text{VI.139})$$

est formellement définie par :

$$UH^2 \ni (X, U) \mapsto A \in SL(2, \mathbf{R}) \Leftrightarrow AX_0A^{-1} = X \text{ et } AU_0A^{-1} = U. \quad (\text{VI.140})$$

Notons ici que l'action adjointe de  $\mathrm{SL}(2, \mathbf{R})$  sur  $UH^2$  se traduit simplement par la multiplication à gauche sur  $\mathrm{SL}(2, \mathbf{R})$ , grâce à l'identification :

$$(X, U) \sim A \Leftrightarrow (MXM^{-1}, MUM^{-1}) \sim MA.$$

Nous devons donc chercher l'expression la plus générale de 2-forme fermée sur  $UH^2$  invariante par l'action adjointe ; ou, ce qui revient au même, l'expression la plus générale de 2-forme fermée sur  $\mathrm{SL}(2, \mathbf{R})$  invariante par la multiplication à gauche. Nous allons immédiatement vérifier que ces formes constituent un espace vectoriel de dimension 3. Orientons d'abord l'algèbre de Lie  $\mathfrak{sl}(2, \mathbf{R})$  : soit  $\mathrm{vol}$  la forme volume définie par :

$$\mathrm{vol}(A, B, C) = -\frac{1}{2} \langle A, [B, C] \rangle. \quad (\text{VI.141})$$

Elle est telle que la base faite des vecteurs

$$U_0 = (1, 0, 0), \quad J_0 = (0, 1, 0) = \frac{1}{2}[U_0, X_0] \quad \text{et} \quad X_0 = (0, 0, 1)$$

soit directe et de volume 1. Soit  $C^2$  la nappe hyperboloïde des vecteurs de  $\mathfrak{sl}(2, \mathbf{R})$  de carré scalaire 1. Soit  $\mathrm{surf}_{H^2}$  et  $\mathrm{surf}_{C^2}$  les formes de surfaces associées au volume  $\mathrm{vol}$  sur les sous-variétés  $H^2$  et  $C^2$  de  $\mathfrak{sl}(2, \mathbf{R})$ , définies par

$$\mathrm{surf}_{\Sigma}(x)(u, v) = \mathrm{vol}(x, u, v), \quad x \in \Sigma, (u, v) \in T_x \Sigma^2, \quad \Sigma = H^2 \text{ ou } C^2.$$

Avec ces notations, la 2-forme la plus générale invariante définie sur  $UH^2 \sim \mathrm{SL}(2, \mathbf{R})$  dépend de trois constantes  $a, b, c$  et s'exprime de la façon suivante :

$$\begin{aligned} \omega(\delta Z, \delta' Z) &= a[\langle \delta U, \delta' X \rangle - \langle \delta' U, \delta X \rangle] \\ &+ b \mathrm{surf}_{H^2}(X)(\delta X, \delta' X) \\ &+ c \mathrm{surf}_{C^2}(U)(\delta U, \delta' U). \end{aligned} \quad (\text{VI.142})$$

En effet, toute 2-forme différentielle sur  $\mathrm{SL}(2, \mathbf{R})$  invariante par multiplication à gauche est entièrement définie par sa valeur en l'identité. L'espace des 2-formes invariantes sur  $\mathrm{SL}(2, \mathbf{R})$  est donc au plus de dimension 3, dimension de l'espace des 2-formes linéaires sur  $\mathfrak{sl}(2, \mathbf{R}) = T_1 \mathrm{SL}(2, \mathbf{R})$ . Il suffit donc de montrer que la 2-forme précédente est nulle si et seulement si  $a = b = c = 0$ . Cela se vérifie

facilement en introduisant le vecteur  $J = (1/2)[U, X]$ , et en explicitant la forme  $\omega$  sur la base « orthonormée »  $(U, J, X)$ . On vérifie ensuite qu'elle est bien fermée puisque le premier terme est une différentielle et les deux autres termes les images réciproques des éléments de surface sur  $H^2$  et  $C^2$  par les deux projections  $(X, U) \mapsto X$  et  $(X, U) \mapsto U$ .

Il est alors facile de déterminer le noyau de  $\omega$ , qui est de dimension 1 (si  $\omega$  est non-nulle évidemment) :

$$\frac{dZ}{ds} \in \ker \omega \Leftrightarrow \frac{dX}{ds} = \lambda(aU - cJ) \text{ et } \frac{dU}{ds} = \lambda(aX + bJ), \lambda \in \mathbf{R}. \quad (\text{VI.143})$$

Pour que le feuilletage caractéristique de  $\omega$  soit *holonome*, c'est-à-dire pour qu'il soit solution d'une équation différentielle du second degré sur  $H^2$ , il faut que  $c = 0$ . Le noyau est alors engendré par le champ de vecteurs suivant :

$$\left\{ \begin{array}{l} \frac{dX}{ds} = aU \\ \frac{dU}{ds} = aX + bJ \end{array} \right. \quad (\text{VI.144})$$

Alors  $U$  est bien la vitesse unitaire du système dynamique. C'est ce que nous allons supposer dorénavant :

$$\omega(\delta Z, \delta' Z) = a[\langle \delta U, \delta' X \rangle - \langle \delta' U, \delta X \rangle] + b \text{ surf}_{H^2}(X)(\delta X, \delta' X). \quad (\text{VI.145})$$

Comme nous pouvons le voir sur les équations précédentes (VI.144), les solutions sont les courbes de  $H^2$  de courbure constante  $k = -b$ , la courbure étant donnée par  $k = -J \cdot (dU/ds)$  où  $U = dX/ds$  est la vitesse unitaire et  $J$  le vecteur normal, orienté positivement, de la courbe.

Pour résoudre explicitement ce système, nous allons utiliser le fait que l'action de  $\text{SL}(2, \mathbf{R})$ , définie plus haut, est hamiltonienne. Pour calculer le moment  $\mu$ , précisons l'action infinitésimale de  $\text{SL}(2, \mathbf{R})$  sur  $UH^2$ , naturellement définie par :

$$\zeta \in \text{sl}(2, \mathbf{R}), \quad \zeta_{UH^2}(X, U) = ([\zeta, X], [\zeta, U]). \quad (\text{VI.146})$$

Le moment est obtenu en résolvant l'équation différentielle :

$$\omega(\zeta_{UH^2}(X, U), \delta(X, U)) = -\delta(\mu(X, U) \cdot \zeta). \quad (\text{VI.147})$$

Ce qui donne :

$$\mu(X, U) = a[X, U] + 2bX + cst. \quad (\text{VI.148})$$

Si on choisit la constante nulle et que l'on introduit le vecteur  $J$  défini plus haut, il vient :

$$\mu(X, U) = 2bX - 2aJ \quad \Rightarrow \quad \|\mu(X, U)\|^2 = 4(a^2 - b^2). \quad (\text{VI.149})$$

D'après le théorème de Noether, nous savons que  $\mu$  est constant sur les solutions et par homogénéité nous déduisons que (les composantes de) la surface dans  $\mathfrak{sl}(2, \mathbf{R})$  définie par l'équation  $\|\mu\|^2 = 4(a^2 - b^2)$  réalise(nt), grâce à l'application moment, l'espace des mouvement du système dynamique défini par les caractéristiques de  $\omega$ . Nous voyons que trois cas se distinguent, correspondant aux trois types d'orbites coadjointes<sup>13</sup> de  $\text{SL}(2, \mathbf{R})$  :

$$\begin{aligned} |a| < |b| & : \text{ cas elliptique - hyperboloïdes à deux nappes} \\ |a| = |b| & : \text{ cas parabolique - cône isotrope} \\ |a| > |b| & : \text{ cas hyperbolique - hyperboloïdes à une nappe} \end{aligned} \quad (\text{VI.150})$$

Autrement dit, l'application moment nous permet d'interpréter chaque orbite coadjointe comme l'espace des solutions d'un système dynamique du type (VI.143), défini comme les caractéristiques de la 2-forme  $\text{SL}(2, \mathbf{R})$  invariante (VI.145). Si on note  $m = a$  et  $qB = b$ , on peut interpréter physiquement ce système comme le mouvement d'une particule test de masse  $m$  de charge électrique  $q$ , soumise à la gravitation représentée par la métrique de Poincaré et à un champ électromagnétique « constant »<sup>14</sup>  $B$ .

Nous pouvons, grâce au moment  $\mu$ , analyser plus précisément la nature géométrique des caractéristiques de la forme  $\omega$  (ce que nous appelons les *mouvements* du système). Soit  $M = \mu(X_0, U_0)$  la valeur du moment le long du mouvement de conditions initiales  $(X_0, U_0)$ . C'est un point de la nappe normique  $4(a^2 - b^2)$ . On remarque qu'en tout point du mouvement  $M \cdot X = -2b$  et  $M \cdot U = 0$ . Autrement dit, le mouvement est entièrement contenu dans le plan affine  $P_{-2b}$

<sup>13</sup>Les différentes composantes des hyperboloïdes à deux nappes ou composantes du cône isotrope correspondent aux solutions orientées dans un sens ou dans un autre.

<sup>14</sup>C'est-à-dire orthogonal au disque et de carré scalaire constant.

d'équation  $M \cdot X = -2b$ . Puisque d'autre part le mouvement est une courbe de la nappe  $H^2$ , c'est exactement l'intersection  $P_{-2b} \cap H^2$ . La nature de l'intersection de  $P$  avec  $H^2$ , plus précisément le nombre de points à l'infini de cette intersection (0,1 ou 2) nous dit dans quel cas (elliptique, parabolique ou hyperbolique) nous nous trouvons. Il est intéressant d'expliciter ces courbes dans le modèle du disque de Poincaré. Soit

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto Y = \begin{pmatrix} \alpha = \frac{x}{1+z} \\ \beta = \frac{y}{1+z} \end{pmatrix}. \quad (\text{VI.151})$$

La projection stéréographique de la nappe  $H^2$  sur le plan  $z = 0$ , dont l'image est le disque de Poincaré d'équation  $\|Y\|^2 < 1$ , l'inverse étant donné par :

$$x = \frac{2\alpha}{1 - \|Y\|^2} \quad y = \frac{2\beta}{1 - \|Y\|^2} \quad z = \frac{1 + \|Y\|^2}{1 - \|Y\|^2}. \quad (\text{VI.152})$$

Notons  $M = (r, s, t)$  ; l'intersection  $P_{-2b} \cap H^2$  dans le plan  $z = 0$  se projette, grâce à la projection stéréographique ci-dessus, comme la courbe d'équation :

$$\begin{aligned} M \cdot X &= -2b \\ rx + sy - tz &= -2b \\ r \frac{2\alpha}{1 - \|Y\|^2} + s \frac{2\beta}{1 - \|Y\|^2} - t \frac{1 + \|Y\|^2}{1 - \|Y\|^2} &= -2b \\ 2(\alpha r + \beta s) - t(1 + \|Y\|^2) &= -2b(1 - \|Y\|^2). \end{aligned}$$

Notons alors

$$M = \begin{pmatrix} T \\ t \end{pmatrix}, \quad T = \begin{pmatrix} r \\ s \end{pmatrix} \in \mathbf{R}^2.$$

En supposant  $2b + t \neq 0$ , l'équation du mouvement dans le plan devient alors, en regroupant les termes comme il se doit :

$$\begin{aligned} 2T \cdot Y - t(1 + \|Y\|^2) &= -2b(1 - \|Y\|^2) \\ \|Y - C\|^2 &= \|C\|^2 + \frac{2b - t}{2b + t} \\ \|Y - C\|^2 &= \rho^2 \end{aligned} \quad (\text{VI.153})$$

où on a posé :

$$C = \frac{T}{2b+t} \quad \text{et} \quad \rho = \left| \frac{2a}{2b+t} \right|. \quad (\text{VI.154})$$

On reconnaît ainsi l'équation d'un cercle de centre  $C$  et de rayon  $\rho$ . Nous pouvons retrouver les trois types de solution elliptique, parabolique et hyperbolique en calculant le nombre de points d'intersection du cercle défini par (VI.153) avec le bord du disque  $\|Y\| = 1$ , c'est-à-dire le nombre de points à l'infini de la courbe (VI.153). Le calcul est laissé en exercice au lecteur :

$$\begin{aligned} 0 \text{ point à l'infini} & : |a| < |b| \text{ elliptique} \\ 1 \text{ point à l'infini} & : |a| = |b| \text{ parabolique} \\ 2 \text{ point à l'infini} & : |a| > |b| \text{ hyperbolique.} \end{aligned} \quad (\text{VI.155})$$

Dans le premier cas, les solutions sont des cercles complètement contenus dans le disque ; dans le second cas ils sont tangents à l'infini : ce sont les horicycles du disque de Poincaré ; enfin dans le troisième cas, il coupe le bord à l'infini en deux points distincts.

On peut vérifier (exercice) que dans le cas hyperbolique, *i.e.*  $|a| > |b|$ , l'angle que fait la trajectoire avec le bord à l'infini est constant. En orientant correctement les trajectoires et le bord, on trouve que cet angle  $\phi$  est donné (au signe près) par :

$$\cos(\phi) = \frac{b}{a}. \quad (\text{VI.156})$$

On retrouve en particulier, lorsque  $b = 0$  (le champ magnétique est nul), la représentation des géodésiques du disque de Poincaré comme les cercles qui coupent le bord à l'infini à angle droit. Si on pose  $b = qB$  et  $a = m$  comme précédemment, alors  $\cos(\phi) = qB/m$  est la *pulsation cyclotron*<sup>15</sup> de la particule.

Lorsque  $2b + t = 0$ , les mouvements sont solutions de l'équation  $Y \cdot T = -2b$ , mais d'autre part  $\|M\|^2 = \|T\|^2 - t^2 = 4(a^2 - b^2) \Rightarrow \|T\| = 2|a|$ . On en déduit que les mouvements sont les droites affines qui passent à la distance  $|b/a|$  de l'origine. Autrement dit, les trajectoires rectilignes, solutions du système dynamique, sont les tangentes au cercle de rayon  $|b/a|$ .

<sup>15</sup>Merci à Christian Duval pour m'avoir précisé ce fait.

Nous venons de voir comment l'application moment nous a permis de traiter complètement l'étude des courbes à courbure prescrite sur le disque de Poincaré, et comment chaque espèce de ces courbes est naturellement l'espace des caractéristiques d'une 2-forme fermée  $SL(2, \mathbf{R})$ -invariante sur le fibré des directions du disque  $H^2$  et hérite donc naturellement d'une structure de variété symplectique. Grâce à l'application moment, en accord avec le théorème de Kirillov-Souriau, chacun de ces espaces est identifié à une *orbite coadjointe* de  $SL(2, \mathbf{R})$  selon la valeur courbure.

---

# Lagrange, Poisson, quelle histoire !

---

Dans les cours de mécanique actuels, il est plus fréquent de voir cités les *crochets de Poisson* plutôt que les *crochets* ou les *parenthèses de Lagrange* ; or, il ne fait aucun doute que l'introduction des divers types de crochets/parenthèses en mécanique analytique est due à J.-L. Lagrange. Quel a donc été le rôle de Poisson dans cette affaire ? On peut lire dans le premier article de Lagrange sur ce sujet [Lag5, pp. 717–718], le texte suivant :

« Dans un mémoire lu à l'académie de Berlin en 1776, je considérai d'une manière directe les variations auxquelles peut être sujet le grand axe d'une planète par les forces perturbatrices provenant de l'action des autres planètes, et je réduisis ces variations à une formule générale et très simple qui, ne dépendant que de la différentielle partielle d'une fonction finie relativement au mouvement moyen de la planète, fait voir tout de suite que le grand axe ne peut jamais contenir aucun terme proportionnel au temps, quelque loin qu'on continue l'approximation par rapport aux excentricités et aux inclinaisons des orbites, mais en s'arrêtant à la première approximation par rapport aux termes proportionnels aux masses des planètes.

On n'avait pas été plus loin sur ce point ; mais M. Poisson y a fait un pas de plus dans le *Mémoire* qu'il a lu il y a deux mois à la classe,

---

*sur les inégalités séculaires des moyens mouvements des planètes*, et dont nous avons fait le Rapport dans la dernière séance. Il a poussé l'approximation de la même formule jusqu'aux termes affectés des carrés et des produits des masses, en ayant égard dans cette formule à la variation des éléments que j'avais regardés comme constants dans la première approximation.

[...]

Cette découverte de M. Poisson a réveillé mon attention sur un objet qui m'avait autrefois beaucoup occupé, et que j'avais ensuite totalement perdu de vue. Il me parut que le résultat qu'il venait de trouver par le moyen des formules qui représentent le mouvement elliptique était un résultat analytique dépendant de la forme des équations différentielles et des conditions de la variabilité des constantes, et que l'on devait y arriver par la seule force de l'analyse, sans connaître les expressions particulières des quantités relatives à l'orbite elliptique.

[...]

J'ai obtenu des formules qui donnent les différentielles de ces variations sous une forme plus simple que celle des formules connues jusqu'à présent.

[...]

Ces formules ont de plus l'avantage que, étant appliquées aux variations du grand axe, on en voit naître tout de suite des expressions analogues à celles auxquelles M. Poisson n'est parvenu que par des réductions heureuses des formules déduites de la considération du mouvement elliptique. »

Ces formules, auxquelles Lagrange fait allusion, sont celles qui ont été exposées dans notre première partie, chapitre II.2, et qui expriment les forces de perturbation relativement aux éléments des planètes en fonction de la variation des constantes et des parenthèses de Lagrange (éq. II.29). Quant à l'article de Poisson auquel Lagrange fait référence, il est publié dans le cahier n° 15 du *Journal de*

*l'école polytechnique* [Poi1]. Comme on peut le constater à sa lecture, cet article de Poisson ne comporte aucune définition des *crochets de Poisson*. Comme le dit Lagrange, la méthode est traditionnelle, les calculs des perturbations y sont simplement poussés davantage que dans les articles précédents de Laplace ou de Lagrange. Alors comment s'explique cette apparition des crochets de Poisson ?

Depuis longtemps, Lagrange avait l'habitude d'utiliser des *parenthèses* ou *crochets* dans les développements en séries perturbatives, où ils apparaissent comme les coefficients des différentielles des éléments perturbés, Lagrange y regroupait l'ensemble des variables dont ces coefficients dépendaient. Ces notations sont utilisées, en particulier, dans son article de 1781 sur la *Théorie des variations séculaires* [Lag4]. Il n'est donc pas étonnant de les voir apparaître à nouveau dans ce *Mémoire sur la théorie des variations des éléments des planètes* en 1808. Seulement, comme nous l'avons vu dans les chapitres précédents, dans le cas particulier de la variation des constantes dans les problèmes de la mécanique, cette notation, qui n'était qu'une commodité jusqu'alors, révèle une structure sous-jacente tout à fait particulière, connue aujourd'hui comme *structure symplectique* de l'espace des solutions du système des planètes. Plus précisément : ces parenthèses sont les coordonnées covariantes de cette structure symplectique.

Quelques mois après son premier mémoire, Lagrange comprend la portée universelle de sa découverte et généralise sa méthode à tous les problèmes de la mécanique. Pour apprécier ensuite l'apport de Poisson, il est nécessaire d'entrer à nouveau au cœur de cette méthode : Lagrange développe les forces perturbatrices par rapport aux éléments kepleriens des planètes (et de façon plus générale par rapport aux constantes d'intégration du système). Ces forces sont des combinaisons linéaires des variations des éléments eux-mêmes et les coefficients de ces variations sont les parenthèses de Lagrange. Pour obtenir la variation des éléments des planètes, en fonctions des forces perturbatrices, il faut alors inverser le système linéaire ainsi obtenu dont les coefficients sont les parenthèses. Lagrange dit lui-même [Lag5, p. 731] :

« Il s'ensuit qu'on pourra également exprimer les différentielles  $\frac{da}{dt}$ ,  $\frac{db}{dt}$ ,  $\frac{dc}{dt}$ , ... par les différences partielles de la fonction  $\Omega$  relatives aux éléments  $a, b, c, \dots$ , multipliées par de simples fonctions de ces quan-

tités sans  $t$  ; car il n'y aura qu'à déduire les valeurs de ces différentielles des six équations précédentes par les méthodes ordinaires de l'élimination, et il est visible qu'elles seront toutes de la forme

$$\left( A \frac{d\Omega}{da} + B \frac{d\Omega}{db} + C \frac{d\Omega}{dc} + F \frac{d\Omega}{df} + G \frac{d\Omega}{dg} + H \frac{d\Omega}{dh} \right) dt,$$

dans laquelle les coefficients  $A, B, C, F, \dots$  ne seront donnés que par les coefficients  $(a, b), (a, c), (b, c), \dots$ , et ne seront par conséquent que de simples fonctions de ces éléments sans  $t$  ; ce qui fournit un Théorème très-important et très-utile dans la théorie des perturbations des planètes. »

Ce sont évidemment ces coefficients  $A, B, C, F, \dots$  qui, considérés par Poisson comme fonctions des conditions initiales du mouvement<sup>1</sup> (temps, position, vitesse) vont devenir les *crochets de Poisson* des constantes  $a, b, c, \dots$  du mouvement. Poisson donne leur construction dans son mémoire intitulé *Sur la variation des constantes arbitraires dans les questions de Mécanique*, publié à la suite du mémoire *Sur la théorie générale de la variation des constantes arbitraires dans tous les problèmes de la mécanique* de Lagrange [Lag6]. Voici ce que Poisson écrit dans ce mémoire, lu à l'Institut le 16 octobre 1809 et publié dans le *Journal de l'école polytechnique* [Poi2, pp. 268–269] :

« Les formules du dernier Mémoire de M. Lagrange sont inverses des nôtres : elles donnent les différences partielles de la même fonction, au moyen des différentielles des constantes arbitraires. Je les rapporte dans la suite de ce mémoire afin de les comparer à celles que nous avons obtenues, et de montrer la singulière analogie qui existe entre ces deux genres de formule. »

Il ajoute plus loin, aux page 289–290 du même mémoire, faisant référence au Théorème de Lagrange cité plus haut :

<sup>1</sup>Remarquez la différence de point de vue : pour Lagrange, un ensemble de constantes d'intégration  $(a, b, c, \dots)$  définit un mouvement : ce sont les formules (II.36) et (II.37) ; les positions et vitesses sont alors fonctions des constantes d'intégration. Poisson considère les constantes d'intégration comme fonctions des conditions initiales. C'est cette différence de point de vue qui conduit naturellement à des formules inverses les unes des autres.

« M. Lagrange démontre directement que les quantités  $[a, b]$ ,  $[a, c]$ ,  $[a, b]$ , &c. sont des fonctions de  $a, b, c, e, f, g, h$ , qui ne renferment pas le temps d'une manière explicite ; d'où l'on conclut ensuite que, par des éliminations entre les équations que nous citons, on obtiendra des valeurs de  $da, db, dc, de, df, dg, dh$ , exprimés au moyen des différences partielles de  $\Omega$  prises par rapport à  $a, b, c$ , &c., et multipliées par des fonctions de ces constantes. Nos formules ont l'avantage de donner immédiatement ces valeurs. »

En effet, les formules de Poisson, données à la page 282 du mémoire en question, sont bien les coefficients  $A, B, C, F, \dots$  de Lagrange, donnés comme fonctions des conditions initiales. Il faut noter qu'au passage, Poisson note les parenthèses de Lagrange par des crochets, et ses propres crochets par des parenthèses, ce qui ajoute à la confusion.

La riposte de Lagrange ne tarde pas. Réalisant qu'il a laissé échapper un aspect relativement important de sa propre théorie, Lagrange publie quelques mois après un second mémoire *sur la théorie de la variation des constantes arbitraires dans les problèmes de mécanique*. Dans ce mémoire, lu à l'Institut le 19 février 1810, il donne sa version personnelle de l'inversion du système des parenthèses qui portent son nom. Il écrit dans l'introduction de ce mémoire [Lag7, p. 810] :

« Mais l'application des formules générales aux Problèmes particuliers demandait encore un long calcul, à cause des éliminations qu'il fallait faire pour obtenir séparément l'expression de la variation de chacune des constantes devenues variables. Heureusement une considération très-simple, que je vais exposer et qui m'avait échappé, facilite et simplifie extrêmement cette application et ne laisse plus rien à désirer dans la Théorie analytique de la variation des constantes, relativement aux questions de la Mécanique. »

Toutefois, le mémoire de Lagrange ne se réduit pas à celui de Poisson, comme il l'écrit à la page 812 :

« M. Poisson a lu, le 16 octobre dernier à cette Classe, un *Mémoire sur la variation des constantes arbitraires dans les questions de Mécanique*,

lequel est imprimé dans le volume qui vient de paraître du *Journal de l'École Polytechnique*. Ce Mémoire contient une savante analyse qui est comme l'inverse de la mienne, et dont l'objet est d'éviter les éliminations que celle-ci exigeait. L'Auteur parvient en effet, par un calcul assez long et délicat, à des formules qui donnent directement les valeurs des différentielles des constantes arbitraires devenues variables. Ces formules ne coïncident pas immédiatement avec celles que je donne dans ce Mémoire, parce qu'elles renferment les constantes arbitraires en fonction des variables du Problème et de leurs différentielles, au lieu que les nôtres ne renferment ces constantes qu'en fonction d'autres constantes ; mais il est facile de se convaincre *à priori* qu'elles conduisent aux mêmes résultats. »

Autrement dit, Lagrange exprime la matrice inverse de son système de parenthèses en termes d'un autre système de coordonnées de l'espace des solutions du problème. Ce qui, effectivement, ne coïncide pas immédiatement avec l'analyse de Poisson. Ce que publie Lagrange, c'est plutôt l'expression générale des *transformations canoniques*. Que faut-il alors retenir de ces deux points de vue, de Poisson et de Lagrange, de ces analyses qui, tout en étant différentes, sont si semblables ? Voici quelques remarques qui, prises ensemble, peuvent expliquer le « retard » de Lagrange :

1. Les parenthèses de Lagrange, dans son premier mémoire de 1808, sont les coordonnées covariantes de la forme symplectique ; les crochets de Lagrange, dans son mémoire de 1810, en sont les coordonnées contravariantes. Qu'elles soient exprimées en fonction d'autres constantes du mouvement ou en tant que fonction des conditions initiales (temps, position, vitesse), c'est en tant que coordonnées de la forme symplectique qu'elles sont exploitées pour étudier les problèmes des inégalités séculaires<sup>2</sup> (théorie des perturbations).
2. C'est la matrice des parenthèses, fonction des constantes du mouvement, qui est utile, et utilisée (voir paragraphe II.3), pour l'intégration des perturba-

---

<sup>2</sup>L'apparition des crochets de Poisson comme méthode pour obtenir de nouvelles constantes du mouvement en fonction de certaines connues est légèrement postérieure.

tions. Les crochets de Poisson sont difficilement exploitables pour cet usage puisqu'ils sont exprimés en fonction des conditions initiales.

3. En revanche, si les parenthèses de Lagrange mettent en évidence la structure symplectique de l'espace des solutions des systèmes de la mécanique, le théorème de Poisson met en évidence la structure d'algèbre de Lie des fonctions sur cette variété symplectique.

Pourquoi l'histoire a-t-elle retenu davantage les crochets de Poisson plutôt que les parenthèses de Lagrange<sup>3</sup>, tout au moins dans les manuels de mécanique ? Peut-être parce que la seule façon de bien comprendre ces constructions de Lagrange c'est de les placer dans le cadre formel de la géométrie symplectique et que la géométrie de l'époque n'était pas encore prête ? Peut-être aussi parce que, si la construction de Lagrange est bien adaptée aux systèmes dont on connaît *a priori* l'intégration complète dans certaines conditions (on pourrait parler de *complète intégrabilité*), elle l'est moins lorsque l'on ne peut pas considérer le système réel comme une perturbation d'un système intégré. Mais, dans ces conditions, l'approche poissonnienne est-elle mieux adaptée ? On peut réfléchir aussi à la question suivante : pourquoi a-t-il fallu plus d'un siècle et demi pour redécouvrir l'œuvre de Lagrange, effacée devant celle de ses épigones ? Je laisse là ces questions à la méditation du lecteur.

NOTE. Siméon Denis Poisson est né à Phitiviers en 1781. Il fut l'élève de Laplace et de Lagrange. Il enseigna à l'École polytechnique à partir de 1802 et fut nommé astronome au bureau des longitudes en 1808. Il meurt à Sceaux en 1840.

NOTE. Joseph-Louis Lagrange est né à Turin en 1736, d'une famille originaire de France. Il succède à Euler à Berlin en 1766, et y reste jusqu'en 1787, date à laquelle il vient à Paris en tant que membre vétérane de l'Académie des sciences. Il publie en 1788<sup>4</sup> la première version de sa *Mécanique Analytique*, puis publie

<sup>3</sup>Pour ajouter à la confusion, entre les premiers articles de Lagrange et sa *Mécanique Analytique* les parenthèses deviennent crochets et les crochets, parenthèses. Alors que dans l'article de Poisson, ce que nous appelons *crochets de Poisson* est noté par des parenthèses. Autrement dit, l'influence de Poisson sur Lagrange aura quand même conduit ce dernier à changer de notations.

<sup>4</sup>Prête semble-t-il en 1782.

en 1811 la première partie de la deuxième version qui contient pour la première fois des éléments de calcul symplectique. Il meurt à Paris en 1813 et le deuxième volume de sa *Mécanique* n'est publié qu'après sa mort en 1816.

Enfin, je conseille vivement la lecture de l'ensemble des textes de Lagrange : trois mémoires déjà cités [Lag5, Lag6, Lag7] et la *Mécanique Analytique* [Lag8]. L'écriture de Lagrange est agréable, explicative et plus moderne qu'on ne peut l'imaginer *a priori*. En particulier, les introductions des chapitres sont de pures merveilles et des textes essentiels pour la compréhension de son œuvre et de l'histoire de la mécanique. Il vaut mieux éviter toutefois, en première lecture, les notes de Joseph Bertrand qui émaillent le texte de la *Mécanique Analytique* dans sa troisième édition. Elles portent sur des détails souvent peu importants et ne commentent pas les idées nouvelles qui y fourmillent.

---

# Histoire d'H

---

Dans cette annexe, j'essaye de faire le point sur les origines de la notation  $H$  que l'on utilise ordinairement, dans les cours de mécanique, pour désigner le hamiltonien. Doit-on la considérer comme l'initiale de Hamilton ? Ou bien comme celle de Huygens ainsi qu'on peut le lire parfois<sup>1</sup> ? Pour répondre à cette question il faut (re)lire Lagrange et Huygens ; on y découvre une partie du cheminement des idées qui ont conduit à cette découverte fondamentale du *principe de conservation des forces vives*, connu aujourd'hui sous le nom de *théorème de conservation de l'énergie*. C'est le plaisir que j'ai eu à parcourir ces quelques pages d'histoire que j'aimerais partager ici.

## Le premier H

Dans les manuels de mécanique analytique, la lettre  $H$  désigne traditionnellement une fonction appelée *hamiltonien* et qui représente l'énergie totale du système étudié, somme de son énergie cinétique et du potentiel des forces d'interaction. Cette notation est ainsi présentée comme un hommage des modernes à Sir Rowan Hamilton. Mais cette interprétation est erronée et résulte d'un malentendu historique. Ainsi que le fait remarquer Souriau dans ses divers textes sur la mécanique symplectique [Sou5], c'est Lagrange qui désigna pour la première fois, par la lettre  $H$ , la constante des forces vives<sup>2</sup>, non comme un hommage à *Sir Rowan Hamilton*

---

<sup>1</sup>Plus précisément dans les divers textes de Souriau sur la mécanique [Sou5].

<sup>2</sup>Si la *force vive* désigne le double de l'énergie cinétique, la *constante des forces vives*, elle, représente l'énergie totale.

mais à *Christiaan Huyghens*. Que ces deux grands mathématiciens Huygens et Hamilton, partageant la même initiale a conduit à ce malheureux malentendu. Le *hamiltonien* devrait être ainsi renommé : *huyghensien* !

Chacun peut vérifier que, dans sa *Mécanique Analytique*, Lagrange désigne bien par la lettre  $H$  la constante des forces vives. On peut y lire, à l'article 33, page 268 [Lag8, tome I, seconde partie, deuxième édition] :

« L'équation précédente devient

$$\mathbf{S} \left( \frac{dx^2 + dy^2 + dz^2}{dt^2} + d\Pi \right) m = 0$$

dont l'intégrale est

$$\mathbf{S} \left( \frac{dx^2 + dy^2 + dz^2}{2dt^2} + \Pi \right) m = \mathbf{H}$$

dans laquelle la lettre  $\mathbf{H}$  désigne une constante arbitraire, égale à la valeur du premier membre de l'équation à un instant donné.

Cette dernière équation renferme le principe connu sous le nom de *Conservation des forces vives*. En effet,  $dx^2 + dy^2 + dz^2$  étant le carré de l'espace que le corps parcourt dans l'instant  $dt$ ,  $\frac{dx^2 + dy^2 + dz^2}{dt^2}$  sera le carré de sa vitesse et  $\frac{dx^2 + dy^2 + dz^2}{dt^2} m$  sa force vive. Donc  $\mathbf{S} \frac{dx^2 + dy^2 + dz^2}{dt^2} m$  sera la somme des forces vives de tous les corps, ou la force vive de tout le système ; et l'on voit par l'équation dont il s'agit, que cette force vive est égale à la quantité  $2\mathbf{H} - 2\mathbf{S}\Pi m$ , laquelle dépend simplement des forces accélératrices qui agissent sur les corps, et nullement de leur liaison mutuelle, de sorte que la force vive du système est à chaque instant la même que les corps auraient acquises si, étant animés des mêmes puissances, ils s'étaient mus librement chacun sur la ligne qu'il a décrite. C'est ce qui fait donner le nom de *Conservation des forces vives* à cette propriété du mouvement. »

## Le principe des forces vives

C'est dans ce texte de sa *Mécanique Analytique*, et non dans les mémoires fondateurs [Lag5, Lag6, Lag7], que pour la première fois dans l'histoire de la mécanique la lettre  $H$  désigna la *constante des forces vives*, c'est-à-dire l'énergie totale du système. Nous continuons à observer cette convention mais avec une idée erronée de son origine.

Le *Principe de conservation des forces vives* est connu aujourd'hui sous le nom de *Théorème de conservation de l'énergie* ; mais ce changement de dénomination cache un glissement sémantique, justement suggéré par l'apparition de cette constante. En effet, comme on le voit clairement dans l'équation précédente, et comme il l'est exprimé (suggéré sans être dit), la force vive n'est évidemment pas conservée ! La tentative de justification de cette « conservation » reste confuse, et la question n'est vraiment réglée qu'avec l'introduction de la constante d'intégration  $H$ .

D'autre part, à la date de parution de cet ouvrage, en 1811, Hamilton n'avait que cinq ans. Le fait est indéniable :  $H$  n'est pas pour Hamilton, mais Lagrange n'explique pas, à cet endroit du texte (tome I, seconde partie, article 33), la raison de son choix. Il faut aller chercher ailleurs l'exposé de ses motivations, plus précisément dans l'introduction à la seconde partie de la mécanique, première section : *Sur les principes de la dynamique*. Outre son intérêt immédiat, sa lecture éclaire ce petit mystère. Voici ce qui est écrit à l'article 14 de la première section de la seconde partie du tome I de la mécanique de Lagrange :

« Le premier de ces quatre principes, celui de la conservation des forces vives, a été trouvé par Huygens, mais sous une forme un peu différente de celle qu'on lui donne présentement ; et nous en avons déjà fait mention à l'occasion du problème des centres d'oscillations.

[...]

Ainsi le principe de Huygens se réduit à ce que, dans le mouvement des corps pesants, la somme des produits des masses par les carrés des

vitesses à chaque instant est la même soit que les corps se meuvent conjointement d'une manière quelconque, ou qu'ils parcourent librement les mêmes hauteurs verticales.

[...]

Il [Huygens] donna ainsi à ce principe le nom de *conservation des forces vives*, et il s'en servit avec succès pour résoudre quelques problèmes qui ne l'avaient pas encore été, et dont il paraissait difficile de venir à bout par des méthodes directes. »

### **Le centre d'oscillations du pendule composé**

Au nombre des problèmes résolus par Huygens grâce au principe de conservation des forces vives, le plus important pour notre propos est celui du *centre d'oscillations*, qui lui fut posé par *le très savant Mersenne* (dixit Huygens [Huy, page 90 de l'édition originale, citation extraite de la traduction française]) :

Le très savant Mersenne me proposa jadis, presque encore enfant, et à de nombreux autres, la recherche du Centre d'Oscillations ou d'Agitation, question tout à fait fameuse parmi les Géomètres de ce temps. . .

Il s'agissait de déterminer le *pendule simple* isochrone à un *pendule composé*, c'est-à-dire le pendule simple dont le battement est identique à celui d'un pendule composé donné. Laissons de nouveau à Lagrange le soin de nous exposer le problème [article 7 de la première section de la seconde partie du tome I de l'ouvrage sus-cité] :

« Un fil considéré comme une ligne inflexible, sans pesanteur et sans masse, étant attaché par un bout à un point fixe, et chargé à l'autre bout d'un petit poids qu'on puisse regarder comme réduit à un point, forme ce qu'on appelle un pendule simple ; et la loi des vibrations de ce pendule dépend uniquement de sa longueur, c'est-à-dire de sa distance entre le poids et le point de suspension. Mais si à ce fil on attache encore

un ou plusieurs poids à différentes distances du point de suspension, on aura alors un pendule composé, dont le mouvement devra tenir une espèce de milieu entre ceux des différents pendules simples que l'on aurait, si chacun de ces poids était suspendu seul au fil. Car la force de gravité tendant d'un côté à faire descendre tous les poids également dans le même temps, et de l'autre l'inflexibilité du fil les contraignant à décrire des arcs inégaux et proportionnels à leur distance du point de suspension, il doit se faire entre ces poids une espèce de compensation, et de répartition de leurs mouvements, en sorte que les poids qui sont les plus proches du point de suspension hâteront les vibrations des plus éloignés, et ceux-ci, au contraire, retarderont les vibrations des premiers. Ainsi il y aura dans le fil un point ou un corps étant placé, son mouvement ne serait ni accéléré ni retardé par les autres poids, mais serait le même que s'il était seul suspendu au fil. Ce point sera donc le vrai centre d'oscillation du pendule composé, et un tel centre doit se trouver dans tout corps solide de quelque figure que ce soit, qui oscille autour d'un axe horizontal. »

Le problème est bien posé, et il est d'importance puisque les horloges sont réglées justement grâce à des poids répartis sur leur balancier. Continuons notre lecture de Lagrange :

« Huygens vit qu'on ne pouvait déterminer ce centre d'une manière rigoureuse sans connaître la loi suivant laquelle les différents poids du pendule composé altèrent mutuellement les mouvements que la gravité tend à leur imprimer à chaque instant ; mais au lieu de chercher à déduire cette loi des principes fondamentaux de la Mécanique, il se contenta d'y suppléer par un principe indirect, lequel consiste à supposer que si plusieurs poids attachés, comme on le voudra, à un pendule, descendent par la seule action de la gravitation, et que, dans un instant quelconque, ils soient détachés et séparés les uns des autres, chacun

---

d'eux, en vertu de la vitesse acquise pendant sa chute, pourra remonter à une telle hauteur, que le centre commun de gravité se trouvera remonté de la même hauteur d'où il était descendu. »

Voilà donc le principe, prémisse de la conservation des forces vives. Et Lagrange d'ajouter :

« On ne saurait deviner ce qui a donné à l'auteur l'idée d'un tel principe  
[...]

Quoi qu'il en soit, ce principe fournit une équation entre la hauteur verticale, d'où le centre de gravité du système est descendu dans un temps quelconque, et les différentes hauteurs verticales auxquelles les corps qui composent le système pourraient remonter avec leurs vitesses acquises, et qui, par le théorème de Galilée, sont comme le carré de ces vitesses. »

Voilà qui montre clairement comment la conservation des forces vives résout le problème du pendule composé. En vérité, le problème n'était pas encore considéré comme définitivement réglé. Les principes posés par Huygens n'étaient pas tout à fait ceux énoncés plus haut, et il fallait les fonder davantage. C'est Jacques Bernoulli qui définitivement satisfait les mécaniciens de l'époque. Mais tout cela est admirablement conté dans l'introduction à *La dynamique de la Mécanique* de Lagrange (articles 6 à 14) et nous y renvoyons le lecteur, pour son plaisir.

## Horologium Oscillatorium

Cela dit, avant de clore définitivement ce chapitre il nous faut encore préciser où et comment Huygens a présenté lui-même sa découverte. Pour cela, il faut lire la proposition V de la quatrième partie de l'*Horologium Oscillatorium* [Huy, page 99 de l'édition originale, citation extraite de la traduction française] :

## PROPOSITION V

« Étant donné un pendule composé des poids que l'on veut, si chacun est multiplié par le carré de sa distance à l'axe d'oscillation, et que la somme des produits soit divisée par le produit fait de la somme des poids par la distance au même axe d'oscillation du centre de gravité commun à tous ; il paraît la longueur du pendule simple isochrone au composé, soit la distance entre l'axe et le centre d'oscillation même du pendule composé. »

C'est dans la démonstration de cette proposition, qui donne la solution du problème du pendule composé, qu'apparaît, cachée mais bien là, la constante des forces vives. On peut lire en effet, à la page 100 de l'ouvrage sus-cité :

« Mais comme le carré de la vitesse du point L qu'il possède en P est au carré de de la vitesse du point A en T, ainsi est la hauteur à laquelle L peut monter par la première vitesse, à la hauteur à laquelle A peut monter par la deuxième vitesse... »

Accordons définitivement à Huygens le mérite de la découverte de la conservation de l'énergie (comme on nomme aujourd'hui ce principe).

**Conclusion**

Ainsi donc le mystère de la lettre  $H$  est éclairci, mais il est important de souligner qu'après les Huygens, Bernoulli et autres, celui qui a donné son sens moderne à ce principe, qui l'a établi dans toute sa rigueur analytique, c'est Lagrange. Et c'est, encore aujourd'hui, la même façon qu'il a eu de le démontrer, il y a deux cent ans, que ce principe est enseigné dans nos facultés.

---

# Éléments de géométrie symplectique

---

Cette annexe est destinée à présenter les éléments essentiels fondant la géométrie symplectique moderne. Ce n'est toutefois pas un cours de géométrie symplectique ; bien entendu le lecteur peut se reporter aux ouvrages classiques cités en référence pour un exposé systématique. Pour ce qui suit, nous considérons le lecteur familier à la fois avec l'algèbre linéaire ordinaire et les éléments de base de géométrie différentielle.

## Espaces vectoriels symplectiques

Soit  $E$  un espace vectoriel réel de dimension finie. On appelle *forme bilinéaire alternée* (ou antisymétrique) de  $E$  toute application

$$\omega : E \times E \rightarrow \mathbf{R} \tag{C.1}$$

telle que :

$$\begin{aligned} \omega(u + v, w) &= \omega(u, w) + \omega(v, w) \\ \omega(\lambda u, v) &= \lambda \omega(u, v) \\ \omega(u, v) &= -\omega(v, u), \end{aligned} \tag{C.2}$$

pour tout  $u, v, w$  vecteurs de  $E$ , et  $\lambda$  réel.

On appelle *noyau* de la forme  $\omega$  le sous-espace vectoriel  $\ker \omega \subset E$  défini par

$$\ker \omega = \{u \in E \mid \omega(u, v) = 0 \quad \forall v \in E\}. \quad (\text{C.3})$$

On considère en même temps la forme  $\omega$  comme une application linéaire de  $E$  dans son dual grâce à l'identification :

$$\begin{aligned} \omega : E &\longrightarrow E^* \\ u &\longmapsto \bar{u} : v \mapsto \omega(u, v) \end{aligned} \quad (\text{C.4})$$

On note aussi  $\bar{u} = \omega(u, \cdot)$  ou encore  $\bar{u} = \omega(u)$ , de sorte que  $\omega(u, v) = \bar{u}(v) = \omega(u)(v)$ . Le noyau de  $\omega$  est le noyau de cette application linéaire.

**DÉFINITION.** *Étant donnée une forme bilinéaire alternée  $\omega$  définie sur l'espace vectoriel  $E$ , on dit qu'elle est symplectique si elle est non dégénérée, c'est-à-dire de noyau nul :  $\ker \omega = \{0\}$ .*

Puisque nous avons une application bilinéaire sur  $E$ , nous avons par la même occasion les notions d'orthogonalité, de sous-espace orthogonal à un sous-espace vectoriel donné :

**DÉFINITION.** *Étant donnés deux vecteurs  $u$  et  $v$  de  $E$ , nous disons qu'ils sont orthogonaux pour  $\omega$  (ou tout simplement orthogonaux) si  $\omega(u, v) = 0$ . Étant donné un sous-espace  $V \subset E$ , nous appelons sous-espace orthogonal à  $V$  l'ensemble :*

$$\text{Orth}(V) = \{u \in E \mid \omega(u, v) = 0 \quad \forall v \in V\}.$$

Cela permet de réinterpréter le noyau  $\ker \omega$  comme l'orthogonal de  $E$  tout entier :  $\ker \omega = \text{Orth}(E)$ . Dans ce point de vue,  $\omega$  est symplectique si  $\text{Orth}(E) = \{0\}$ .

L'orthogonalité d'une forme bilinéaire alternée s'accompagne naturellement des notions d'*isotropie* et de *co-isotropie* :

DÉFINITION. *Nous disons qu'un sous-espace vectoriel  $V$  de  $E$  est isotrope s'il est contenu dans son orthogonal :  $V \subset \text{Orth}(V)$ . Nous disons qu'il est co-isotrope s'il contient son orthogonal :  $\text{Orth}(V) \subset V$ .*

À titre d'exemples : le noyau de  $\omega$  est évidemment isotrope ; toute droite vectorielle est isotrope (appliquer l'antisymétrie) ; tout hyperplan est co-isotrope. Considérons la restriction de  $\omega$  sur un sous-espace vectoriel  $V$ , son noyau est donné par :

$$\ker \omega|_V = \text{Orth}(V) \cap V. \quad (\text{C.5})$$

L'espace quotient  $V/\ker \omega|_V = V/(\text{Orth}(V) \cap V)$  est donc muni naturellement par projection d'une forme symplectique. On appelle cette méthode de construction d'espace vectoriel symplectique la *méthode de réduction symplectique*. Lorsque  $V$  est co-isotrope alors  $\text{Orth}(V) \cap V = \text{Orth}(V)$  et le quotient est simplement  $V/\text{Orth}(V)$ .

On obtient encore, comme conséquence de cette dernière égalité, que si la restriction de  $\omega$  sur  $V$  est symplectique alors  $V$  et  $\text{Orth}(V)$  sont complémentaires, et même supplémentaires compte tenu du théorème suivant :

THÉORÈME. *Soit  $E$  un espace vectoriel muni d'une forme symplectique  $\omega$ , soit  $V$  un sous-espace vectoriel quelconque de  $E$ , alors :*

$$\dim V + \dim \text{Orth}(V) = \dim E. \quad (\text{C.6})$$

DÉMONSTRATION. *Soit  $\mathcal{V} = (v_1, \dots, v_r)$  une base de  $V$ ,  $\dim V = r$ . Considérons l'application linéaire  $\tilde{\mathcal{V}} : E \rightarrow \mathbf{R}^r$  définie par :*

$$\tilde{\mathcal{V}}(w) = (\bar{v}_1(w), \dots, \bar{v}_r(w)). \quad (\text{C.7})$$

*En remarquant que le noyau de cette application est naturellement l'orthogonal symplectique de  $V$  et en appliquant le théorème classique d'algèbre linéaire :*

$$\begin{aligned} \dim E &= \dim \text{im } \tilde{\mathcal{V}} + \dim \ker \tilde{\mathcal{V}} \\ &= \dim V + \dim \text{Orth}(V), \end{aligned}$$

*on établit l'égalité annoncée. ■*

Mais attention, cela ne veut pas dire que pour tout sous-espace  $V$ ,  $V$  et  $\text{Orth}(V)$  sont supplémentaires ; par exemple, si  $V$  est une droite, son orthogonal est un hyperplan contenant  $V$ . Cette formule indique aussi que la dimension d'un espace isotrope (ou co-isotrope) est nécessairement inférieure, ou égale, à la moitié de la dimension de l'espace. Ce qui conduit à la définition suivante :

**DÉFINITION.** *On appelle sous-espace vectoriel lagrangien tout sous-espace vectoriel isotrope saturé<sup>1</sup>.*

La dimension d'un espace lagrangien est nécessairement de dimension moitié de celle de l'espace symplectique comme l'implique le théorème fondamental suivant :

**THÉORÈME. (DE LA BASE CANONIQUE)** *Soit  $\omega$  une forme bilinéaire alternée définie sur un espace vectoriel  $E$ , il existe une base*

$$S = (u_1, \dots, u_n, v_1, \dots, v_n, k_1, \dots, k_r)$$

de  $E$  telle que :

1.  $\ker \omega = \langle k_1, \dots, k_r \rangle$ ,
2.  $\omega(u_i, u_j) = 0$ ,  $\omega(v_i, v_j) = 0$  et  $\omega(u_i, v_j) = \delta_{ij}$ ,

où  $\delta$  est le symbole de Kronecker et les symboles  $\langle \dots \rangle$  désignent le sous-espace vectoriel engendré. Une telle base est appelée base canonique.

**DÉMONSTRATION.** *Soit  $k_1, \dots, k_r$  une base de  $\ker \omega$  et  $F$  un supplémentaire de  $\ker \omega$ . Alors  $\omega|_F$  est symplectique et on peut compléter  $k_1, \dots, k_r$  par une base de  $F$  pour en faire une base de  $E$ . Le problème se ramène donc au cas symplectique. Supposons donc que  $E$  soit symplectique. Soit  $u_1$  un vecteur non nul, il existe alors un vecteur  $v_1$  tel que  $\omega(u_1, v_1) \neq 0$ , sinon  $u_1$  serait dans le noyau de  $\omega$ . On*

<sup>1</sup>C'est sous ce nom : Variétés isotropes saturées que les variétés lagrangiennes (sous-variétés tangentes en chaque point à un espace vectoriel lagrangien) sont apparues originellement en géométrie symplectique [Sou1].

peut alors choisir  $v_1$  pour que  $\omega(u_1, v_1) = 1$ . Soit  $E_2$  l'orthogonal symplectique du sous-espace vectoriel engendré par le couple  $(u_1, v_1)$ ,  $E_2 = \text{Orth}(\langle u_1, v_1 \rangle)$ . Les espaces  $E_2$  et  $\langle u_1, v_1 \rangle$  sont supplémentaires et la restriction de  $\omega$  à  $E_2$  est donc symplectique, comme nous l'avons vu plus haut. Il suffit donc d'appliquer cette procédure un nombre fini de fois puisque la dimension de  $E = E_1$  est finie et qu'à chaque étape la dimension des  $E_i$  successifs décroît de 2. Il suffit alors de vérifier que les  $u_i$  et  $v_i$  ainsi choisis vérifient bien les conditions du théorème. ■

On remarque que si  $E$  est symplectique, sa dimension est paire ; autrement dit, le rang d'une forme antisymétrique est pair. Dans une base canonique, la matrice de Gramm  $\Omega$  de la forme  $\omega$  s'écrit :

$$\Omega = \begin{pmatrix} \mathbf{0}_n & \mathbf{1}_n \\ -\mathbf{1}_n & \mathbf{0}_n \end{pmatrix}. \quad (\text{C.8})$$

EXEMPLE. L'exemple standard d'espace vectoriel symplectique est  $\mathbf{R}^{2n}$  : soient

$$X = (u, v) \text{ et } X' = (u', v')$$

deux vecteurs de  $\mathbf{R}^{2n} = \mathbf{R}^n \times \mathbf{R}^n$ , la forme symplectique canonique  $\omega$  est définie par

$$\omega(X, X') = u.v' - u'.v, \quad (\text{C.9})$$

où le point désigne le produit scalaire ; autrement dit :

$$\omega(X, X') = \sum_{i=1}^n u_i v'_i - u'_i v_i. \quad (\text{C.10})$$

La base canonique  $(e_1, \dots, e_n, e_{n+1}, \dots, e_{2n})$  est symplectique. La forme symplectique  $\omega$  s'appelle la *forme symplectique canonique* de  $\mathbf{R}^{2n}$ .

REMARQUE. On peut généraliser certaines de ces constructions à la dimension infinie mais il faut faire attention à choisir les bonnes définitions. Pour le lecteur averti, voici un exemple amusant qui mériterait de plus amples développements. On considère une surface compacte orientée  $\Sigma$ , de genre  $g$ . On considère l'espace

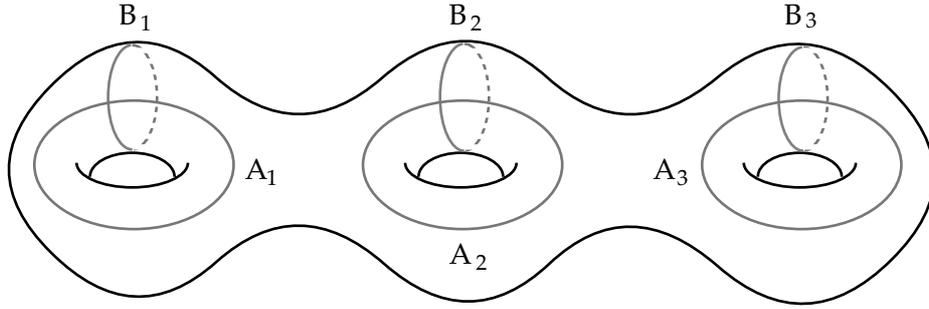


Figure C.1 : Surface de genre 3 et base symplectique de l'homologie  $(A_i, B_i)$

vectorel  $E$  de ses 1-formes différentielles. C'est un espace de dimension infini. On définit sur cet espace vectoriel la forme bilinéaire suivante :

$$\omega(a, b) = \int_{\Sigma} a \wedge b. \quad (\text{C.11})$$

Cette définition a bien un sens puisque : 1) la surface  $\Sigma$  est compacte et donc l'intégrale converge ; 2) la surface  $\Sigma$  est orientée. La forme bilinéaire  $\omega$  est antisymétrique, conséquence des lois du produit extérieur. Elle est non dégénérée dans le sens ordinaire :

$$\omega(a, b) = \int_{\Sigma} a \wedge b = 0 \quad \forall b \quad \Rightarrow a = 0. \quad (\text{C.12})$$

On peut donc dire que  $\omega$  est une forme symplectique sur  $E$ . Il n'est pas question d'établir dans ce cas le théorème de la base symplectique, tout au plus peut-on espérer montrer l'existence de sous-espaces vectoriels, isotropes, co-isotropes et finalement lagrangiens.

Considérons le sous-espace vectoriel  $B$  des formes différentielles exactes (dérivées extérieures de fonctions réelles). Cet espace est isotrope, en effet, pour tout couple  $f$  et  $g$  de fonctions :

$$\omega(df, dg) = \int_{\Sigma} df \wedge dg = \int_{\Sigma} d(fdg) = \int_{\partial\Sigma} fdg = 0. \quad (\text{C.13})$$

Montrons que son orthogonal symplectique est le sous-espace des formes différentielles fermées  $Z$  de  $\Sigma$  :

$$\omega(df, a) = \int_{\Sigma} df \wedge a = \int_{\Sigma} d(fa) - \int_{\Sigma} f da = - \int_{\Sigma} f da; \quad (\text{C.14})$$

donc  $\omega(df, a) = 0$  pour tout  $f$  implique  $\int_{\Sigma} f da = 0$ , c'est-à-dire  $da = 0$ , ce qui se résume ainsi :

$$\text{Orth}(B) = Z \quad \Leftrightarrow \quad \text{Orth}(Z) = B. \quad (\text{C.15})$$

Nous avons vu plus haut que la méthode de réduction symplectique munit le quotient d'un espace co-isotrope par son orthogonal d'une structure symplectique canonique. Ici, le sous-espace co-isotrope choisi est  $Z$  et le quotient symplectique  $Z/B$  est le premier espace de cohomologie de De Rham  $H^1(\Sigma, \mathbf{R})$  :

$$Z/(B = \text{Orth}(Z)) = H^1(\Sigma, \mathbf{R}). \quad (\text{C.16})$$

Nous savons par ailleurs, grâce à la théorie des surfaces, que cet espace est de dimension finie :  $\dim H^1(\Sigma, \mathbf{R}) = 2g$ , autrement dit  $Z/(B = \text{Orth}(Z)) \simeq \mathbf{R}^{2g}$ . Nous savons que  $\mathbf{R}^{2g}$  possède des espaces vectoriels lagrangiens, que l'on peut d'ailleurs tous construire grâce à la méthode de la base symplectique. Soit  $L$  un de ces espaces et  $\tilde{L}$  sa pré-image par la projection canonique  $Z \rightarrow H^1(\Sigma, \mathbf{R})$ . Compte tenu de ce qui précède, le lecteur peut vérifier lui-même que ce sous-espace vectoriel  $\tilde{L}$  est lagrangien, c'est-à-dire que c'est un sous-espace vectoriel isotrope saturé de l'espace  $E$  des 1-formes différentielles sur  $\Sigma$ .

Il ne saurait être question de géométrie symplectique sans l'introduction des « groupes symplectiques ».

**DÉFINITION.** On appelle groupe symplectique de l'espace vectoriel symplectique  $(E, \omega)$  le sous-groupe  $\text{Sp}(E, \omega)$  du groupe linéaire  $\text{GL}(E)$  des éléments préservant  $\omega$ , c'est-à-dire :

$$\text{Sp}(E, \omega) = \{A \in \text{GL}(E) \mid \omega(Au, Av) = \omega(u, v) \quad \forall (u, v) \in E \times E\}. \quad (\text{C.17})$$

Ce groupe est noté  $\text{Sp}(2n, \mathbf{R})$  lorsque  $E = \mathbf{R}^{2n}$  muni de sa forme symplectique canonique.

Il est bien évident que les groupes symplectiques d'espaces vectoriels symplectiques de même dimension sont isomorphes, puisque les espaces vectoriels symplectiques le sont eux-mêmes, d'après le théorème de la base symplectique. Il est inutile de préciser que ce groupe joue un rôle fondamental en géométrie symplectique, et nous renvoyons aux ouvrages spécialisés pour plus de détails.

Venons en à la *géométrie différentielle symplectique*. La définition d'une *variété symplectique* copie la définition des espaces vectoriels symplectiques dans le sens qu'une structure symplectique sur une variété est définie par la donnée sur chaque espace tangent d'une forme symplectique, à ceci près que l'on demande, en plus, à cette forme d'être fermée. Donnons la définition formelle :

DÉFINITION. *Soit  $X$  une variété différentielle, on appelle forme symplectique sur  $X$  toute 2-forme différentielle  $\omega$  non dégénérée fermée :*

$$\forall x \in X \quad \ker \omega_x = \{0\} \quad \text{et} \quad d\omega = 0.$$

*Une variété munie d'une forme symplectique s'appelle variété symplectique.*

Dans cette définition,  $\omega_x$  désigne la valeur de la forme  $\omega$  au point  $x$ , c'est donc une forme bilinéaire antisymétrique définie sur l'espace vectoriel tangent  $T_x X$ , et  $d\omega$  désigne la dérivée extérieure de  $\omega$ . Pour ces définitions de géométrie différentielle ordinaire, nous renvoyons le lecteur aux ouvrages classiques, ou bien à l'annexe suivante D pour la notion de différentielle extérieure.

EXEMPLE. L'exemple classique de variété différentielle est celui de l'espace cotangent  $T^*Q$  à une variété quelconque  $Q$ . Nous avons déjà rencontré cet exemple au chapitre III.2. Nous n'y reviendrons pas.

NOTE. Les notions de *sous-variété isotrope*, *sous-variété co-isotrope*, *sous-variété lagrangienne* se généralisent naturellement à partir de l'algèbre linéaire : ce sont des sous-variétés (plongée ou immergée, suivant les cas) qui sont tangentes en chaque point à un sous-espace vectoriel, respectivement, isotrope, co-isotrope ou lagrangien de l'espace tangent.

Le théorème essentiel de la géométrie symplectique est le théorème suivant, dû à Gaston Darboux, qui dit que toute variété symplectique est non seulement infinitésimalement isomorphe à  $\mathbf{R}^{2n}$  canonique, comme nous pourrions nous y attendre, mais même *localement* isomorphe, c'est-à-dire sur un voisinage entier et non seulement en un point.

THÉORÈME. (DARBOUX) *Soit  $(X, \omega)$  une variété symplectique de dimension  $2n$ . Soit  $x_0 \in X$ , il existe une carte  $F : U \rightarrow X$ , où  $U$  est un voisinage de  $0 \in \mathbf{R}^{2n}$ , envoyant  $0$  sur  $X_0$ , et telle que  $F^*(\omega)$  coïncide, sur son domaine de définition, avec la forme symplectique canonique de  $\mathbf{R}^{2n}$ . Une telle carte est appelée carte de Darboux.*

Pour démontrer ce théorème fondamental, nous utiliserons le suivant, dû à Moser :

THÉORÈME. (MOSER) *Soit  $X$  une variété symplectique compacte (sans bord), soient  $\omega$  et  $\omega'$  deux formes symplectiques sur  $X$  homotopes à travers la même classe de cohomologie. Il existe alors un difféomorphisme  $\phi$  de  $X$ , homotope à l'identité, tel que  $\phi^*\omega' = \omega$ .*

Nous admettrons le lemme suivant

LEMME. *Soit  $X$  une variété compacte (sans bord). Soit  $[t \mapsto \omega_t]$  un arc différentiable dans l'espace des  $p$ -formes exactes, dans le sens suivant<sup>2</sup> : l'application  $[(t, x) \mapsto \omega_t(x)]$  est différentiable sur un voisinage ouvert de  $[0, 1] \times X$  dans l'espace des  $p$ -formes et prend ses valeurs dans le sous-espace des formes fermées exactes. Il existe alors un arc différentiable  $[t \mapsto \beta_t]$  de  $(p-1)$ -formes tel que pour tout  $t$  :  $\omega_t = \omega_0 + d\beta_t$ .*

DÉMONSTRATION. (THÉORÈME DE MOSER) *On suppose l'homotopie  $[t \mapsto \omega_t]$ , définie sur l'intervalle  $] - \varepsilon, 1 + \varepsilon[$ ,  $\varepsilon > 0$ , ce qui est une conséquence de la différentiabilité de l'arc et notons :*

$$\omega_t = \omega_0 + d\beta_t,$$

---

<sup>2</sup>C'est la différentiabilité au sens des *espaces différentiables*, voir annexe D.

où  $[t \mapsto \beta_t]$  est un arc différentiable de 1-formes (d'après le lemme précédent). Soit  $Y$  la variété  $] -\varepsilon, 1 + \varepsilon[ \times X$  sur laquelle on prolonge chaque 2-forme  $\omega_t$  par zéro sur  $dt$  ; soit  $\Omega$  ce prolongement :

$$\Omega_{(t,x)}((s, u), (s', u')) = (\omega_t)_x(u, u').$$

Considérons le champ de vecteurs  $\Xi$  défini sur  $Y$  par :

$$\Xi \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} 1 \\ \xi_t(x) \end{pmatrix} \text{ avec } \xi_t(x) = -(\omega_t^{-1})_x \left( \frac{\partial \beta_t(x)}{\partial t} \right).$$

Pour tout  $t$  dans l'intervalle  $] -\varepsilon, 1 + \varepsilon[$ , notons  $j_t$  l'injection  $x \mapsto (t, X)$  de  $X$  dans  $\{t\} \times X$ .

Démontrons que l'image réciproque de la dérivée de Lie  $\mathcal{L}_\Xi \Omega$  par  $j_t$  est nulle pour tout  $t$ . Soient  $x \in X$ ,  $u$  et  $u'$  deux vecteurs tangents en  $x$  à  $X$ . Nous avons :

$$j_t^*(\mathcal{L}_\Xi \Omega)_x(u, u') = \mathcal{L}_\Xi \Omega_{(t,x)} \begin{pmatrix} 0 \\ u \end{pmatrix} \begin{pmatrix} 0 \\ u' \end{pmatrix}.$$

Appliquons la formule de Cartan :  $\mathcal{L}_\Xi \Omega = d\Omega(\Xi) + d(\Omega(\Xi))$ , et calculons les deux termes de droite de l'égalité suivante :

$$j_t^*(\mathcal{L}_\Xi \Omega)_x(u, u') = d\Omega(\Xi)_{(t,x)} \begin{pmatrix} 0 \\ u \end{pmatrix} \begin{pmatrix} 0 \\ u' \end{pmatrix} + d(\Omega(\Xi))_{(t,x)} \begin{pmatrix} 0 \\ u \end{pmatrix} \begin{pmatrix} 0 \\ u' \end{pmatrix}.$$

Pour cela calculons d'abord  $d\Omega$ . Exprimons  $\Omega$  dans une carte adaptée :

$$\Omega_{(t,x)} = \sum_{i < j} (\omega_t)_{ij}(x) dx_i \wedge dx_j.$$

On déduit la dérivée extérieure

$$d\Omega_{(t,x)} = d\omega_t + \sum_{i < j} \frac{\partial (\omega_t)_{ij}(x)}{\partial t} dt \wedge dx_i \wedge dx_j = d\omega_t + dt \wedge \frac{\partial \omega_t(x)}{\partial t}$$

mais  $\omega_t$  étant fermée, par hypothèse, il reste :

$$d\Omega_{(t,x)} = dt \wedge \frac{\partial \omega_t(x)}{\partial t}.$$

Contracté avec  $\Xi$  et appliqué au couple de vecteurs  $(0, u)$ ,  $(0, u')$ , il vient le premier terme de  $j_t^*(\mathcal{L}_\Xi \Omega)_x(u, u')$  :

$$\begin{aligned} d\Omega_{(t,x)}(\Xi(t,x)) \begin{pmatrix} 0 \\ u \end{pmatrix} \begin{pmatrix} 0 \\ u' \end{pmatrix} &= d\Omega_{(t,x)} \begin{pmatrix} 1 \\ \xi_t(x) \end{pmatrix} \begin{pmatrix} 0 \\ u \end{pmatrix} \begin{pmatrix} 0 \\ u' \end{pmatrix} \\ &= \frac{\partial \omega_t(x)}{\partial t}(u, u'). \end{aligned}$$

D'autre part, pour tout vecteur  $(s, u)$  tangent à  $Y$  au point  $(t, x)$ , nous avons :

$$\begin{aligned} \Omega_{(t,x)}(\Xi(t,x)) \begin{pmatrix} s \\ u \end{pmatrix} &= \Omega_{(t,x)} \begin{pmatrix} 1 \\ \xi_t(x) \end{pmatrix} \begin{pmatrix} s \\ u \end{pmatrix} \\ &= (\omega_t)_x(\xi_t(x))(u) \\ &= (\omega_t)_x \left( -(\omega_t^{-1})_x \left( \frac{\partial \beta_t(x)}{\partial t} \right) \right) (u) \\ &= -\frac{\partial \beta_t(x)}{\partial t}(u). \end{aligned}$$

On en déduit, puisque le terme de dérivée seconde en  $t$  n'intervient pas, que :

$$\begin{aligned} d(\Omega(\Xi)) \begin{pmatrix} 0 \\ u \end{pmatrix} \begin{pmatrix} 0 \\ u' \end{pmatrix} &= -d \left( \frac{\partial \beta_t}{\partial t} \right)_x (u, u') \\ &= -\frac{\partial (d\beta_t)_x}{\partial t}(u, u'). \end{aligned}$$

On a obtenu ainsi les deux termes de la dérivée de Lie recherchée :

$$\begin{aligned} j_t^*(\mathcal{L}_\Xi \Omega)_x(u, u') &= \frac{\partial \omega_t(x)}{\partial t}(u, u') - \frac{\partial (d\beta_t)_x}{\partial t}(u, u') \\ &= \frac{\partial}{\partial t} \left[ (\omega_t)_x(u, u') - (d\beta_t)_x(u, u') \right]. \end{aligned}$$

En comparant avec l'hypothèse  $\omega_t = \omega_0 + d\beta_t$ , on conclut :

$$j_t^*(\mathcal{L}_\Xi \Omega)_x = 0.$$

Autrement dit, l'image réciproque de la forme  $\Omega$  par le flot de  $\Xi$  et  $\Omega$  coïncident en restriction à chaque sous-variété  $\{t\} \times X$ . On a alors en particulier :

$$j_0^* \left( (e^{s\Xi})^* \Omega \right) = (e^{s\Xi} \circ j_0)^* \Omega = j_0^* \Omega.$$

On reconnaît  $\omega_0$  dans le terme de droite. D'autre part  $e^{s\Xi} \circ j_0$  est un difféomorphisme  $\phi_s$  de  $X$  sur  $\{s\} \times X \simeq X$ , donc :

$$(e^{s\Xi} \circ j_0)^* \Omega = \phi_s^*(\Omega|_{\{s\} \times X}) = \phi_s^*(\omega_s).$$

Enfin, le champ de vecteurs  $\Xi$  est complet sur le produit  $[0, 1] \times X$ , grâce à la compacité de  $X$  et parce qu'il est transverse aux fibres  $\{t\} \times X$ . Il échange en particulier les fibres  $\{t\} \times X$ , plus précisément  $e^{s\Xi}(\{t\} \times X) = \{t+s\} \times X$ . Nous pouvons alors choisir  $s = 1$  ; le difféomorphisme  $e^\Xi$  envoie  $\{0\} \times X$  sur  $\{1\} \times X$  et nous avons le résultat attendu :  $\phi^* \omega_1 = \omega_0$ . ■

Venons en à la démonstration du théorème de Darboux :

DÉMONSTRATION. (THÉORÈME DE DARBOUX) Soit  $F : U \rightarrow X$  une carte telle que  $F(0) = x_0$ , l'image réciproque  $F^* \omega$  est une forme symplectique définie sur un ouvert  $U$  de  $\mathbf{R}^{2n}$ . Le problème est donc ramené au cas où  $X = U$ , ce que nous supposons maintenant. Il s'agit alors de trouver un difféomorphisme  $\phi$  d'un voisinage  $V \subset U$  qui fixe 0 et qui transforme la forme symplectique  $\omega$  en la forme canonique sur tout  $V$ . On peut déjà supposer que la valeur  $\omega_0 = \omega(0)$ , est la forme canonique, en appliquant le théorème de la base symplectique. Considérons le chemin  $t \mapsto \omega_t = t\omega + (1-t)\omega_0$  qui joint  $\omega_0$ , défini sur tout  $U$ , et  $\omega$ . Comme  $(\omega_t)(0) = t\omega(0) + (1-t)\omega_0(0) = \omega_0$  pour tout  $t$ , il existe une boule  $B$ , voisinage de 0 dans  $U$ , telle que  $\omega_t$  soit symplectique sur  $B$  pour tout  $t$  dans un certain voisinage de  $[0, 1]$ . En effet, la condition de fermeture est évidemment vérifiée. Reste la régularité, qui s'exprime par la condition :  $\det \Omega_t(x) \neq 0$ . Mais le déterminant étant continu, puisque  $\det(\Omega_t)(0) \neq 0$  pour tout  $t$ , il existe (pour tout  $t$ ) une boule  $B_t \subset \mathbf{R} \times V$ , centrée en  $(t, 0)$ , telle que  $\det \Omega_t(x) \neq 0$  sur  $B_t$ . On a donc un recouvrement ouvert de  $[0, 1] \times \{0\}$ . Par compacité on en extrait un recouvrement fini  $(B_i)$ , et il suffit ensuite de prendre l'intersection  $B$  des boules  $B_i \cap (\{i\} \times U)$ . Nous sommes ainsi ramené à l'énoncé du théorème de Moser, à ceci près que la boule  $B$  n'est pas une variété compacte et que nous ne pouvons pas invoquer sa compacité pour en déduire la complétude du champ de vecteur  $\Xi$  de la

démonstration précédente, défini sur  $[0, 1] \times B$  (le rôle de la variété  $X$  est maintenant joué par  $B$ ). Mais la complétude du champ  $\Xi$ , condition suffisante, n'est pas nécessaire dans notre cas, il suffit que nous connaissions la courbe intégrale du champ de vecteur  $\Xi$  passant par le point  $(0, 0) \in [0, 1] \times B$ , c'est-à-dire l'application  $s \mapsto e^{s\Xi}(0, 0) = (s, 0)$ , qui envoie le point  $(0, 0)$  sur  $(1, 0)$ . En appliquant alors le théorème de Cauchy sur les équations différentielles ordinaires, on montre qu'il existe une sous-variété transverse  $W = \{0\} \times B'$  à  $\Xi$  telle que l'exponentielle restreinte à  $W$  soit un difféomorphisme de  $W$  sur son image  $e^{s\Xi}(W) \subset \{1\} \times B$  qui est un voisinage de  $0$  dans  $B$ . La démonstration du théorème de Moser s'applique alors sans modification formelle à cette situation. ■

REMARQUE. Toute variété symplectique  $(X, \omega)$  possède un volume<sup>3</sup>, il suffit de choisir  $\text{vol} = \omega^n$ , où  $\dim X = 2n$ . Réciproquement, toute variété orientable de dimension paire n'est pas nécessairement symplectique. Pour les sphères de dimension paire par exemple, seule la sphère  $S^2$  est symplectique. En effet, toute variété compacte  $X$  symplectique a son deuxième groupe d'homologie  $H_2(X, \mathbf{Z})$  non nul : si l'on suppose le contraire alors  $\omega = d\alpha$  donc  $\text{vol} = d(a \wedge \alpha^{n-1})$  et  $\int_X \text{vol} = 0$  ce qui est absurde. La seule sphère ayant son deuxième groupe d'homologie non nul est  $S^2$ . Remarquons à ce propos que la seule condition pour qu'une surface (variété de dimension 2) soit symplectique est d'être orientable. Mais le cas des surfaces est plutôt un cas dégénéré de la géométrie symplectique. Les théorèmes symplectiques y sont soit faux soit triviaux<sup>4</sup>.

REMARQUE. Le théorème de Darboux montre que la géométrie symplectique différentielle n'a pas d'invariants locaux, comme c'est le cas pour d'autres types de géométrie différentielle — par exemple la géométrie riemannienne où l'on a à sa disposition la courbure et autres caractéristiques de ce type. Toute structure symplectique est localement *plate*. Les seuls invariants symplectiques sont donc

<sup>3</sup>Toute variété symplectique est donc orientable et naturellement orientée.

<sup>4</sup>Selon une expression que j'aime beaucoup de Michèle Audin.

des invariants globaux. Par exemple le *nombre de Darboux* d'une variété symplectique est le nombre minimal de cartes nécessaires à un atlas de Darboux, un atlas de Darboux étant évidemment un atlas constitué exclusivement de cartes de Darboux. Mais ce nombre est souvent difficile à calculer explicitement. Un des jeux des géomètres symplecticiens est d'inventer justement de tels invariants globaux et d'essayer ensuite de les calculer (ou de les faire calculer par leurs élèves). C'est exactement cela qui rend la géométrie différentielle symplectique contradictoire : elle est à la fois facile, localement tout se ressemble, et difficile du point de vue de la structure globale.

Il n'est pas possible d'achever ce paragraphe sans évoquer à la fois le groupe des *symplectomorphismes* d'une variété symplectique  $(X, \omega)$  et la notion de gradient symplectique d'une fonction réelle  $f$ , notions qui sont à la base de la géométrie différentielle symplectique.

**DÉFINITION.** *On appelle symplectomorphisme d'une variété symplectique  $(X, \omega)$  tout difféomorphisme  $\phi$  de  $X$  tel que  $\phi^*\omega = \omega$ . L'ensemble des symplectomorphismes de  $X$  constitue le groupe des symplectomorphismes de  $X$ , qui est noté  $\text{Symp}(X)$ .*

La forme symplectique  $\omega$  au point  $x$  réalise un isomorphisme entre  $T_x X$  et  $T_x^* X$ . Son inverse, noté  $\omega_x^{-1}$ , est l'isomorphisme qui permet la définition du *gradient symplectique*.

**DÉFINITION.** *Soit  $f$  une fonction réelle de  $X$ , sa dérivée extérieure  $df$  et, en chaque point  $x$ , un covecteur :  $df(x) \in T_x^* X$ . On appelle gradient symplectique de  $f$  le champ de vecteurs sur  $X$  défini en tout point  $x$  par :*

$$\text{grad}_\omega(f)(x) = \omega_x^{-1}(df(x)).$$

Il n'est pas dans les objectifs de cet ouvrage d'énoncer toutes les vertus et propriétés du gradient symplectique, mais celle-ci est importante :

---

PROPOSITION. Soit  $(X, \omega)$  une variété symplectique compacte et  $f$  une fonction différentiable réelle sur  $X$ , alors son gradient symplectique est complet et son exponentielle  $\exp(s \operatorname{grad}_\omega f)$  est un symplectomorphisme pour tout  $s$ .

Cette proposition (dont on laisse la démonstration au lecteur) montre clairement que le groupe des symplectomorphismes est de dimension infinie.

$$\dim \operatorname{Symp}(X, \omega) = \infty.$$

REMARQUE. Il est tentant de vouloir comparer la géométrie symplectique et la géométrie riemannienne. Mais cette comparaison est déraisonnable. Même si l'on est tenté de jouer sur l'aspect contradictoire *tenseur symétrique* versus *tenseur anti-symétrique*, cette comparaison rejette dans l'ombre un aspect essentiel de la structure symplectique qui n'a pas son équivalent côté riemannien : la *fermeture* de la forme symplectique  $\omega$ , *i.e.*  $d\omega = 0$ . Cette propriété rapproche davantage la géométrie symplectique de la géométrie des *1-tenseurs* que de la géométrie riemannienne. Bien que cela n'en était pas l'objectif principal, c'est ce que j'ai montré, plus ou moins, tout au long de ce livre. Il est donc inutile d'insister sur l'aspect  $\dim \operatorname{Symp}(X, \omega) = \infty$  versus  $\dim \operatorname{Isom}(X, g) < \infty$ , où  $g$  est une métrique riemannienne sur  $X$ .

---

# Espaces différentiables

---

Nous utilisons dans cette annexe une terminologie empruntée à la théorie des *espaces différentiables* ou *espaces difféologiques*, dont on trouvera deux différentes variantes (équivalentes) dans [Che] et [Sou4], et à la théorie des fibrées de cette catégorie d'espaces développée dans [Ig12]. Nous en rappelons quelques éléments.

Un paramétrage d'un ensemble  $X$  est une application  $P$  d'un ouvert  $U$  d'un espace numérique quelconque  $\mathbf{R}^n$  dans  $X$ . Un *espace différentiable* est un ensemble  $X$  pour lequel on a choisi un ensemble de paramétrages dits différentiables, qui sont appelés *plaques*, et qui vérifient les propriétés suivantes :

1. Les paramétrages constants sont des plaques ;
2. le plus petit prolongement commun d'une famille compatible de plaques est encore une plaque ;
3. le composé d'une plaque par un paramétrage  $C^\infty$  de sa source est encore une plaque.

Les variétés sont évidemment des espaces différentiables. Mais les quotients de variétés sont aussi des espaces différentiables : une plaque du quotient est un paramétrage qui admet un relevé différentiable local au voisinage de tout point. C'est muni de cette structure quotient que le tore des périodes  $T_\omega$  est considéré.

Une application  $F$  d'un espace différentiable  $X$  dans un autre  $Y$  est dite *différentiable* si le composé d'une plaque de  $X$  par  $F$  est une plaque de  $Y$ . Les difféomorphismes de  $X$  à  $Y$  sont évidemment les applications différentiables bijectives dont les inverses sont aussi différentiables.

L'espace  $C^\infty(X, Y)$  des applications différentiables de  $X$  dans  $Y$  est muni naturellement d'une structure différentiable appelée *structure fonctionnelle*. Les plaques de cette structure sont les paramétrages  $r \mapsto f$  tels que l'application  $(r, x) \mapsto f(x)$  soit différentiable. C'est de cette structure différentiable dont est muni l'espace  $\text{Arc}(X)$ .

On sait définir sur ces espaces la notion de fibrés différentiables et la théorie de l'homotopie qui l'accompagne [Igl2]. La relation d'homotopie est définie à partir des familles à un paramètre d'arcs différentiables, indépendamment du choix d'une topologie *a priori*<sup>1</sup>.

- Une projection d'espaces différentiables  $\pi : Y \rightarrow X$  est une *fibration* si, pour toute plaque  $P : U \rightarrow X$ , l'image réciproque  $P^*(Y) \rightarrow U$  de la projection  $P$  est localement triviale.

On montre que tout espace différentiable  $X$  possède un revêtement universel, simplement connexe, unique à équivalence près, de groupe structural  $\pi_1(X)$ . Par exemple, la projection  $\text{cl}_\omega : \mathbf{R} \rightarrow T_\omega$  réalise le revêtement universel de  $T_\omega$ .

La notion de groupe différentiable est immédiate : un groupe différentiable est un groupe  $G$  muni d'une structure différentiable compatible avec la multiplication et l'inversion. L'*algèbre de Lie*  $\mathfrak{G}$  de  $G$  est défini comme l'espace des homomorphismes différentiables de  $\mathbf{R}$  dans  $G$  :  $\mathfrak{G} = \text{Hom}^\infty(\mathbf{R}, G)$ . Dans le cas du tore  $T_\omega$ , les homomorphismes différentiables de  $\mathbf{R}$  dans  $G$  se relèvent au revêtement universel  $\text{cl}_\omega : \mathbf{R} \rightarrow T_\omega$  (théorème de monodromie des espaces différentiables [Igl2]) en des homomorphismes de  $\mathbf{R}$  dans  $\mathbf{R}$  et donc  $\mathfrak{T}_\omega \simeq \mathbf{R}$ .

Une  $p$ -forme différentielle  $\omega$  sur un espace différentiable  $X$  est un procédé, qui à toute plaque  $P$  de  $X$  associe une  $p$ -forme notée  $\omega(P)$  ou encore  $P^*\omega$ , définie

<sup>1</sup>Il existe toutefois, sur les espaces différentiables, une topologie remarquable (appelée *D-topologie*), la plus fine qui rende les plaques continues.

sur le domaine  $U$  de  $P$ , telle que pour tout paramétrage  $C^\infty$ ,  $F$  de  $U$  :

$$(P \circ F)^*\omega = F^*(P^*\omega).$$

La dérivée extérieure de  $\omega$  est alors définie par

$$P^*[d\omega] = d[P^*\omega].$$

Cette définition donne lieu à un calcul différentiel extérieur sur  $X$  qui coïncide avec le calcul extérieur ordinaire lorsque  $X$  est une variété. L'image réciproque  $f^*\omega$  d'une  $p$ -forme  $\omega$  sur  $X$  par une application différentiable  $f : Y \rightarrow X$  est définie par :  $P^*[f^*\omega] = [f \circ P]^*\omega$ , où  $P$  est une plaque de  $Y$ .

**PROPOSITION. (QUOTIENT DE FORME)** *Soit  $X$  un espace différentiable et  $\pi : X \rightarrow Y$  une surjection,  $Y$  étant muni de sa structure quotient. Pour qu'une forme différentielle  $\alpha$  définie sur  $X$  passe au quotient, c'est-à-dire pour qu'il existe une forme  $\beta$  sur  $Y$  telle que  $\alpha = \pi^*\beta$ , il faut et il suffit que pour tout couple de plaques  $(P, P')$  de  $X$  telles que  $\pi \circ P = \pi \circ P'$  on ait :  $\alpha(P) = \alpha(P')$ .*

**DÉMONSTRATION.** *Par définition de la structure quotient, toute plaque  $Q$  de  $Y$  s'écrit :  $Q = \sup \pi \circ P_i$ , où les  $P_i$  sont des plaques de  $X$  et  $\sup$  désigne le plus petit prolongement commun. La forme  $\beta$  est définie par :  $\beta(Q) = \sup \alpha(P_i)$ . En effet, soit  $U_i$  le domaine de définition de  $P_i$  ; restreintes à  $U_{ij} = U_i \cap U_j$ , les plaques  $P_i$  et  $P_j$  vérifient  $\pi \circ P_i = \pi \circ P_j$ , et donc :  $\alpha(P_i | U_{ij}) = \alpha(P_j | U_{ij})$ . Les formes  $\alpha(P_i)$  sont compatibles, et ont donc un plus petit prolongement commun  $\sup \alpha(P_i)$ . Par un raisonnement analogue, on montre que  $\beta(Q)$  ne dépend pas du choix des relevés  $P_i$  de  $Q$ . Les autres vérifications sont immédiates. ■*

Il existe sur les espaces différentiables une formule de Cartan pour la dérivée de Lie des formes différentielles. Soit  $h$  un groupe à un paramètre de difféomorphismes de  $X$  :  $h \in \text{Hom}^\infty(\mathbf{R}, \text{Diff}(X))$ , la dérivée de Lie d'une  $p$ -forme  $\omega$  par  $h$  est définie par :

$$\mathcal{L}_h \omega = \frac{\partial}{\partial t} \left\{ h(t)^*\omega \right\}_{t=0}. \quad (\text{D.1})$$

Le contracté  $i_h(\omega)$  de la  $p$ -forme  $\omega$  par le groupe à un paramètre  $h$  est donné par la construction suivante. Soit  $P : U \rightarrow X$  une plaque de  $X$  et  $h \cdot P$  la plaque définie sur  $\mathbf{R} \times U$  par :

$$h \cdot P : (t, r) \rightarrow h(t) \circ P(r) = h(t)(P(r)). \quad (\text{D.2})$$

La  $(p-1)$ -forme  $i_h(\omega)$ , évaluée sur la plaque  $P$ , est donnée par la formule :

$$P^*[i_h(\omega)] = i_{\partial/\partial t}[(h \cdot P)^*\omega]_{\{0\} \times U} \quad (\text{D.3})$$

Cette formule étend la construction ordinaire sur les variétés, et préserve la formule de Cartan<sup>2</sup>.

THÉORÈME. *Soit  $X$  un espace différentiable,  $\omega$  une  $p$ -forme sur  $X$  et  $h$  un groupe à un paramètre de difféomorphismes de  $X$ . On a l'identité :*

$$\mathcal{L}_h \omega = d[i_h(\omega)] + i_h[d\omega]. \quad (\text{D.4})$$

DÉMONSTRATION. *Soit  $P$  une plaque de  $X$  définie sur un domaine  $U$ , par définition :*

$$[\mathcal{L}_h \omega](P) = \frac{\partial}{\partial t} \left\{ [h(t)^*\omega](P) \right\}_{t=0} = \frac{\partial}{\partial t} \left\{ [h(t) \circ P]^*\omega \right\}_{t=0}.$$

Soit  $j : U \mapsto \{0\} \times U$  définie par  $j(r) = (0, r)$ , on a :

$$[i_h(\omega)](P) = i_{\partial/\partial t}[\omega(h \cdot P)]_{\{0\} \times U},$$

où  $h \cdot P$  est définie par (D.2), c'est-à-dire  $[i_h(\omega)](P) = j^*[i_{\partial/\partial t}(\omega(h \cdot P))]$ . On peut donc écrire :

$$\begin{aligned} [d(i_h\omega)](P) + [i_h(d\omega)](P) &= d\{j^*(i_{\partial/\partial t}[\omega(h \cdot P)])\} + j^*\{i_{\partial/\partial t}(d[\omega(h \cdot P)])\} \\ &= j^*\{d \circ i_{\partial/\partial t}[\omega(h \cdot P)] + i_{\partial/\partial t} \circ d[\omega(h \cdot P)]\} \\ &= j^*\{\mathcal{L}_{\partial/\partial t}[\omega(h \cdot P)]\} \\ &= j^*\left\{\frac{\partial}{\partial t}[h(t)^*\omega(h \cdot P)]_{t=0}\right\} \\ &= \frac{\partial}{\partial t}\left\{[h(t) \circ h \cdot P \circ j]^*\omega\right\}_{t=0}. \end{aligned}$$

<sup>2</sup>Nous verrons une autre démonstration de cette formule de Cartan dans l'annexe F.

Or  $h(t) \circ h \cdot P \circ j(r) = h(t)(P(r))$ , c'est-à-dire  $h(t) \circ h \cdot P \circ j = h(t) \circ P$  ; on en déduit donc :

$$\begin{aligned} [d(i_h \omega)](P) + [i_h(d\omega)](P) &= \frac{\partial}{\partial t} \left\{ [h(t) \circ P]^* \omega \right\}_{t=0} \\ &= \frac{\partial}{\partial t} \left\{ h(t)^* [\omega(P)] \right\}_{t=0}, \end{aligned}$$

c'est-à-dire

$$[d(i_h \omega)] + [i_h(d\omega)] = \frac{\partial}{\partial t} \left\{ h(t)^* \omega \right\}_{t=0},$$

c'est ce qu'il fallait démontrer. ■

REMARQUE. Dans le cas d'une 1-forme  $\alpha$ , le contracté avec un homomorphisme  $h$  est une fonction réelle. Pour en donner une expression simple, considérons l'application orbite  $\hat{x}_h$  :

$$\begin{aligned} \hat{x}_h &: \mathbf{R} \longrightarrow X \\ t &\longmapsto h(t)(x) \end{aligned}$$

On vérifie alors que :

$$i_h(\alpha)(x) = \hat{x}_h^*(\alpha)_{t=0}(1). \quad (\text{D.5})$$

L'application  $\hat{x}_h$  étant définie de  $\mathbf{R}$  dans  $X$ , l'image réciproque  $\hat{x}_h^*(\alpha)$  est une 1-forme sur  $\mathbf{R}$  ; calculée au point  $t = 0$  et appliquée au vecteur 1, elle donne la valeur de  $i_h(\alpha)$  au point  $x$ . La vérification de cette formule est une application directe de la formule (D.2) à ce cas particulier.

---

# Opérateur de chaîne-homotopie

---

Soit  $X$  une variété différentiable connexe, on désigne par  $\text{Arc}(X)$  son espace des arcs, c'est-à-dire l'espace des applications différentiables de  $\mathbf{R}$  à valeurs dans  $X$  :

$$\text{Arc}(X) = C^\infty(\mathbf{R}, X). \quad (\text{E.1})$$

Les applications *source* et *but*, notées  $s$  et  $t$ , associent respectivement à tout arc  $\gamma$  les points  $\gamma(0)$  et  $\gamma(1)$ . Le choix de  $\mathbf{R}$  plutôt que l'intervalle  $[0, 1]$  est technique, pour assurer la différentiabilité de la juxtaposition des arcs et l'action du groupe des translations.

L'espace des arcs de  $X$  est muni de sa structure fonctionnelle d'espace différentiable (voir annexe D).

À toute  $p$ -forme  $\omega$  de  $X$ , on associe la  $p$ -forme sur  $\text{Arc}(X)$  définie par l'intégration de  $\omega$  le long des arcs, ce que nous noterons :

$$\tilde{\omega}_\gamma = \int_\gamma \omega. \quad (\text{E.2})$$

L'expression de  $\tilde{\omega}$  dans toute plaque (paramétrage différentiable)  $P$  de l'espace  $\text{Arc}(X)$  est donnée par :

$$P^* \tilde{\omega} = \int_0^1 P_t^* \omega \, dt, \quad (\text{E.3})$$

où  $P_t$  est la plaque de  $X$  définie par  $P_t(r) = P(r)(t)$ . L'application linéaire  $\Phi : \omega \mapsto \tilde{\omega}$  ainsi définie, est un morphisme du complexe différentiel de De Rham de  $X$

dans celui de  $\text{Arc}(X)$  :

$$\Phi : \Omega^*(X) \rightarrow \Omega^*(\text{Arc}(X)) \quad \text{et} \quad d \circ \Phi = \Phi \circ d. \quad (\text{E.4})$$

Les translations de  $\mathbf{R}$  agissent comme un groupe à un paramètre  $\eta$  de difféomorphismes de  $\text{Arc}(X)$ , pour tout  $\gamma \in \text{Arc}(X)$  et tout  $a \in \mathbf{R}$  :

$$\eta(a) : \gamma \mapsto \gamma \circ T_a \quad \text{où} \quad T_a : t \mapsto t + a. \quad (\text{E.5})$$

PROPOSITION. *La dérivée de Lie de la p-forme  $\tilde{\omega}$  par  $\eta$  vérifie :*

$$\mathcal{L}_\eta \tilde{\omega} = \text{but}^* \omega - \text{source}^* \omega. \quad (\text{E.6})$$

DÉMONSTRATION. *En effet, pour toute plaque  $P$  de  $\text{Arc}(X)$  :*

$$[\mathcal{L}_\eta \tilde{\omega}](P) = \frac{\partial}{\partial t} \left\{ [\eta(t)^* \tilde{\omega}](P) \right\}_{t=0} = \frac{\partial}{\partial t} \left\{ (\eta(t)^* \circ P)^* \tilde{\omega} \right\}_{t=0}.$$

Mais  $\eta(t) \circ P : r \mapsto P(r) \circ T_t$  et donc :

$$(\eta(t) \circ P)^* \tilde{\omega} = \int_0^1 (\eta(t) \circ P)_s^* \omega \, ds = \int_0^1 [r \mapsto P(r)(t+s)]^* \omega \, ds.$$

En posant  $u = t + s$ , on obtient

$$(\eta(t) \circ P)^* \tilde{\omega} = \int_t^{1+t} [r \mapsto P(r)(u)]^* \omega \, du,$$

ce qui donne en dérivant :

$$[\mathcal{L}_\eta \tilde{\omega}](P) = [r \mapsto P(r)(1)]^* \omega - [r \mapsto P(r)(0)]^* \omega = [\text{but}^* \omega - \text{source}^* \omega](P).$$

C'est ce qu'il fallait démontrer. ■

En appliquant alors la formule de Cartan (D.4), on a :

$$d[i_\eta \Phi(\omega)] + i_\eta \circ d[\Phi(\omega)] = \text{but}^* \omega - \text{source}^* \omega.$$

En posant alors :

$$K = i_\eta \circ \Phi, \quad K : \Omega^*(X) \rightarrow \Omega^{*-1}(\text{Arc}(X, x_o)), \quad (\text{E.7})$$

on obtient l'identité :

$$K \circ d + d \circ K = \text{but}^* - \text{source}^*. \quad (\text{E.8})$$

DÉFINITION. *Le morphisme  $K$  ci-dessus est appelé opérateur de chaîne-homotopie.*

Lorsque  $X$  est une variété, on peut donner une expression de l'opérateur  $K$  en termes d'intégrales de chemins. Soit  $P$  une plaque de  $\text{Arc}(X)$  définie sur un domaine  $U$ , soit  $r \in U$  et  $\delta r$  un vecteur tangent en  $r$ . Notons

$$\gamma = P(r) \quad \text{et} \quad \delta\gamma : t \mapsto D[\gamma(t)](r)(\delta r),$$

alors  $K\omega$ , calculé au point  $\gamma$  et appliqué à  $p - 1$  variations  $\gamma_i$  associées à des vecteurs  $\delta r_i$  est donné par :

$$K\omega_\gamma(\delta\gamma_1, \dots, \delta\gamma_{p-1}) = \int_0^1 \omega_{\gamma(t)}(\dot{\gamma}(t), \delta\gamma_1, \dots, \delta\gamma_{p-1}) dt. \quad (\text{E.9})$$

---

# Variation d'intégrales multiples dans les espaces différentiables

---

Nous allons montrer, dans cette annexe, comment la variation d'une famille d'intégrales, définie sur un espace différentiable, permet de généraliser à ces espaces la formule de Stokes et donne une nouvelle méthode de démonstration de la formule de Cartan. Nous choisirons pour cela d'utiliser l'homologie cubique, plus pratique, même s'il n'est pas difficile de réécrire ce paragraphe en utilisant l'homologie simpliciale ordinaire, adaptée naturellement aux espaces différentiables.

Soit  $X$  un espace différentiable, nous appellerons  $p$ -cube de  $X$  toute application différentiable  $c$  définie sur un voisinage  $U$  de  $I^p$  à valeurs dans  $X$ , avec  $I = [0, 1]$  (c'est en particulier une  $p$ -plaque de  $X$ ). L'homologie cubique est définie de façon ordinaire, en considérant les différents groupes abéliens libres  $C_p(X)$  engendrés par les  $p$ -cubes,  $p \in \mathbf{N}$ . Les éléments de  $C_p(X)$  sont les  $p$ -chaînes cubiques de  $X$ .

Soit  $\epsilon \in \{0, 1\}$ , on notera  $j_i^\epsilon$  l'application de  $\mathbf{R}^{p-1}$  dans  $\mathbf{R}^p$  définie par :

$$j_i^\epsilon : (t_1, \dots, t_{p-1}) \rightarrow (t_1, \dots, t_{i-1}, \epsilon, t_{i+1}, \dots, t_{p-1}). \quad (\text{F.1})$$

À chaque indice  $i = 1, \dots, p$  correspond les deux faces du cube  $c$  :

$$c_i^\epsilon = c \circ j_i^\epsilon, \quad i = 0, 1, \quad (\text{F.2})$$

ce qui permet de définir le bord  $\partial c$  d'un  $p$ -cube  $c$  par :

$$\partial c = \sum_{\epsilon=0}^1 (-1)^\epsilon \sum_{i=1}^p (-1)^i c_i^\epsilon. \quad (\text{F.3})$$

L'opérateur  $\partial$  est ensuite prolongé par linéarité sur  $C_p(X)$  tout entier ; il est à valeurs dans  $C_{p-1}$  ( $p \geq 1$ ).

L'intégrale d'une  $p$ -forme  $\omega \in \Omega^p(X)$  sur une  $p$ -chaîne  $\sigma \in C_p(X)$  est obtenue en prolongeant par linéarité l'intégrale de  $\omega$  sur les  $p$ -cubes :

$$\sigma = \sum_{c \in C_p(X)} n_{c_p} c \Rightarrow \int_\sigma \omega = \sum_{c \in C_p(X)} n_{c_p} \int_c \omega, \quad (\text{F.4})$$

où les entiers  $n_c$  sont tous nuls sauf un nombre fini d'entre eux, avec :

$$\int_c \omega = \int_{I^p} c^* \omega. \quad (\text{F.5})$$

Considérons maintenant une famille différentiable de  $p$ -cubes  $c_\alpha$ , c'est-à-dire une application différentiable définie sur  $] - \epsilon, \epsilon[ \times I^p$  à valeurs dans  $X$ . L'application  $\alpha \mapsto \int_{c_\alpha} \omega_\alpha$  est alors différentiable, et nous noterons :

$$\delta \int_c \omega = \frac{\partial}{\partial \alpha} \left\{ \int_{c_\alpha} \omega \right\}_{\alpha=0} \quad (\text{F.6})$$

la variation première de l'intégrale de  $\omega$  sur  $c$  associée à la famille  $c_\alpha$ .

Soit  $\bar{c} : ] - \epsilon, \epsilon[ \times U \rightarrow X$  et  $\bar{\omega}$  les  $(p+1)$ -plaques et  $p$ -formes définies respectivement par :

$$\bar{c} = (\alpha, t) \mapsto c_\alpha(t) \quad \text{et} \quad \bar{\omega} = \omega(\bar{c}). \quad (\text{F.7})$$

Soit  $t_0 = \alpha$  et  $e_0$  le vecteur de base  $\partial/\partial \alpha = (1, 0, \dots, 0)$  de  $\mathbf{R} \times \mathbf{R}^p$ , nous introduirons pour la commodité des expressions finales les quantités suivantes :

$$\int_c d\omega(\delta c) = \int_{\{0\} \times I^p} [d\bar{\omega}](e_0) \quad (\text{F.8})$$

et

$$\int_{\partial c} \omega(\delta c) = \sum_{\epsilon=0}^1 (-1)^\epsilon \sum_{i=0}^p \int_{I^{p-1}} j_i^{\epsilon*} \bar{\omega}(e_0), \quad (\text{F.9})$$

où la notation  $\bar{\omega}(e_0)$  désigne le contracté de  $\bar{\omega}$  par le vecteur  $e_0$  (*idem* pour  $d\bar{\omega}$ ). Il serait possible de définir formellement  $\delta c$ , mais nous dirons simplement que cela représente la variation infinitésimale du cube  $c$  associée à la famille  $c_\alpha$ .

Le lecteur peut vérifier que, compte tenu de ces définitions, la démonstration de la proposition suivante se réduit à un calcul de géométrie différentielle ordinaire sur  $\mathbf{R}^p$ .

PROPOSITION. *Soit  $X$  un espace différentiable,  $\omega$  une  $p$ -forme sur  $X$  et  $c$  un  $p$ -cube de  $X$ . Soit  $c_\alpha$  une famille différentiable de  $p$ -cubes telle que  $c_0 = c$ . La variation première de l'intégrale de  $\omega$  sur  $c$  associée à la famille  $c_\alpha$  est donnée par la formule :*

$$\delta \int_c \omega = \int_c d\omega(\delta c) + \int_{\partial c} \omega(\delta c). \quad (\text{F.10})$$

On obtient en particulier comme corollaire la formule de Stokes étendue aux espaces différentiables.

COROLLAIRE. *Soit  $\omega$  une  $(p-1)$ -forme définie sur un espace différentiable  $X$ , et  $\sigma$  une  $p$ -chaîne cubique de  $X$ , on a :*

$$\int_\sigma d\omega = \int_{\partial\sigma} \omega. \quad (\text{F.11})$$

DÉMONSTRATION. *En effet, il suffit de le démontrer pour un  $p$ -cube  $c$ . On a  $\delta \int_c d\omega = \int_{\partial c} d\omega(\delta c)$  et aussi  $\delta \int_{\partial c} \omega = \int_{\partial c} d\omega(\delta c)$ , car  $dd\omega = 0$  et  $\partial\partial c = 0$ . On en déduit que  $\delta[\int_c d\omega - \int_{\partial c} \omega] = 0$ , et donc que  $\int_c d\omega - \int_{\partial c} \omega$  est constante sur l'espace des  $p$ -cubes. Comme l'espace des  $p$ -cubes est contractile, il suffit d'évaluer cette différence sur le  $p$ -cube constant, ce qui donne le résultat. ■*

On peut étendre évidemment cette proposition par linéarité à toute  $p$ -chaîne  $\sigma$ , limite de  $p$ -chaînes cubiques, pour laquelle l'intégrale converge. On peut l'étendre aussi sans difficulté à des domaines polyédraux en utilisant les  $p$ -chaînes simpliciales plutôt que cubiques.

La formule (F.10) ci-dessus peut se généraliser immédiatement au cas où la forme  $\omega$  est aussi variable. Soit  $\omega_\alpha$  une famille différentiable de  $p$ -formes telle

que  $\omega_0 = \omega$ . On dira que la famille de  $p$ -formes  $\omega_\alpha$  est différentiable si pour toute plaque  $P$  de  $X$  l'application  $(\alpha, r) \mapsto [\omega_\alpha(P)]_r$  est différentiable, autrement dit, si pour toute famille de  $p$ -vecteurs  $(u_1, \dots, u_p)$  dans  $\mathbf{R}^p$  l'application  $(\alpha, r) \mapsto [\omega_\alpha(P)]_r(u_1, \dots, u_p)$  est différentiable. En notant  $\delta\omega$  la  $p$ -forme définie par :

$$\delta\omega : P \mapsto \frac{\partial}{\partial\alpha} \{\omega_\alpha(P)\}_{\alpha=0}, \quad (\text{F.12})$$

on peut écrire :

$$\delta \int_c \omega = \int_c d\omega(\delta c) + \int_{\partial c} \omega(\delta c) + \int_c \delta\omega. \quad (\text{F.13})$$

La formule de Stokes permet de montrer en particulier, comme dans le cas des variétés, l'invariance par homotopie de la cohomologie de De Rham. Elle permet aussi, sous sa forme étendue (F.13), de donner une autre démonstration de la formule de Cartan (annexe D).

**DÉMONSTRATION.** (FORMULE DE CARTAN) *Soit  $h$  un groupe à un paramètre de difféomorphismes de  $X$ ,  $\omega$  une  $p$ -forme sur  $X$  et  $c$  un  $p$ -cube. Notons  $c_t = h(t) \circ c$  et  $\omega_t = h(t)_*\omega = h(-t)^*\omega$ . En appliquant (F.13) à ces familles à un paramètre et en notant que  $\int_{c_t} \omega_t = \int_c \omega$ , on obtient :*

$$\int_c \mathcal{L}_h \omega = \int_c i_h[d\omega] + \int_c d[i_h\omega]. \quad (\text{F.14})$$

*Cette égalité étant vérifiée pour tout  $p$ -cube  $c$  de  $X$ , on en déduit la formule de Cartan :  $\mathcal{L}_h \omega = i_h[d\omega] + d[i_h\omega]$ . ■*

# Index

- éléments kepleriens, 11, 26
- énergie, 20, 35
- éphémérides, 9
- époque, 11, 25
- équation aux variations, 49
- équation de Kepler, 24
- équation de Newton, 19
  
- action hamiltonienne, 96
- action lagrangienne, 49
- anomalie excentrique, 24, 35
- anomalie vraie, 22
- antisymétrique, 30
- application but, 122, 185
- application de Legendre, 52
- application moment, 90, 101, 102
- application orbite, 121
- application source, 122, 185
- arcs, 185
- arcs homologues, 123
  
- base canonique, 168
  
- caractéristiques, 90
- co-isotrope, 167
- cobord, 94
- cocycle, 93, 120
  
- cocycle de Bott-Thurston, 140
- cocycle de Gelfand-Fuchs, 142
- composantes contravariantes, 38
- composantes covariantes, 38
- conditions initiales, 51
- constantes d'intégration, 10
- contrainte (compatibilité), 46
- courbes extrémales, 49
- crochets de Lagrange, 32, 34
- crochets de Poisson, 32
  
- demi-droites, 56
- difféomorphisme hamiltonien, 129
- disque de Poincaré, 142
- défaut d'invariance, 93
- déplacements euclidiens, 82
  
- ellipse, 11
- ellipse osculatrice, 29
- espace cotangent, 52
- espace d'évolution, 51, 59
- espace des arcs, 122
- espace des configurations, 53
- espace des mouvements, 12
- espace difféologique, 180
- espace différentiable, 122, 180

- espace dual, 52  
espace tangent, 52  
espaces des directions, 56  
excentricité, 22  
extension centrale, 119
- fibré d'homologie, 115  
force, 16  
force centripète, 17  
force d'inertie, 47  
force perturbatrice, 27, 31  
forces vives, 21, 26, 160  
forme bilinéaire, 165  
forme canonique, 39, 169  
forme de Cartan, 54  
forme différentielle, 38  
forme exacte, 109  
forme symplectique, 14, 38  
formule de Cartan, 183, 191  
foyer, 11, 16, 22, 24, 25
- gradient symplectique, 41, 71, 178  
groupe de Galilée, 96  
groupe de Heisenberg, 138  
groupe de Lie, 88  
groupe des périodes, 121  
groupe difféologique, 122  
groupe différentiable, 122  
groupe symplectique, 14  
géodésiques de la sphère, 73  
géodésiques du disque  $H^2$ , 76
- hamiltonien, 62, 67  
holonome, 145  
homogénéisation (du lagrangien), 52  
homomorphisme de Calabi, 97
- immersion du cercle, 141  
inertie, 16  
inégalités (des mouvements), 37  
isotrope, 167
- lagrangien, 48, 50  
loi des aires, 16  
lois de Kepler, 16, 18
- moment, 90, 101, 102  
moment cinétique, 21  
moment des 1-formes, 112  
mouvements kepleriens, 10, 12, 23  
multiplicateur de Lagrange, 63  
mécanique, 16
- nœuds (ligne des), 26
- obstruction (au moment), 103  
opérateur d'homotopie, 185  
orbite, 12  
orbite coadjointe, 149  
orbite elliptique, 23  
orbite hyperbolique, 23  
orbite parabolique, 23  
orthogonal symplectique, 167
- paramétrage différentiable, 122

- parentèses de Lagrange, 13, 30, 31  
perturbation, 26  
plaques, 180  
Poincaré, 42  
point matériel, 45, 61  
potentiel, 29  
potentiel de perturbation, 29  
principe de d'Alembert, 49  
principe des travaux virtuels, 45  
problème à deux corps, 15  
présymplectique, 58  
périhélie, 26
- relevé holonome, 54  
revêtement d'homologie, 116  
réduction symplectique, 167
- section conique, 22  
solide, 81, 86  
sphère, 46  
stabilité (du grand axe), 34  
stabilité séculaire, 13  
symplectique, 13, 60  
symplectomorphisme, 179  
symétries, 88  
système conservatif, 35  
système intégrable, 10  
série de Fourier, 35
- théorème de Crofton, 64  
théorème de Darboux, 39, 173  
théorème de de Rham, 119  
théorème de Lagrange, 31, 36  
théorème de Moser, 173  
théorème de Noether, 90  
tore, 135  
tore des périodes, 121  
trajectoire, 12  
transformation de jauge, 130  
travail (d'une force), 46
- variation, 46, 54  
variation des constantes, 12, 27, 32  
variété des droites, 62  
variété symplectique, 172  
vecteur de Laplace, 22

# Bibliographie

- [Arn1] V. I. Arnold. *Méthodes mathématiques de la mécanique classique*. Éditions MIR, Moscou, 1976.
- [Arn2] V. I. Arnold. *Chapitres supplémentaires à la théorie des équations différentielles*. Éditions MIR, Moscou, 1980.
- [Aud1] M. Audin. *The topology of torus actions on symplectic manifolds*, volume 93. Birkäuser, 1991.
- [Aud2] M. Audin. *Spinning Tops*, volume 51. Cambridge University Press, Cambridge 1996.
- [Bot] R. Bott. « On some formulas for the characteristic classes of group actions, differential topology, foliations and Gelfand-Fuchs cohomology ». In *Proceed. Rio de Janeiro, 1976*, volume 652 of *Springer Lectures Notes*. Springer Verlag, New-York 1978.
- [Che] K. T. Chen. « Iterated path integrals ». *Bull. of Am. Math. Soc.*, 83(5) : 831–879, New-York 1977.
- [Del1] T. Delzant. « Hamiltoniens périodiques et images convexes de l'application moment ». *Bulletin de la SMF*, 116, Paris 1988.
- [Del2] T. Delzant. « Classifications des actions hamiltoniennes complètement intégrables des groupes de rang 2 ». *Annals of Global Analysis and Geometry*, 8(1) : 87–112, 1990.

- [dGa] F. de Gandt. *Force et géométrie*. Thèse de doctorat, Paris I, 1987.
- [GF] I. Gelfand and D. Fuchs. « Cohomology of the lie algebra of vector fields on the circle ». *Functional Analysis and Applications*, 2 : 342–343, 1968.
- [Huy] C. Huygens. *Horologium oscillatorium*. Éditions Bergeret, rue Leyteire, Bordeaux, 1980. Traduit du latin par Jean Peyroux. Édité originellement à Paris, en 1673.
- [Igl1] P. Iglesias. « Classification des  $SO(3)$ -variétés symplectiques de dimension 4 ». Prétirage CPT-84/PE.1673, Centre de physique théorique, Marseille-Luminy, 1984.
- [Igl2] P. Iglesias. *Fibrés difféologiques et homotopie*. Thèse de doctorat d'état, Université de Provence, Marseille, 1985.
- [Igl3] P. Iglesias. « Les  $SO(3)$ -variétés symplectiques et leur classification en dimension 4 ». *Bull. Soc. Math. France*, 119(4) : 371–396, 1991.
- [KM] M. Kargapalov and Y. Merzliakov. *Éléments de la théorie des groupes*. Éditions MIR, Moscou, 1985.
- [Lag1] J.-L. Lagrange. « Sur les intégrales particulières des équations différentielles ». In *Œuvres de Lagrange*, volume IV. Gauthier-Villars, Paris, 1877. Nouveaux Mémoires de l'Académie royale des Sciences et Belles Lettres de Berlin, 1774.
- [Lag2] J.-L. Lagrange. « Recherches sur les suites récurrentes ». In *Œuvres de Lagrange*, volume IV. Gauthier-Villars, Paris, 1877. Nouveaux Mémoires de l'Académie royale des Sciences et Belles Lettres de Berlin, 1775.
- [Lag3] J.-L. Lagrange. « Sur différentes questions d'analyse relatives à la théorie des intégrales particulières ». In *Œuvres de Lagrange*, volume IV. Gauthier-Villars, Paris, 1877. Nouveaux Mémoires de l'Académie royale des Sciences et Belles Lettres de Berlin, 1779.

- [Lag4] J.-L. Lagrange. « Théorie des variations séculaires des éléments des planètes ». In *Œuvres de Lagrange*, volume V, pp. 128-207. Gauthier-Villars, Paris, 1877. Nouveaux Mémoires de l'Académie royale des Sciences et Belles Lettres de Berlin, 1781.
- [Lag5] J.-L. Lagrange. « Sur la théorie des variations des éléments des planètes et en particulier des variations des grands axes de leurs orbites ». In *Œuvres de Lagrange*, volume VI, pp. 713-768. Gauthier-Villars, Paris, 1877. Lu le 22 août 1808 à l'Institut de France.
- [Lag6] J.-L. Lagrange. « Sur la théorie générale de la variation des constantes arbitraires dans tous les problèmes de la mécanique ». In *Œuvres de Lagrange*, volume VI, pp. 771-805. Gauthier-Villars, Paris, 1877. Lu le 13 mars 1809 à l'Institut de France.
- [Lag7] J.-L. Lagrange. « Second mémoire sur la théorie générale de la variation des constantes arbitraires dans tous les problèmes de la mécanique ». In *Œuvres de Lagrange*, volume VI, pages 809-816. Gauthier-Villars, Paris, 1877. Lu le 19 février 1810 à l'Institut de France.
- [Lag8] J.-L. Lagrange. *Mécanique analytique*. Librairie Albert Blanchard, Paris, 1965. *Fac-simile* de la troisième édition. Deuxième édition par Mme Veuve Courcier, Paris 1811.
- [New1] I. Newton. *De motu corporum in gyrum*. Manuscrit envoyé à Halley, Cambridge, en 1684, original introuvable. ULC Add.3965, fol.55-62v. Traduit dans *De la gravitation, suivi de Du mouvement des corps* par M.-F. Biarnais et F. de Gandt, Gallimard, Paris 1995.
- [New2] I. Newton. *Les principes Mathématiques de la Philosophie Naturelle*. Traduction de la Marquise du Chastelet. Librairie Albert Blanchard, Paris, 1966. *Fac-simile* d'une édition de 1756.

- [New3] I. Newton. *De philosophiae naturalis principia mathematica*. Traduction nouvelle par Marie-Françoise Biarnais. Christian Bourgeois Editeur, Paris 1985.
- [McL] S. McLane. *Homology*. Springer Verlag, New-York Heidelberg Berlin, 1967.
- [Poi1] S. D. Poisson. « Sur les inégalités séculaires des moyens mouvements des planètes ». *Journal de l'école polytechnique*, 15 : 1–56, 1808. Lu à l'Institut le 20 juin 1808.
- [Poi2] S. D. Poisson. « Sur la variation des constantes arbitraires dans les questions de mécanique ». *Journal de l'école polytechnique*, 15 : 266–344, 1809. Lu à l'Institut le 16 octobre 1809.
- [Poin] H. Poincaré. *Les nouvelles méthodes de la mécanique céleste*. Gauthier-Villars, Paris, 1892.
- [Sou1] J.-M. Souriau. « Géométrie symplectique différentielle. applications ». In *Coll. Int. CNRS*, page 53. CNRS, Strasbourg, 1953.
- [Sou2] J.-M. Souriau. *Structure des systèmes dynamiques*. Dunod, Paris, 1970.
- [Sou3] J.-M. Souriau. « Géométrie globale du problème à deux corps ». *Modern Developments in Analytical Mechanics*. Accademia della Scienza di Torino, 1983, pp. 369–418. Supplemento al vol. 117, Atti della Accademia della Scienza di Torino.
- [Sou4] J.-M. Souriau. « Groupes différentiels et physique mathématique ». In *Feuillets et quantification géométrique, Séminaire Sud-Rhodanien de géométrie II*, P. Dazord et N. Desolneux-Moulis (Dir.) *Collection travaux en cours* Hermann, pages 73–119, Paris, 1984.
- [Sou5] J.-M. Souriau. « La structure symplectique de la mécanique décrite par Lagrange en 1811 ». *Math. Sci. hum.*, (94) : 45–54, 1986.

[Ste] S. Sternberg. *Celestial Mechanics*. W. A. Benjamin Inc., New-York, 1969.

[Wey] H. Weyl. *Classical Groups*. Princeton University Press, 1939–1946.

# Table des matières

<b>I</b>	<b>Introduction</b>	<b>5</b>
<b>II</b>	<b>Lagrange et les origines du calcul symplectique</b>	<b>9</b>
II.1	Espace des mouvements d'une planète . . . . .	15
II.2	La méthode de la variation des constantes . . . . .	26
II.3	Application à la stabilité séculaire du grand axe . . . . .	34
II.4	Structure de l'espace des mouvements keplériens . . . . .	37
II.5	Une remarque de Poincaré . . . . .	42
<b>III</b>	<b>Les principes de la mécanique analytique</b>	<b>45</b>
III.1	Le principe des travaux virtuels . . . . .	45
III.2	Homogénéisation du lagrangien . . . . .	51
III.3	L'espace des directions tangentes . . . . .	56
III.4	La structure symplectique . . . . .	59
III.5	Le point de vue hamiltonien . . . . .	67
<b>IV</b>	<b>Quelques exemples</b>	<b>73</b>
IV.1	Les géodésiques de la sphère . . . . .	73
IV.2	Les géodésiques du disque de Poincaré . . . . .	76
IV.3	Les mouvements du solide . . . . .	81
<b>V</b>	<b>Invariance du lagrangien et théorème de Noether</b>	<b>88</b>
V.1	Lagrangien invariant par un groupe de symétries . . . . .	88
V.2	Lagrangien « presque » invariant . . . . .	91

---

<b>VI L'application moment</b>	<b>101</b>
VI.1 Conditions d'existence d'une application moment . . . . .	102
VI.2 Définition alternative du moment . . . . .	104
VI.3 Le cas particulier des formes exactes . . . . .	109
VI.4 Classes de cohomologie associées au moment . . . . .	114
VI.5 Le cas général des 2-formes fermées quelconques . . . . .	120
VI.5.1 Intégration d'une 2-forme fermée . . . . .	121
VI.5.2 Les difféomorphismes hamiltoniens revisités . . . . .	129
VI.5.3 Quelques exemples . . . . .	135
<b>ANNEXES</b>	<b>149</b>
<b>A Lagrange, Poisson, quelle histoire !</b>	<b>150</b>
<b>B Histoire d'H</b>	<b>158</b>
<b>C Éléments de géométrie symplectique</b>	<b>165</b>
<b>D Espaces différentiables</b>	<b>180</b>
<b>E Opérateur de chaîne-homotopie</b>	<b>185</b>
<b>F Variation d'intégrales multiples dans les espaces différentiables</b>	<b>188</b>

Patrick Iglesias  
Centre de Mathématiques et Informatique  
39, rue F. Joliot-Curie  
13453 Marseille cedex 13  
`patrick.iglesias@laposte.net`