

Mécanique Analytique

Cyril Falvo & Guillaume Roux
Université Paris Saclay

2024–2025

Table des matières

1	Rappels : principes de la mécanique Newtonienne	9
I	Cinématique Galiléenne	9
II	Axiomes de la mécanique Newtonienne	10
III	Formulation énergétique de la mécanique classique	12
IV	Repères tournants	13
I	Formulation Lagrangienne	17
2	Équations de Lagrange	19
I	Contraintes et principe de d'Alembert	19
1	Principe de d'Alembert au travers d'un exemple	19
2	Généralisation à N particules et classification des contraintes	21
II	Coordonnées généralisées	24
1	Définition et exemples	24
2	Réécriture des déplacements virtuels	26
III	Équations de Lagrange	29
1	Démonstration	29
2	Forces dérivant d'un potentiel - Lagrangien	31
3	Potentiel généralisé et force de Lorentz	31
IV	Covariance : invariance de la forme des équations de Lagrange	33
3	Principe de moindre action	35
I	Calcul des variations	35
1	Position du problème : le brachistochrone	35
2	Fonctionnelle	36
3	Extrémalisation et dérivée fonctionnelle	36
4	Formule d'Euler	38
5	Principe de moindre action de Hamilton	39
II	Prise en compte des contraintes par des multiplicateurs de Lagrange	40
1	Principe : théorème des extréma liés	40
2	Principe de moindre action de Hamilton en présence de contraintes	42

3	Exemple : Le pendule plan	43
C1.	Solution du problème du brachistochrone	44
4	Formulation lagrangienne - Grandeurs physiques	47
I	Moment généralisé	47
1	Définition	47
2	Coordonnées cycliques	48
II	Énergie	49
1	Théorème d'Euler sur les fonctions homogènes	49
2	Définitions	50
3	Dépendance en temps et conservation de l'énergie	50
III	Invariance de Jauge du Lagrangien	51
1	Invariance de jauge du lagrangien	51
2	Transformation de jauge électromagnétique	52
3	Particule dans un référentiel en translation	52
IV	Rôle des symétries dans la formulation lagrangienne	53
1	Forme du lagrangien d'une particule libre	53
2	Théorème de Noether	54
5	Oscillations	57
I	Linéarisation des équations du mouvement	57
1	Exemples et motivations	57
2	Voisinage de l'équilibre	57
3	Équations du mouvement	58
II	Pulsations propres et modes propres	58
1	Systèmes aux valeurs propres généralisé	58
2	Modes propres ou modes normaux	59
III	Exemple : Chaîne linéaire de 3 particules	59
C1.	Problème aux valeurs propres généralisé	62
II	Formulation Hamiltonienne	65
6	Formalisme hamiltonien	67
I	Transformée de Legendre	67
1	Approche intuitive	67
2	Approche plus rigoureuse	68
3	Généralisation à plusieurs variables	69
II	Équations de Hamilton	70
1	Fonction de Hamilton ou hamiltonien	70

2	Équations du mouvement	70
3	Principe de Hamilton	71
III	Exemples (voir TDs)	71
1	Le pendule simple	71
2	Particule chargée dans un champ électromagnétique	71
3	Hamiltonien relativiste	72
4	Coordonnées cylindriques et sphériques	72
7	Dynamique hamiltonienne	73
I	Évolution temporelle et espace des phases	73
1	Système dynamique	73
2	Déterminisme	73
3	Système hamiltonien conservatif et portrait de phase	74
4	Théorèmes généraux	75
II	Crochets de Poisson	77
1	Définition	77
2	Propriétés mathématiques (TDs)	77
3	Évolution temporelle avec les crochets de Poisson	78
4	Liouvillien et opérateur d'évolution	79
5	Intégrales premières et systèmes intégrables	80
8	Transformations canoniques	83
I	Définition et fonctions génératrices	83
1	Définition	83
2	Fonctions génératrices	84
II	Structure symplectique et invariance des crochets de Poisson	85
1	Formulation symplectique de la dynamique hamiltonienne	85
2	Application aux transformations canoniques	86
III	Exemples de transformations canoniques	88
1	Transformations d'échange	88
2	Oscillateur harmonique	88
3	Transformations infinitésimales (voir TD)	89
9	Équation de Hamiton-Jacobi	91
I	La fonction principale de Hamilton	91
II	Équation de Hamilton-Jacobi	92
III	Systèmes conservatifs et séparables	94
IV	Transformation canonique de Jacobi	95
V	Variables angle-action	96
1	Principe	96

2	Systèmes multiples périodiques	97
3	Représentation des coordonnées en fonction des variables angle-action	101
VI	Théorie des perturbations adiabatiques	102

Introduction

La *mécanique newtonienne* est la théorie physique qui traite du mouvement des corps macroscopiques dont la vitesse reste faible par rapport à la vitesse de la lumière. Cette théorie fut introduite par Isaac Newton en 1687 dans son œuvre *Philosophiæ naturalis principia mathematica*.¹ Utilisée conjointement avec la théorie universelle de la gravitation, elle a permis d'unifier deux domaines de la physique jusque là séparé : la chute des corps et le mouvement des corps célestes. Cette théorie a prévalu pendant plus de 200 ans jusqu'au tout début du 20ème siècle où une refondation complète de la physique s'est opérée avec l'émergence de trois nouvelles théories : la *physique statistique*, la *théorie de la relativité*, et la *mécanique quantique*.²

On pourrait croire que pendant ces deux siècles entre l'établissement des trois lois de Newton et la révolution d'Einstein et de Schrödinger, la mécanique classique n'a pas évolué de manière significative. Pourtant, elle s'est profondément transformée sous l'impulsion notamment de Pierre Louis Maupertuis, Leonhard Euler, Joseph-Louis Lagrange, William Rowan Hamilton, et Carl Gustav Jacob Jacobi (Fig. 1) pour aboutir à une formulation analytique par opposition à la formulation vectorielle de Newton. Pour illustrer la force de la formulation analytique on pourra citer les quelques mots de Lagrange dans le préface de son manuscrit "Mécanique Analytique" publié en 1788.

On ne trouvera point de Figures dans cet Ouvrage. Les méthodes que j'y expose ne demandent ni constructions, ni raisonnements géométriques ou mécaniques, mais seulement des opérations algébriques, assujetties à une marche régulière et uniforme. Ceux qui aiment l'Analyse, verront avec plaisir la Mécanique en devenir une nouvelle branche, et me sauront gré d'en avoir étendu ainsi le domaine.

— Joseph-Louis, Lagrange

La *mécanique analytique* est une expression multiforme de la *mécanique classique* : elle peut s'écrire sous les différentes formulation dites *Lagrangienne*, *Hamiltonienne*, ou bien encore avec les formulations de *Hamilton–Jacobi* et *symplectique*. Ce cours abordera l'ensemble de ces formulations avec plus ou moins de détails. La *mécanique analytique* peut être vu comme juste une simplification mathématique de la mécanique newtonienne sans réel apport aux principes fondateurs de la mécanique classique. C'est sans compter sur toute la puissance du formalisme de la mécanique analytique qui permet la résolution de problèmes très complexes et l'utilisation d'approximations permettant des prédictions très précises notamment en mécanique céleste. D'autre part, la formulation analytique de la mécanique a offert aux précurseurs de la physique moderne le cadre théorique adaptée à une refondation des principes de base de la physique. La théorie de la relativité d'Einstein s'exprime naturellement avec la formulation

1. Principes mathématiques de la philosophie naturelle

2. Il existe également une théorie quantique relativiste : la théorie quantique des champs



FIGURE 1. – De gauche à droite, portraits de Pierre-Louis Moreau de Maupertuis (1698–1759), Leonhard Euler (1707–1783), Joseph-Louis Lagrange (1736–1813), William Rowan Hamilton (1805–1865) et Carl Gustav Jacob Jacobi (1804–1851).

Lagrangienne de la mécanique, la physique statistique avec la formulation Hamiltonienne et la mécanique quantique dans sa formulation originelle repose sur les formulations Hamiltonienne et de Hamilton-Jacobi. La *mécanique analytique* est donc un formalisme essentiel pour la compréhension des théories moderne de physique fondamentale.

Ces notes de cours sont largement inspirés d’ouvrages de référence qui peuvent venir compléter cette introduction à la mécanique analytique. Voici une liste non exhaustive d’ouvrage en langue Française et Anglaise. Chacun de ces livres comporte un point de vue différent et complémentaire de la mécanique analytique. Ces notes de cours ont également été fortement influencées par le cours de Mécanique Analytique donné par J.-M. Rax en Licence 3 de Physique Fondamentale.

Bibliographie

- Arnold, V. I. (1989). *Mathematical Methods of Classical Mechanics*. Springer-Verlag, second edition.
- Charlier, A., Bérard, A., and Charlier, M.-F. (1989). *Mécanique Analytique*. Ellipses.
- Golstein, H., Safko, J. L., and Poole, C. P. (2013). *Classical Mechanics*. Pearson, third edition.
- José, J. V. and Saletan, E. J. (2002). *Classical Dynamics, a contemporary approach*. Cambridge University Press.
- Lanczos, C. (1970). *The Variational Principles of Mechanics*. Dover, fourth edition.
- Landau, L. and Lifchitz, E. (1998). *Mécanique*. Physique Théorique. Ellipses.

1. Rappels : principes de la mécanique Newtonienne

I. Cinématique Galiléenne

La cinématique a pour objectif de décrire le mouvement d'un ensemble de point au cours du temps. En cinématique Galiléenne, nous allons supposer que l'espace ambiant s'apparente à un espace vectoriel Euclidien de dimension trois E^3 et que le temps est un espace unidimensionnel pouvant être décrit par l'ensemble des réels \mathbb{R} . Puisque l'espace est de dimension 3, la position dans l'espace de chaque point est donnée par rapport à un repère de manière unique par le biais de trois coordonnées. Pour décrire le temps, on introduit le paramètre t mesuré par un observateur associé au repère d'espace. L'union des repères d'espace et de temps forme alors un *référentiel* associé à cet observateur. En cinématique Galiléenne, on considère que le temps est un paramètre absolu partagé par l'ensemble des observateurs et qu'il s'écoule de manière identique dans l'ensemble des référentiels.

La position du point est déterminé par le vecteur \mathbf{r} qui relie le point au centre du repère. Par exemple, dans un repère Cartésien de centre O les trois composantes du vecteur \mathbf{r} par rapport aux axes Ox , Oy et Oz sont les composantes x , y et z du point

$$\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z, \quad (1.1)$$

où \mathbf{e}_x , \mathbf{e}_y , et \mathbf{e}_z sont les vecteurs unitaires dans les directions Ox , Oy , et Oz . Il est également courant de noter ces trois composantes sous la forme x_i avec $i = 1, 2, 3$ si bien que le vecteur \mathbf{r} s'écrit

$$\mathbf{r} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 = \sum_{i=1}^3 x_i\mathbf{e}_i. \quad (1.2)$$

Chaque point du système évolue au cours du temps. Pour décrire cette évolution, le vecteur \mathbf{r} devient une fonction du temps t

$$\mathbf{r} = \mathbf{r}(t). \quad (1.3)$$

On supposera que le paramètre t évolue de manière monotone et croissante de sorte que le vecteur $\mathbf{r}(t)$ suive les différentes positions de manière chronologiquement ordonnée. La vitesse et l'accélération d'un point par rapport à un référentiel est défini par

$$\mathbf{v} = \dot{\mathbf{r}} = \frac{d\mathbf{r}}{dt}, \quad \mathbf{a} = \dot{\mathbf{v}} = \frac{d^2\mathbf{r}}{dt^2} \quad (1.4)$$

Dans un repère cartésien, la vitesse et l'accélération s'écrivent

$$\mathbf{v} = \sum_i \dot{x}_i(t) \mathbf{e}_i, \quad \mathbf{a} = \sum_i \ddot{x}_i(t) \mathbf{e}_i \quad (1.5)$$

Le repère Cartésien n'est pas l'unique manière de décrire la position et le mouvement d'un point. Il est également possible d'utiliser des repères tournants comme les repères cylindriques et sphériques. Dans ce cas, les vecteurs unitaires deviennent dépendant des coordonnées du point x_i . Les propriétés des repères tournants est rappelé en annexe de ce chapitre. Dans le cas du repère cylindrique $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z)$, la position du système est caractérisée par les trois coordonnées r , θ et z .

$$\mathbf{r} = r\mathbf{e}_r + z\mathbf{e}_z \quad (1.6)$$

La vitesse et l'accélération s'écrivent alors

$$\mathbf{v} = \dot{r}\mathbf{e}_r + r\dot{\theta}\mathbf{e}_\theta + \dot{z}\mathbf{e}_z, \quad (1.7)$$

$$\mathbf{a} = \left(\ddot{r} - r\dot{\theta}^2\right) \mathbf{e}_r + \frac{1}{r} \frac{d}{dt} \left(r^2\dot{\theta}\right) \mathbf{e}_\theta + \ddot{z}\mathbf{e}_z. \quad (1.8)$$

Dans le cas du repère sphérique $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi)$, la position du point est caractérisée par les trois coordonnées r , θ et φ . La vitesse et l'accélération s'écrivent

$$\mathbf{v} = \dot{r}\mathbf{e}_r + r\dot{\theta}\mathbf{e}_\theta + r\dot{\varphi} \sin \theta \mathbf{e}_\varphi, \quad (1.9)$$

$$\begin{aligned} \mathbf{a} = & \left(\ddot{r} - r\dot{\theta}^2 - r\dot{\varphi}^2 \sin^2 \theta\right) \mathbf{e}_r + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta} - r\dot{\varphi}^2 \sin \theta \cos \theta\right) \mathbf{e}_\theta \\ & + \left(r\ddot{\varphi} \sin \theta + 2\dot{r}\dot{\varphi} \sin \theta + 2r\dot{\theta}\dot{\varphi} \cos \theta\right) \mathbf{e}_\varphi \end{aligned} \quad (1.10)$$

Dans le repère de Frenet-Serret $(\mathbf{t}, \mathbf{n}, \mathbf{b})$, la position du point le long d'une courbe gauche est caractérisée par l'abscisse curviligne s . La vitesse et l'accélération sont alors données par

$$\mathbf{v} = \dot{s}\mathbf{t}, \quad (1.11)$$

$$\mathbf{a} = \ddot{s}\mathbf{t} + \gamma\dot{s}^2\mathbf{n}, \quad (1.12)$$

où γ est la courbure de la courbe.

II. Axiomes de la mécanique Newtonienne

La mécanique Newtonienne est en général introduite par les trois lois de Newton. Cependant ces trois lois n'offrent pas de bonne définition du concept de masse et du concept de force. Nous préférons plutôt présenter les principes et axiomes établies à partir d'observations et d'expériences dont dérive les lois de Newton. Nous allons passer en revue ces axiomes dans le cadre du mouvement de points matériels, c'est à dire des particules dont la taille est négligeable. La description de systèmes étendus se fait de naturellement à partir des axiomes établis pour les points matériels.

Principe d'inertie Il existe une classe d'équivalence de référentiels dit Galiléen qui possèdent la propriété suivante : dans un référentiel Galiléen, une particule *isolée* est soit immobile, soit en mouvement rectiligne uniforme. Le passage d'un référentiel Galiléen à un autre se fait par l'intermédiaire d'une transformation dite Galiléenne. Ce principe permet d'une part de définir la notion de référentiel Galiléen, c'est à dire un référentiel dans lequel une particule isolée se déplace en ligne droite et d'autre part la notion de temps associé à ce référentiel qui s'écoule de manière linéaire le long du chemin de cette particule. Ce point soulève déjà des questions sur la définition du temps. Le temps est une notion naturelle et il semble futile de devoir définir la façon dont celui-ci s'écoule. Aujourd'hui le temps est défini de la manière suivante : *La seconde est la durée de 9 192 631 770 périodes de la radiation correspondant à la transition entre les niveaux hyperfins de l'atome de césium 133*. Cette définition correspond à une transition quantique entre niveaux excités d'un atome. D'un point de vue classique le temps peut être défini par rapport au mouvement d'un objet. Par exemple, la rotation de la terre ou encore le mouvement d'un pendule.¹ Ainsi du principe d'inertie on peut en déduire que le temps s'écoule de manière uniforme dans un référentiel Galiléen. Ceci sous entend que ce n'est pas nécessairement le cas pour un référentiel non Galiléen.

Conservation de l'impulsion et définition de la masse Soit un ensemble de N particules de vitesses \mathbf{v}_i , $i = 1, \dots, N$. Ces particules peuvent interagir entre elles mais sont isolées du reste de l'univers. Il existe un ensemble de quantités associées aux particules et noté m_i tel que dans un référentiel Galiléen alors

$$\sum_{i=1}^N m_i \mathbf{v}_i = \text{cste.} \quad (1.13)$$

Les quantités m_i sont les masses des particules. Ce principe permet d'introduire d'une part le concept de masse mais également d'impulsion $\mathbf{p}_i = m_i \mathbf{v}_i$, en exprimant la conservation de l'impulsion totale d'un système isolé. En général ce principe est démontré à partir des lois de Newton. Ici on préférera admettre ce principe comme étant un axiome de la mécanique et dont on peut déduire les lois de Newton.

Principe fondamental de la dynamique : L'état initial d'un système mécanique composé par l'ensemble des positions et des vitesses des points matériels le constituant détermine de manière unique son mouvement au cours du temps. L'expérience montre en effet qu'il n'est pas nécessaire de connaître ni l'accélération ni ses dérivées temporelles à l'instant initial pour connaître l'évolution du système. A partir de ce principe on peut en déduire que l'accélération d'un point matériel est nécessairement une fonction des positions et des vitesses.

$$\ddot{\mathbf{r}} = f(\mathbf{r}, \dot{\mathbf{r}}, t). \quad (1.14)$$

1. Le pendule est modélisé par un oscillateur harmonique qui est l'équivalent classique d'une transition atomique en physique quantique.

Ceci permet d'introduire la notion de force. La force appliquée à une particule de masse m est définie par

$$\mathbf{F} = m f(\mathbf{r}, \dot{\mathbf{r}}, t) = m \ddot{\mathbf{r}} \quad (1.15)$$

Cette force est nécessairement une fonction de la position de la vitesse et du temps $\mathbf{F} = \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}, t)$. Ce principe doit être compris comme une définition de la force plus que comme une loi fondamentale. La connaissance des forces n'est en effet possible que par l'effet qu'elles produisent.

L'ensemble de ces trois principes est en fait équivalent aux trois lois de Newton dans sa forme habituelle. Par exemple la troisième loi de Newton, la loi de l'action et la réaction, se trouve en fait caché dans le principe de conservation de l'impulsion pour le cas de deux particules isolées.

III. Formulation énergétique de la mécanique classique

C'est Gottfried Wilhelm Leibniz (1646–1716) qui inventa le concept d'énergie cinétique sous le nom de *force vive*. Dans le cadre de chocs élastiques, Leibniz s'opposait à René Descartes (1596–1650) en soutenant que la force vive devait se conserver et non pas la quantité de mouvement. Pourtant on sait que ces deux concepts ne s'opposent pas. La formulation énergétique de la mécanique classique est complètement équivalente à la mécanique Newtonienne elle possède cependant un certain nombre d'avantages. En particulier on montre par cette approche que tout système unidimensionnel et conservatif est intégrable.² De plus la mécanique analytique repose complètement sur le concept d'énergie. Pour un système à une particule, l'approche énergétique s'exprime grâce au théorème de l'énergie cinétique

$$T(t_2) - T(t_1) = W_{12} = \int_{t_1}^{t_2} \mathbf{F} \cdot \mathbf{v} dt, \quad (1.16)$$

où $T(t)$ est l'énergie cinétique au temps t et W_{12} est le travail de la force \mathbf{F} . Si la force est indépendante de la vitesse et du temps alors l'expression du travail s'exprime comme une intégrale le long du chemin suivi

$$T(t_2) - T(t_1) = \int_C \mathbf{F}(\mathbf{r}) \cdot d\mathbf{l} \quad (1.17)$$

On peut encore simplifier le problème dans le cas de forces qui dérivent d'un potentiel

$$\mathbf{F}(\mathbf{r}) = -\nabla V(\mathbf{r}). \quad (1.18)$$

Dans ce cas le travail s'écrit $W_{12} = V(\mathbf{r}_1) - V(\mathbf{r}_2)$ et on peut définir le concept d'énergie mécanique $E_m = T + V = \text{cste}$.

2. D'une manière générale, on montre que pour un système de n degrés de liberté, le système est intégrable s'il existe n constantes du mouvement indépendantes.

Application aux problèmes unidimensionnels Soit une particule de masse m dont la position est décrite par la coordonnée x . La particule est sujet à l'énergie potentielle $V(x)$. Le système est conservatif si bien que l'énergie

$$E = \frac{1}{2}m\dot{x}^2 + V(x), \quad (1.19)$$

est une quantité conservée. Cette propriété va permettre de trouver une solution par quadrature au problème pour tout potentiel $V(x)$. Exprimons la vitesse \dot{x} en fonction de la position

$$\dot{x} = \pm \sqrt{\frac{2}{m}} \sqrt{E - V(x)}, \quad (1.20)$$

où le signe \pm exprime le fait que pour un système conservatif, pour toute solution, la solution inverse obtenue par renversement du temps est également solution. On peut donc écrire la variation du temps en fonction de la variation de l'espace

$$dt = \pm \sqrt{\frac{m}{2}} \frac{dx}{\sqrt{E - V(x)}} \quad (1.21)$$

Cette expression s'intègre pour donner le temps en fonction de la position x

$$t(x) = \pm \sqrt{\frac{m}{2}} \int_{x_0}^x \frac{dx'}{\sqrt{E - V(x')}} \quad (1.22)$$

A partir de l'expression de $t(x)$ on peut en principe trouver la fonction $x(t)$. La solution (1.22) constitue la solution sous la forme de quadrature d'un système 1D (c'est à dire sous la forme d'une intégrale).

IV. Repères tournants

Application antisymétrique Soit E^n un espace Euclidien de dimension n . Soit f une application de linéaire E^n dans E^n . L'application est antisymétrique lorsque la propriété suivante est vérifiée

$$\forall \mathbf{u} \text{ et } \mathbf{v} \in E^n, \quad \mathbf{u} \cdot f(\mathbf{v}) = -\mathbf{v} \cdot f(\mathbf{u}) \quad (1.23)$$

Soit $\mathbf{e}_i, i = 1, \dots, n$ une base de E^n . Le vecteur \mathbf{u} se décompose sur la base

$$\mathbf{u} = \sum_{i=1}^n u_i \mathbf{e}_i \quad (1.24)$$

et l'application f est représenté par la matrice \mathbf{A}

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad (1.25)$$

les coefficients de la matrice \mathbf{A} sont donnés par $a_{ij} = \mathbf{e}_i \cdot f(\mathbf{e}_j)$. La matrice d'une application antisymétrique est antisymétrique

$$a_{ij} = -a_{ji} \quad (1.26)$$

Pour $n = 3$, il existe une dualité entre application linéaire antisymétrique et produit vectoriel. En effet pour $n = 3$, le nombre d'éléments indépendants et non nuls de la matrice \mathbf{A} est $n(n-1)/2 = 3$. Il est donc possible de représenter l'application linéaire grâce à un vecteur de E^3 . Si l'on introduit les nombres $R_1 = a_{32}$, $R_2 = a_{13}$ et $R_3 = a_{21}$, la matrice \mathbf{A} s'écrit

$$\mathbf{A} = \begin{pmatrix} 0 & -R_3 & R_2 \\ R_3 & 0 & -R_1 \\ -R_2 & R_1 & 0 \end{pmatrix} \quad (1.27)$$

Si bien que l'on a

$$\forall \mathbf{u} \in E^3 \quad f(\mathbf{u}) = \mathbf{R} \times \mathbf{u} \quad (1.28)$$

A partir de cette expression on retrouve très simplement la propriété d'anti-symétrie (Eq. 1.23).

Repère mobile Dans le cas d'un repère mobile, les directions des vecteurs unitaires dépendent des coordonnées du point. Cependant la relation d'orthonormalité reste vérifiée

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}. \quad (1.29)$$

La différentielle de cette équation donne

$$d\mathbf{e}_i \cdot \mathbf{e}_j + \mathbf{e}_i \cdot d\mathbf{e}_j = 0 \quad (1.30)$$

Ceci montre que la différentielle d peut être vu comme une application linéaire antisymétrique et on peut écrire

$$d\mathbf{e}_i = d\mathbf{R} \times \mathbf{e}_i, \quad (1.31)$$

où $d\mathbf{R}$ est le vecteur de rotation infinitésimale.

Repère cylindrique Dans le cas du repère cylindrique $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_z)$, la position du système est caractérisée par les trois coordonnées r , θ et z . La rotation du repère est caractérisé par le vecteur de rotation $d\mathbf{R} = d\theta \mathbf{e}_z$. On obtient les différentielles des vecteurs unitaires

$$d\mathbf{e}_r = \mathbf{e}_\theta d\theta, \quad (1.32)$$

$$d\mathbf{e}_\theta = -\mathbf{e}_r d\theta. \quad (1.33)$$

Repère sphérique Dans le cas du repère sphérique $(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi)$, la position du point est caractérisée par les trois coordonnées r, θ et φ . Le vecteur de rotation infinitésimale s'écrit

$$d\mathbf{R} = d\theta\mathbf{e}_\varphi + d\varphi\mathbf{e}_z = d\theta\mathbf{e}_\varphi + \cos\theta d\varphi\mathbf{e}_r - \sin\theta d\varphi\mathbf{e}_\theta \quad (1.34)$$

et les différentielles des vecteurs unitaires peuvent s'écrire sous la forme

$$d\mathbf{e}_r = \mathbf{e}_\theta d\theta + \sin\theta\mathbf{e}_\varphi d\varphi, \quad (1.35)$$

$$d\mathbf{e}_\theta = -\mathbf{e}_r d\theta + \cos\theta\mathbf{e}_\varphi d\varphi, \quad (1.36)$$

$$d\mathbf{e}_\varphi = -\cos\theta\mathbf{e}_\theta d\varphi - \sin\theta\mathbf{e}_r d\varphi. \quad (1.37)$$

Repère de Frenet-Serret Pour finir, on peut considérer le repère de Frenet-Serret $(\mathbf{t}, \mathbf{n}, \mathbf{b})$. Dans ce cas la position est donnée par l'abscisse curviligne s le long de la courbe gauche que suit le point. Le vecteur tangent \mathbf{t} est défini comme le vecteur unitaire tangent à la courbe.

$$\mathbf{t} = \frac{d\mathbf{r}}{ds} \quad (1.38)$$

La variation du vecteur unitaire \mathbf{t} par une variation de l'abscisse curviligne le long du cercle osculateur de rayon de courbure $R(s) = 1/\gamma(s)$ donne la direction du vecteur normal \mathbf{n}

$$\frac{d\mathbf{t}}{ds} = \gamma\mathbf{n}. \quad (1.39)$$

Les vecteurs \mathbf{t} et \mathbf{n} forment alors le plan osculateur. Le vecteur binormal \mathbf{b} permet de terminer le trièdre il est défini par

$$\mathbf{b} = \mathbf{t} \times \mathbf{n}. \quad (1.40)$$

Les vecteurs \mathbf{t} et \mathbf{b} forment le plan rectifiant et les vecteurs \mathbf{n} et \mathbf{b} forment le plan normal. Dans le cas de courbes gauches (non planes) la variation du vecteur binormal est donnée en fonction de la torsion $\kappa(s)$ tel que

$$\frac{d\mathbf{b}}{ds} = -\kappa\mathbf{n}. \quad (1.41)$$

On peut définir le vecteur de Darboux $\boldsymbol{\omega}$.

$$\frac{d\mathbf{t}}{ds} = \boldsymbol{\omega} \times \mathbf{t}, \quad (1.42)$$

$$\frac{d\mathbf{n}}{ds} = \boldsymbol{\omega} \times \mathbf{n}, \quad (1.43)$$

$$\frac{d\mathbf{b}}{ds} = \boldsymbol{\omega} \times \mathbf{b} \quad (1.44)$$

A partir des résultats précédents on montre facilement que ce vecteur s'écrit

$$\boldsymbol{\omega} = \gamma\mathbf{b} + \kappa\mathbf{t}. \quad (1.45)$$

On obtient alors les relations suivantes

$$\frac{d\mathbf{t}}{ds} = \gamma\mathbf{n}, \quad (1.46)$$

$$\frac{d\mathbf{n}}{ds} = -\gamma\mathbf{t} + \kappa\mathbf{b}, \quad (1.47)$$

$$\frac{d\mathbf{b}}{ds} = -\kappa\mathbf{n}. \quad (1.48)$$

Première partie

Formulation Lagrangienne

2. Équations de Lagrange

Objectifs : le but de ce chapitre est de discuter une reformulation de la mécanique newtonienne pour traiter les problèmes avec contraintes. Ce faisant, nous allons aboutir à un jeu d'équations, les équations de Lagrange qui s'exprime en fonction de coordonnées généralisées. Nous verrons que ces équations sont covariantes : elles conservent la même forme quelque soit le jeu de variables choisi.

I. Contraintes et principe de d'Alembert

1. Principe de d'Alembert au travers d'un exemple

Nous considérons une particule de masse m , repérée par le vecteur position $\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z$ et astreinte à se déplacer sur une surface support. Le traitement du problème en mécanique newtonienne se fait en utilisant le principe fondamental de la dynamique (PFD)

$$m\ddot{\mathbf{r}} = \mathbf{F} + \mathbf{R} \quad (2.1)$$

dans lequel on introduit une force de réaction du support \mathbf{R} qui est normale localement à la surface et qui empêche la particule de passer au travers. \mathbf{F} est la somme des forces agissant sur la particule, incluant possiblement les forces de frottements parallèles au mouvement. Notons que le PFD est une équation vectorielle et que les variables position et vitesse sont des vecteurs. Dans les cas simples (plan, sphère, pendule, ...), une projection du PFD avec un jeu de coordonnées bien choisi suffit à traiter la contrainte.

Une surface à 2D dans un espace à 3D se définit par une équation implicite de la forme

$$f(\mathbf{r}) = 0 \quad (2.2)$$

Déplacement virtuel : un déplacement virtuel, noté $\delta\mathbf{r}$ est un déplacement infinitésimal qui satisfait à la contrainte (ici, qui reste sur la surface mais pas forcément associé à la trajectoire).

Ainsi, par définition, si l'on considère deux points séparés de façon infinitésimale sur la surface, nous aurons, par définition de f :

$$f(\mathbf{r} + \delta\mathbf{r}) - f(\mathbf{r}) = 0 = \delta f = \nabla f(\mathbf{r}) \cdot \delta\mathbf{r} \quad (2.3)$$

Ces remarques simples nous permettent tout d'abord de voir que localement en un point de la surface, la direction du vecteur $\nabla f(\mathbf{r})$ donne la normale à la surface en ce point, de sorte que la force de réaction doit lui être proportionnelle

$$\mathbf{R}(t) = \lambda(t) \nabla f(\mathbf{r}(t)) \quad (2.4)$$

Ainsi, une formulation du problème dans sa forme "newtonienne" consiste à écrire les équations pour la trajectoire $\mathbf{r}(t)$ sous la forme

$$\begin{aligned} m\ddot{\mathbf{r}} &= \mathbf{F} + \lambda \nabla f(\mathbf{r}) \\ f(\mathbf{r}) &= 0 \end{aligned} \quad (2.5)$$

Il y a alors 3 + 1 équations à résoudre et 3 + 1 inconnues qui sont $\mathbf{r}(t)$ et $\lambda(t)$ qui sont des fonctions du temps et dont l'obtention résout le problème. En général, seule la trajectoire $\mathbf{r}(t)$ est recherchée.

D'Alembert propose dans son principe éponyme une reformulation équivalente faisant intervenir la contrainte à travers le concept de déplacement virtuel, caractérisé par (2.3). Les équations s'obtiennent en projetant le PFD selon tout déplacement virtuel qui, eux, satisfont automatiquement la contrainte. La force de réaction disparaît alors et l'on obtient

$$\begin{aligned} (\mathbf{F} - m\ddot{\mathbf{r}}) \cdot \delta\mathbf{r} &= 0 \quad \forall \delta\mathbf{r} \text{ tel que } \nabla f(\mathbf{r}) \cdot \delta\mathbf{r} = 0 \\ f(\mathbf{r}) &= 0 \end{aligned} \quad (2.6)$$

On peut vérifier qu'il y a équivalence entre les deux formulations.

Exemple : bille sur un plan incliné. – Pour illustrer le principe de D'Alembert on peut considérer le cas simple d'une masse ponctuelle astreinte à se déplacer sur le plan incliné $f(\mathbf{r}) = x + y + z = 0$, avec $\nabla f = (1, 1, 1)$, et soumise à l'accélération gravitationnelle $\mathbf{F} = -mg\mathbf{e}_z$. Dans ce cas, le principe de D'Alembert s'écrit

$$\ddot{x}\delta x + \ddot{y}\delta y + (g + \ddot{z})\delta z = 0 \quad \forall \delta\mathbf{r} \text{ tel que } \delta x + \delta y + \delta z = 0 \quad (2.7)$$

On peut exprimer δz en fonction de δx et δy ce qui donne

$$(\ddot{x} - \ddot{z} - g)\delta x + (\ddot{y} - \ddot{z} - g)\delta y = 0 \quad \forall \delta x, \delta y \quad (2.8)$$

Ici ayant éliminé δz , on peut traiter δx et δy comme variables indépendantes. On obtient

$$\ddot{x} - \ddot{z} - g = 0, \quad \ddot{y} - \ddot{z} - g = 0 \quad (2.9)$$

Au final et en utilisant la contrainte $x + y + z = 0$ on obtient facilement

$$\ddot{x} = \ddot{y} = \frac{g}{3}, \quad \ddot{z} = -\frac{2g}{3} \quad (2.10)$$

Cet exemple très simple montre que grâce au principe de D'Alembert, il n'est pas nécessaire d'introduire la réaction du plan.

2. Généralisation à N particules et classification des contraintes

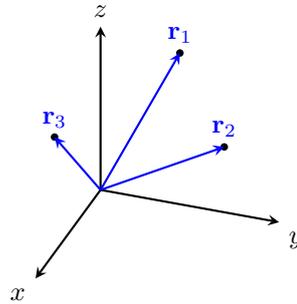


FIGURE 2.1. – En mécanique Newtonienne la description d'un système de trois particules repose sur l'utilisation de trois vecteurs \mathbf{r}_1 , \mathbf{r}_2 , \mathbf{r}_3 de l'espace Euclidien de dimension 3.

Description d'un système à N particules. En mécanique newtonienne, la dynamique de N masses ponctuelles évoluant dans l'espace à trois dimensions est décrite par N vecteurs positions \mathbf{r}_i , avec $i = 1, \dots, N$ (voir Fig. 2.1) auxquelles s'ajoutent les N vitesses \mathbf{v}_i . Il y a donc $3N$ coordonnées pour repérer les positions. Positions \mathbf{r}_i et vitesses \mathbf{v}_i sont des *vecteurs* et les équations de Newton s'écrivent de façon vectorielle dans leur forme générale. On les projette habituellement dans une base bien choisie pour obtenir les équations algébriques à résoudre, même s'il arrive que l'on puisse les résoudre directement sous forme vectorielle.

Classification des contraintes. Physiquement, les systèmes mécaniques présentent différents types de contraintes (voir Fig 2.2)

- Lorsque les masses sont reliées par des liaisons rigides (pendule, bielles, ...) ou qu'elles appartiennent à un solide indéformable. Nous avons alors $\|\mathbf{r}_i - \mathbf{r}_j\| = \text{const.}$

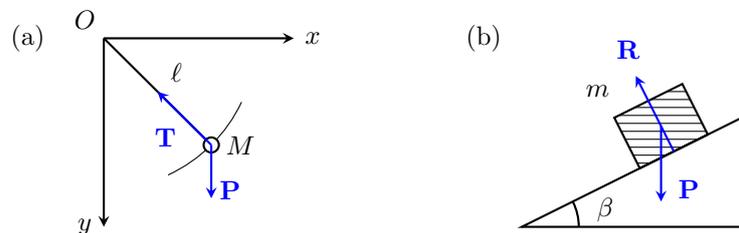


FIGURE 2.2. – (a) Problème du pendule simple : La barre OM est rigide, il existe une force de tension \mathbf{T} qui maintient la masse m le long de la trajectoire circulaire. (b) Problème du bloc glissant le long d'un plan incliné : Pour que le bloc reste sur le plan, il existe une force de réaction \mathbf{R} orthogonale à la surface du plan.

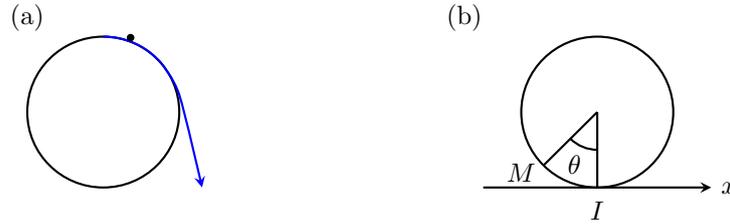


FIGURE 2.3. – (a) Une bille glisse le long d’une sphère. La contrainte est non holonome. (b) Un cerceau roule sans glisser le long d’une droite. La contrainte est semi-holonome.

- Un anneau ou une “perle” sur un fil. La particule suit alors une courbe $y = f(x)$ déterminée.
- Une particule sur une sphère $\|\mathbf{r}\| \geq R^2$ avec pour origine le centre de la sphère ou, plus généralement, sur une surface qu’elle ne peut traverser mais avec la possibilité de décoller (skieur!).
- Des contraintes de non-glissement, comme le fait que sur une sphère qui roule sans glisser, le point de contact avec la surface doit avoir une vitesse nulle.

D’un point de vue mathématique, il existe une nomenclature en fonction de la forme que prend la contrainte.

Holonome Lorsqu’une contrainte peut s’exprimer sous la forme d’une équation algébrique de la forme

$$f(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) = 0, \quad (2.11)$$

on parle alors de *contrainte holonome*. Lorsque cette dernière est indépendante du temps on parle de *contrainte holonome scléronome* et lorsque qu’elle dépend explicitement du temps on parle de *contrainte holonome rhéonome*. Par exemple le pendule simple (Fig. 2.2a) possède deux contraintes holonomes scléronomes. La masse est astreinte à se déplacer dans le plan Oxy et la longueur du pendule est fixe si bien que l’on a les deux relations $z = 0$ et $x^2 + y^2 - \ell^2 = 0$.

Non holonome Toutes les contraintes ne pouvant pas s’exprimer sous la forme de l’Eq. (2.11) sont alors dites *non holonome*. C’est le cas d’une contrainte qui s’exprime sous la forme d’une inégalité (exemple de la particule sur une sphère)

$$f(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) \geq 0, \quad (2.12)$$

ou d’une contrainte qui s’exprime sous la forme d’une équation algébrique faisant intervenir les vitesses (exemple du roulement sans glissement)

$$f(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \dot{\mathbf{r}}_2, \dots, \dot{\mathbf{r}}_N, t) = 0. \quad (2.13)$$

Dans le cadre de ce cours, nous n'allons considérer que les contraintes holonomes.

Le cas des contraintes non holonomes est beaucoup plus difficile à traiter. Le lecteur intéressé pourra trouver plus d'informations sur les contraintes non holonomes dans la plus part des livres traitant du sujet de la mécanique analytique.¹

Complément : Il existe des contraintes non holonomes intégrables, on parle parfois de contraintes semi-holonomes. C'est le cas du roulement sans glissement d'un cerceau se déplaçant le long d'une droite (Fig 2.3b). La condition de roulement sans glissement s'écrit

$$\dot{x} = R\dot{\theta}. \quad (2.14)$$

Cette équation s'intègre facilement pour donner

$$x = R\theta + x_0, \quad (2.15)$$

où x_0 est une constante. Dans le cas de la sphère roulant sur un plan, la contrainte est non-intégrable. Ceci se traduit par le fait que l'orientation de la sphère ne dépend pas uniquement des positions initiales et finales mais elle dépend du chemin suivi le long du plan. Ce dernier exemple constitue donc bien une contrainte non holonome. Il est beaucoup plus simple de traiter les contraintes holonomes il est donc préférable dans le cas de contraintes intégrables d'utiliser la forme intégrée de la contrainte même si le prix à payer est d'avoir une contrainte holonome rhéonome (explicitement dépendante du temps).

Principe de d'Alembert. – Enfin, nous généralisons le principe de d'Alembert au cas de N particules et en notant $\delta\mathbf{r}_i$ des déplacements virtuels compatibles avec les contraintes holonomes

$$\sum_{i=1}^N (\mathbf{F}_i - m_i\ddot{\mathbf{r}}_i) \cdot \delta\mathbf{r}_i = 0, \quad \forall \delta\mathbf{r}_i \text{ compatible avec les contraintes.} \quad (2.16)$$

Ce principe est équivalent à l'ensemble des équations de la dynamique $m_i\ddot{\mathbf{r}}_i = \mathbf{F}_i + \mathbf{R}_i$ pour chacune des particules tout en intégrant les contraintes².

Notez qu'une des difficultés est que les déplacements virtuels $\delta\mathbf{r}_i$ ne sont pas indépendants ce qui motive l'introduction d'un jeu de variables indépendantes et adaptées aux contraintes appelées coordonnées généralisées. Par ailleurs, s'il existe plusieurs contraintes f_1, f_2, \dots , déterminer les déplacements virtuels $\delta\mathbf{r}_i$ compatibles avec toutes peut devenir fastidieux. Nous allons également au prochain chapitre montrer qu'il existe une méthode dite des multiplicateurs de Lagrange pour prendre en compte les contraintes dans les équations sans introduire de déplacements virtuels.

1. Certains ouvrages contiennent parfois des erreurs sur ce sujet. Le cas des contraintes non holonomes linéaires constitue toutefois un cas simple qui est correctement décrit dans le livre de Goldstein, *Classical Mechanics*

2. Si le système possède K contraintes holonomes $f_\beta(\mathbf{r}_1, \dots, \mathbf{r}_N, t) = 0$ indicées par β , les déplacements virtuels satisfont à

$$\sum_{i=1}^N \nabla_i f_\beta \cdot \delta\mathbf{r}_i = 0 \quad \forall \beta$$

II. Coordonnées généralisées

1. Définition et exemples

Soit un système de N masses ponctuelles \mathbf{r}_i , $i = 1, \dots, N$ assujetties à K contraintes holonomes

$$f_\beta(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) = 0, \quad \beta = 1, \dots, K. \quad (2.17)$$

Les positions cartésiennes des masses \mathbf{r}_i ne constituent pas des coordonnées indépendantes puisqu'il existe K relations entre les coordonnées du système. Le nombre de degrés de liberté, ou nombre de variables indépendantes pour décrire le système, est $n = 3N - K$.

Définition : un jeu de n coordonnées *indépendantes* qui permet de décrire le système en satisfaisant automatiquement les K contraintes holonomes (2.17) définit n **coordonnées généralisées** $\{q_\alpha\}_{\alpha=1, \dots, n}$. Les N vecteurs positions sont alors fonction de ces coordonnées selon

$$\mathbf{r}_i = \mathbf{r}_i(q_1, q_2, \dots, q_n, t) \quad i = 1, \dots, N. \quad (2.18)$$

Notez que la dépendance explicite en temps peut provenir d'un paramètre de contrôle qui présente une dépendance temporelle connue.

Au lieu de repérer l'état du système par les positions $\{\mathbf{r}_i\}$, nous pouvons le représenter par un point dans l'espace associé aux coordonnées généralisées appelé **variété des configurations** et de dimension n . D'un point de vue topologique, cet espace n'est pas un espace euclidien en général. Cette variété des configurations peut être décrite à partir des coordonnées de positions en inversant (2.18) selon

$$q_\alpha = q_\alpha(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) \quad \alpha = 1, \dots, n. \quad (2.19)$$

Remarques

- S'il n'y a aucune contrainte, il s'agit simplement d'une projection dans un jeu de coordonnées.
- Les coordonnées généralisées ne sont en général pas des coordonnées cartésiennes, elles peuvent être des distances, des angles, etc. Elles n'ont pas nécessairement la dimension physique d'une distance.
- Les coordonnées généralisées peuvent en principe dépendre explicitement du temps. C'est le cas notamment lorsque nous avons des contraintes holonomes rhéonomes.

Exemples (voir figure 2.4)

Le pendule simple : la longueur ℓ du pendule est un paramètre de contrôle. Les contraintes s'écrivent $x^2 + y^2 - \ell^2 = 0$ et $z = 0$ si bien que le nombre de degrés

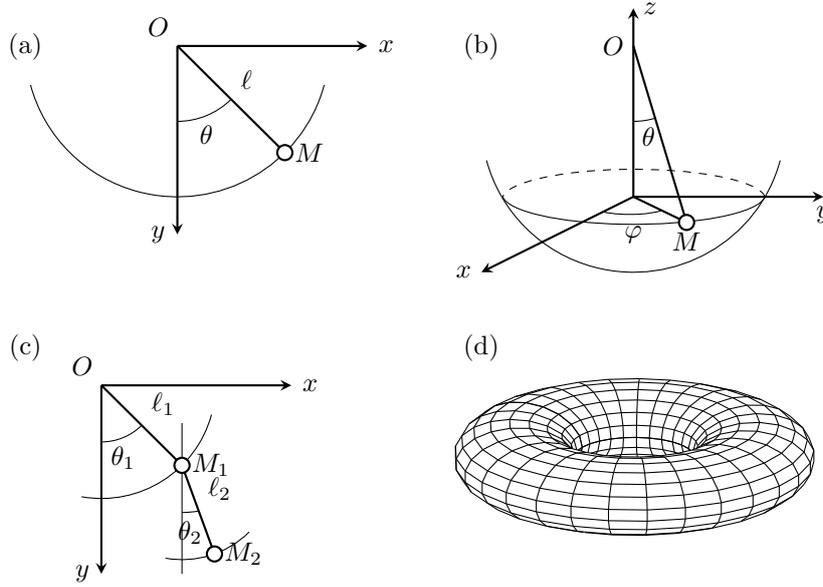


FIGURE 2.4. – (a) Pendule simple : la variété des configurations est le cercle. (b) Pendule sphérique : la variété des configurations est la surface de la sphère. (c-d) Pendule double : repéré par deux angles θ_1 et θ_2 , a variété des configurations est le tore.

de liberté est $n = 3 - 2 = 1$. Le système peut être décrit par l'angle θ défini entre la masse et la verticale. La variété des configurations est le cercle.

Le pendule sphérique : la masse est astreinte à se déplacer sur une sphère de rayon ℓ de sorte que la contrainte s'écrit $\mathbf{r}^2 - \ell^2 = 0$. Le nombre de degrés de liberté est $n = 3 - 1 = 2$. En introduisant un système de coordonnées sphériques, on peut décrire le système par les deux angles θ et φ définis par

$$\begin{aligned} x &= \ell \cos \varphi \sin \theta \\ y &= \ell \sin \varphi \sin \theta \\ z &= -\ell \cos \theta \end{aligned} \quad (2.20)$$

On voit que la variété des configurations est la surface de la sphère.

Perle sur un cerceau : Une perle est astreinte à se déplacer sur un cerceau qui tourne autour de l'axe z selon un mouvement $\varphi(t)$ imposé par l'opérateur (paramètre de contrôle) comme représenté sur la figure 2.5. Le système possède un degré de liberté $q_1 = \theta$ avec les mêmes notations angulaires qu'en coordonnées sphériques, ce qui donne

$$\begin{aligned} x &= \ell \cos \varphi(t) \sin \theta \\ y &= \ell \sin \varphi(t) \sin \theta \\ z &= \ell \cos \theta \end{aligned} \quad (2.21)$$

où l'on fait apparaître explicitement la dépendance en temps.

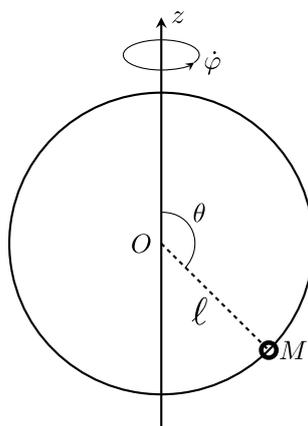


FIGURE 2.5. – Une masse (perle) glisse sans frottement le long d'un cerceau de rayon ℓ en rotation paramétrée par l'angle φ à la vitesse angulaire $\dot{\varphi}$.

Le double pendule (traité en TD) : Le double pendule est constitué d'un premier pendule simple au bout duquel est attaché une masse m_1 . Un second pendule relie la masse m_1 à une masse m_2 . Il constitue l'un des systèmes les plus simple ayant un comportement chaotique. Pour décrire le système on peut utiliser les deux angles θ_1 et θ_2 que font les masses m_1 et m_2 avec la verticale. La variété des configurations est alors la surface d'un tore.

2. Réécriture des déplacements virtuels

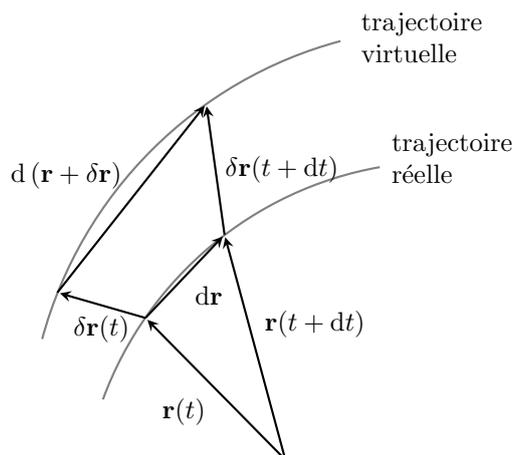


FIGURE 2.6. – Les déplacements virtuels $\delta \mathbf{r}$ ont lieu à un instant t donné alors que les variations naturelles $d\mathbf{r}$ ont lieu entre deux instants différents le long des trajectoires.

Variations naturelles et virtuelles - Comme décrit sur la figure 2.6, on considère deux trajectoires infiniment proches notées $\mathbf{r}(t)$ et $(\mathbf{r} + \delta\mathbf{r})(t)$, séparées par des déplacements virtuels $\delta\mathbf{r}$. Les *variations virtuelles* (qui sont des déplacements virtuels) ont lieu au même temps t de sorte que l'on peut écrire

$$\delta\mathbf{r}(t) = (\mathbf{r} + \delta\mathbf{r})(t) - \mathbf{r}(t) \quad (2.22)$$

$$\delta\mathbf{r}(t + dt) = (\mathbf{r} + \delta\mathbf{r})(t + dt) - \mathbf{r}(t + dt) \quad (2.23)$$

On appelle *variations naturelles* les différentielles le long des trajectoires entre les temps t et $t + dt$:

$$d\mathbf{r}(t) = \mathbf{r}(t + dt) - \mathbf{r}(t) \quad (2.24)$$

$$d(\mathbf{r} + \delta\mathbf{r})(t) = (\mathbf{r} + \delta\mathbf{r})(t + dt) - (\mathbf{r} + \delta\mathbf{r})(t) \quad (2.25)$$

En utilisant l'égalité vectorielle sur le losange de la figure 2.6, on voit que

$$\delta\mathbf{r}(t) + d(\mathbf{r} + \delta\mathbf{r})(t) = d\mathbf{r}(t) + \delta\mathbf{r}(t + dt) = d\mathbf{r}(t) + \delta(\mathbf{r} + d\mathbf{r})(t),$$

Après simplification, l'on obtient la relation

$$d\delta\mathbf{r} = \delta d\mathbf{r}. \quad (2.26)$$

que nous garderons en mémoire pour la suite.

Expression des déplacements virtuels et vitesses généralisées - Comme un déplacement virtuel $\delta\mathbf{r}_i$ opère à *temps constant*, les différences finies dans (2.18) à t fixé, donnent l'expression

$$\delta\mathbf{r}_i = \sum_{\alpha=1}^n \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \delta q_\alpha, \quad (2.27)$$

où les δq_α sont des variations quelconques.

Si l'on considère maintenant la variation naturelle, ou différentielle, le long de la trajectoire, elle se fait entre les temps t et $t + dt$, de sorte que figure 2.6).

$$d\mathbf{r}_i = \frac{\partial \mathbf{r}_i}{\partial t} dt + \sum_{\alpha=1}^n \frac{\partial \mathbf{r}_i}{\partial q_\alpha} dq_\alpha \quad (2.28)$$

En simplifiant (2.28) par dt , la vitesse $\mathbf{v}_i = \frac{d\mathbf{r}_i}{dt}$ s'écrit donc

$$\mathbf{v}_i = \dot{\mathbf{r}}_i = \frac{\partial \mathbf{r}_i}{\partial t} + \sum_{\alpha=1}^n \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \dot{q}_\alpha, \quad (2.29)$$

où la dérivée partielle par rapport au temps s'effectue à coordonnées généralisées fixées. On remarque que la vitesse $\dot{\mathbf{r}}_i$ est à la fois une fonction des coordonnées généralisées, des **vitesses généralisées** $\dot{q}_\alpha = \frac{dq_\alpha}{dt}$ et du temps

$$\dot{\mathbf{r}}_i = \mathbf{r}_i(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t) \quad (2.30)$$

Exemple de la perle sur le cerceau : la dérivation par rapport aux temps de (2.21) donne

$$\begin{aligned}\dot{x} &= \ell \left(-\sin \varphi(t) \sin \theta \dot{\varphi}(t) + \cos \varphi(t) \cos \theta \dot{\theta} \right) \\ \dot{y} &= \ell \left(\cos \varphi(t) \sin \theta \dot{\varphi}(t) + \sin \varphi(t) \cos \theta \dot{\theta} \right) \\ \dot{z} &= \ell(-\sin \theta \dot{\theta})\end{aligned}\tag{2.31}$$

avec à la fois des contributions venant de la dépendance explicite en temps (avec les $\dot{\varphi}$) et des vitesses généralisées $\dot{\theta}$.

Sur l'indépendance des variables en mécanique analytique – Avec des équations du second ordre par rapport au temps, la dynamique newtonienne est caractérisée par la donnée des positions et des vitesses. De même, la dynamique lagrangienne est caractérisée par la donnée des q_α et des \dot{q}_α qui sont traitées comme des *variables indépendantes* pour établir les équations du mouvement.

Une manière de comprendre que positions et vitesses sont des variables indépendantes est de voir qu'il n'existe pas de relation algébriques entre elles à *un instant donné*. Il existe entre elles une relation différentielle mais qui fait intervenir les deux instants t et $t + dt$. La dynamique de Newton peut se réécrire comme un système différentiel du *premier ordre* par rapport au temps dans le jeu de variables *indépendantes* $(\mathbf{r}(t), \mathbf{v}(t))$ selon

$$m \frac{d}{dt} \begin{pmatrix} \mathbf{r} \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} m\mathbf{v} \\ \mathbf{F}(\mathbf{r}, \mathbf{v}, \mathbf{t}) \end{pmatrix}\tag{2.32}$$

La première équation lie la variable vitesse à la dérivée de la position (cinématique), la seconde est l'équation de la dynamique. Il faut préciser les conditions initiales *indépendantes* $(\mathbf{r}(0), \mathbf{v}(0))$ pour déterminer de façon univoque la solution. Le passage aux coordonnées généralisées doit être compris comme le changement de variables $(\mathbf{r}(t), \mathbf{v}(t)) \rightarrow (q_\alpha, \dot{q}_\alpha)$ qui s'inverse a priori en

$$\mathbf{r} = \mathbf{r}(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t)\tag{2.33}$$

$$\mathbf{v} = \mathbf{v}(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t)\tag{2.34}$$

avec le fait que, dans la première équation, les positions ne dépendent pas des vitesses généralisées. Les équations du mouvement pour $(q_\alpha, \dot{q}_\alpha)$ vont être du premier ordre, et l'ordre le plus élevé sera du second ordre, en \ddot{q}_α .

Complément : Pour préciser les choses du point de vue mathématique, bien que ce ne sera pas essentiel pour la suite et la géométrie différentielle pas l'objet de ce cours, notons que la variété des configurations constitue une variété différentiable. Ce n'est pas un espace vectoriel et cela ne fait pas de sens d'additionner des points de l'espace des configurations. C'est cependant un espace courbe pour lequel on peut définir localement un espace vectoriel tangent sur lequel on peut définir des variations élémentaires et en particulier des vitesses

généralisées $\dot{q}_\alpha = \frac{dq_\alpha}{dt}$. On peut ainsi construire un espace de dimension $2n$ regroupant les coordonnées et les vitesses généralisées, traitées comme variables indépendantes (q_α, \dot{q}_α) dans les variables des fonctions et les dérivées partielles.

III. Équations de Lagrange

1. Démonstration

Nous recherchons maintenant les équations du mouvement pour les coordonnées généralisées q_α . Pour ce faire, partons du principe de d'Alembert (2.16) et utilisons (2.27) pour obtenir

$$\sum_{\alpha} \left(Q_{\alpha} - \sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_{\alpha}} \right) \delta q_{\alpha} = 0, \quad \forall \delta q_{\alpha} \quad (2.35)$$

où Q_{α} est la **force généralisée** définie par

$$Q_{\alpha} = \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_{\alpha}}. \quad (2.36)$$

Sachant que les δq_{α} sont indépendantes, on a donc

$$\sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_{\alpha}} = Q_{\alpha}, \quad \alpha = 1, \dots, n \quad (2.37)$$

Nous transformons le terme de gauche via une forme d'intégration par partie

$$\sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_{\alpha}} = \sum_i \frac{d}{dt} \left(m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_{\alpha}} \right) - \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{r}_i}{\partial q_{\alpha}} \right)$$

Pour le premier terme, on remarque que les vitesses cartésiennes sont linéaires dans les vitesses généralisées. À partir de (2.29), on en déduit l'expression

$$\frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_{\alpha}} = \frac{\partial \mathbf{r}_i}{\partial q_{\alpha}}. \quad (2.38)$$

Pour le second terme, en dérivant (2.29) en tenant compte du fait que les variables q_{α} et \dot{q}_{α} sont indépendantes

$$\frac{\partial \dot{\mathbf{r}}_i}{\partial q_{\alpha}} = \frac{\partial^2 \mathbf{r}_i}{\partial t \partial q_{\alpha}} + \sum_{\beta} \frac{\partial^2 \mathbf{r}_i}{\partial q_{\alpha} \partial q_{\beta}} \dot{q}_{\beta}. \quad (2.39)$$

Cette dernière expression se met sous la forme simple³

$$\frac{\partial \dot{\mathbf{r}}_i}{\partial q_{\alpha}} = \frac{d}{dt} \left(\frac{\partial \mathbf{r}_i}{\partial q_{\alpha}} \right). \quad (2.40)$$

3. On remarquera que la dérivée totale par rapport au temps et la dérivée partielle par rapport aux vitesses généralisées ne commutent pas. En effet, on peut voir que

$$\frac{\partial}{\partial \dot{q}_{\alpha}} \frac{d}{dt} \mathbf{r}_i = \frac{\partial \mathbf{r}_i}{\partial q_{\alpha}} \neq \frac{d}{dt} \frac{\partial}{\partial \dot{q}_{\alpha}} \mathbf{r}_i = 0.$$

Il vient finalement

$$\begin{aligned} \sum_i m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} &= \sum_i \frac{d}{dt} \left(m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial \dot{q}_\alpha} \right) - \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial q_\alpha} \\ &= \frac{d}{dt} \left(\frac{\partial}{\partial \dot{q}_\alpha} \left(\sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 \right) \right) - \frac{\partial}{\partial q_\alpha} \left(\sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 \right) \end{aligned} \quad (2.41)$$

où l'on reconnaît ici l'expression de l'énergie cinétique

$$T = \sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^2, \quad (2.42)$$

si bien que le principe de D'Alembert s'écrit sous la forme simple

$$\boxed{\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_\alpha} \right) - \frac{\partial T}{\partial q_\alpha} = Q_\alpha}, \quad \alpha = 1, \dots, n. \quad (2.43)$$

Ce sont les **équations de Lagrange**.

On remarque que, de façon générale, le terme d'énergie cinétique ne dépend pas que des vitesses généralisées \dot{q}_α mais souvent aussi des coordonnées q_α . En effet, nous avons, en simplifiant la discussion au cas d'une transformation $\mathbf{r}_i \rightarrow q_\alpha$ indépendante du temps,

$$T = \sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 = \sum_{\alpha, \beta} \frac{M_{\alpha\beta}}{2} \dot{q}_\alpha \dot{q}_\beta \quad \text{avec} \quad M_{\alpha\beta} = \sum_i m_i \frac{\partial \mathbf{r}_i}{\partial q_\alpha} \cdot \frac{\partial \mathbf{r}_i}{\partial q_\beta} \quad (2.44)$$

des fonctions qui dépendent a priori des variables q_α .

Exemple de la perle sur le cerceau : en utilisant (2.31), on trouve facilement l'expression de l'énergie cinétique

$$T = \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{m\ell^2}{2} (\dot{\theta}^2 + \dot{\varphi}^2(t) \sin^2 \theta) \quad (2.45)$$

qui illustre la remarque précédente. Pour la force généralisée Q_θ associée à la gravité, on écrit simplement

$$-mg\delta z = mg\ell \sin \theta \delta \theta = Q_\theta \delta \theta. \quad (2.46)$$

En appliquant (2.47) à $q_1 = \theta$, il vient

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{\theta}} \right) - \frac{\partial T}{\partial \theta} = \frac{m\ell^2}{2} (2\ddot{\theta} - \dot{\varphi}^2(t) 2 \sin \theta \cos \theta) = Q_\theta = mg\ell \sin \theta \quad (2.47)$$

et l'équation du mouvement

$$\ddot{\theta} - \left(\frac{g}{\ell} + \dot{\varphi}^2(t) \cos \theta \right) \sin \theta = 0 \quad (2.48)$$

c'est-à-dire l'équation d'un pendule modifiée par la force centrifuge⁴. On voit ici la grande simplicité d'utilisation des équations de Lagrange une fois établies l'expression de l'énergie cinétique et des forces généralisées.

4. Noter que $\theta = \pi$ est la position d'équilibre quand $\dot{\varphi} = 0$.

2. Forces dérivant d'un potentiel - Lagrangien

Pour le cas de forces dérivant d'un potentiel ne dépendant que des positions $\mathbf{F}_i = -\nabla_i U$, avec $\nabla_i = (\partial_{x_i}, \partial_{y_i}, \partial_{z_i})$, nous avons conservation de l'énergie mécanique $E_m = T + U$. On s'attend donc à pouvoir combiner U à T dans les équations de Lagrange. D'ailleurs, on observe plus généralement que, si le potentiel dépend des positions et aussi explicitement du temps, on peut écrire

$$Q_\alpha = - \sum_i \nabla_i U \cdot \frac{\partial \mathbf{r}_i}{\partial q_\alpha} = - \frac{\partial U}{\partial q_\alpha}(q_1, \dots, q_n, t), \quad (2.49)$$

avec $U = U(q_1, \dots, q_n, t)$ exprimé en fonction des coordonnées généralisées, les équations de Lagrange s'écrivent simplement

$$\boxed{\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) - \frac{\partial \mathcal{L}}{\partial q_\alpha} = 0}, \quad \alpha = 1, \dots, n \quad (2.50)$$

où $\mathcal{L}(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t) = T(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t) - U(q_1, \dots, q_n, t)$ est la fonction de Lagrange, encore appelée **lagrangien**. On retiendra une notation plus condensée sous la forme

$$\boxed{\mathcal{L}(q_\alpha, \dot{q}_\alpha, t) = T(q_\alpha, \dot{q}_\alpha, t) - U(q_\alpha, t)}. \quad (2.51)$$

On portera une attention particulière au signe négatif qui différencie la fonction de Lagrange de l'énergie mécanique ! Les équations (2.50) sont complètement équivalentes aux équations de Newton, leur expression toutefois permettra d'aborder un certain nombre de problème en mécanique de manière beaucoup plus simple puisque les contraintes géométriques sont naturellement satisfaites dans ce formalisme. Les forces dérivant d'un potentiel (Eq. 2.49) ne sont pas nécessairement conservatives car le potentiel peut dépendre explicitement du temps. Nous préférons donc utiliser le terme de *forces dérivant d'un potentiel*.

En présence de forces ne dérivant pas d'un potentiel, on écrira alors (2.50) sous la forme

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) - \frac{\partial \mathcal{L}}{\partial q_\alpha} = Q'_\alpha, \quad \alpha = 1, \dots, n \quad (2.52)$$

où les Q'_α ne contiennent que les forces ne dérivant pas d'un potentiel.

Exemple de la perle sur le cerceau : avec $U = mgl \cos \theta$, on obtient

$$\mathcal{L}(\theta, \dot{\theta}, t) = \frac{m\ell^2}{2} (\dot{\theta}^2 + \dot{\varphi}^2(t) \sin^2 \theta) - mgl \cos \theta. \quad (2.53)$$

3. Potentiel généralisé et force de Lorentz

La forme des équations de Lagrange (2.50) suggère qu'il est possible de généraliser le concept de potentiel. En effet, il est clair que les équations de Lagrange et le lagrangien conserveront la

forme (2.50) pour toute force généralisée dérivant d'un *potentiel généralisé* $U(q_\alpha, \dot{q}_\alpha, t)$ selon

$$Q_\alpha = \frac{d}{dt} \left(\frac{\partial U}{\partial \dot{q}_\alpha} \right) - \frac{\partial U}{\partial q_\alpha} . \quad (2.54)$$

On parle parfois de couplage gyroscopique pour ce type de forces.

Exemple : la force de Lorentz – Une charge q se déplaçant dans les champs électromagnétiques $(\mathbf{E}(\mathbf{r}, t), \mathbf{B}(\mathbf{r}, t))$ est soumise à la force de Lorentz

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) . \quad (2.55)$$

Nous allons maintenant voir que cette force dérive d'un potentiel généralisé dans le formalisme lagrangien. Pour cela, rappelons que les champs \mathbf{E} et \mathbf{B} dérivent d'un potentiel scalaire noté $V(\mathbf{r}, t)$ et d'un potentiel vecteur $\mathbf{A}(\mathbf{r}, t)$ selon les lois

$$\mathbf{E} = -\nabla_{\mathbf{r}} V - \frac{\partial \mathbf{A}}{\partial t} , \quad \mathbf{B} = \nabla_{\mathbf{r}} \times \mathbf{A} . \quad (2.56)$$

Montrons que la force de Lorentz dérive du potentiel généralisé

$$U(\mathbf{r}, \dot{\mathbf{r}}, t) = q(V(\mathbf{r}, t) - \dot{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r}, t)) . \quad (2.57)$$

En coordonnées cartésiennes, regardons la composante selon x . Nous écrivons, en traitant (x, \dot{x}) comme des variables indépendantes,

$$\frac{\partial U}{\partial x} = q \left(\frac{\partial V}{\partial x} - \dot{x} \frac{\partial A_x}{\partial x} - \dot{y} \frac{\partial A_y}{\partial x} - \dot{z} \frac{\partial A_z}{\partial x} \right) , \quad (2.58)$$

$$\frac{d}{dt} \frac{\partial U}{\partial \dot{x}} = -q \frac{dA_x}{dt} = -q \left(\frac{\partial A_x}{\partial t} + \dot{x} \frac{\partial A_x}{\partial x} + \dot{y} \frac{\partial A_x}{\partial y} + \dot{z} \frac{\partial A_x}{\partial z} \right) . \quad (2.59)$$

On en déduit

$$Q_x = \frac{d}{dt} \left(\frac{\partial U}{\partial \dot{x}} \right) - \frac{\partial U}{\partial x} \quad (2.60)$$

$$= q \left(-\frac{\partial A_x}{\partial t} - \frac{\partial V}{\partial x} + \dot{y} \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) + \dot{z} \left(\frac{\partial A_z}{\partial x} - \frac{\partial A_x}{\partial z} \right) \right) \quad (2.61)$$

$$= q(E_x + (\dot{y}B_z - \dot{z}B_y)) . \quad (2.62)$$

qui est bien la composante selon x de la force de Lorentz. On peut rapidement vérifier que cela marche pour les deux autres composantes, si bien que le lagrangien d'une particule chargée dans un champ électromagnétique pourra s'écrire

$$\boxed{\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) = \frac{1}{2} m \dot{\mathbf{r}}^2 - q(V(\mathbf{r}, t) - \dot{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r}, t))} . \quad (2.63)$$

Complément : les forces d'inertie – Les forces d'inertie dérivent également d'un potentiel généralisé. On considère une particule libre de masse m . On introduit le référentiel Galiléen \mathcal{R}_0 et le référentiel tournant \mathcal{R} . Soit $\boldsymbol{\Omega}$ le vecteur rotation de \mathcal{R} par rapport à \mathcal{R}_0 . La vitesse de la particule dans le référentiel \mathcal{R}_0 est noté \mathbf{v}_0 et s'exprime en fonction de \mathbf{v} la vitesse de la particule dans le référentiel \mathcal{R} selon

$$\mathbf{v}_0 = \mathbf{v} + \boldsymbol{\Omega} \times \mathbf{r} \quad (2.64)$$

où \mathbf{r} est le vecteur position de la particule. La fonction de Lagrange exprimé en fonction de la position \mathbf{r} et de la vitesse \mathbf{v} s'écrit

$$\mathcal{L}(\mathbf{r}, \mathbf{v}) = \frac{1}{2}m\mathbf{v}^2 + m\mathbf{v} \cdot (\boldsymbol{\Omega} \times \mathbf{r}) + \frac{1}{2}m(\boldsymbol{\Omega} \times \mathbf{r})^2 \quad (2.65)$$

On retrouve ici un potentiel généralisé similaire au potentiel d'interaction d'une particule chargée avec le champ électromagnétique. Exprimons les équations de Lagrange

$$\frac{d}{dt}(\nabla_{\mathbf{v}}\mathcal{L}) = \nabla_{\mathbf{r}}\mathcal{L} \quad (2.66)$$

Le gradient par rapport aux vitesses s'exprime selon

$$\nabla_{\mathbf{v}}\mathcal{L} = m\mathbf{v} + m\boldsymbol{\Omega} \times \mathbf{r} \quad (2.67)$$

Le gradient par rapport à la position s'écrit

$$\begin{aligned} \nabla_{\mathbf{r}}\mathcal{L} &= m\nabla_{\mathbf{r}}(\mathbf{r} \cdot (\mathbf{v} \times \boldsymbol{\Omega})) + \frac{1}{2}m\nabla_{\mathbf{r}}(\boldsymbol{\Omega}^2\mathbf{r}^2 - (\boldsymbol{\Omega} \cdot \mathbf{r})^2) \\ &= -m\boldsymbol{\Omega} \times \mathbf{v} + m\boldsymbol{\Omega}^2\mathbf{r} - m(\boldsymbol{\Omega} \cdot \mathbf{r})\boldsymbol{\Omega} = -m\boldsymbol{\Omega} \times \mathbf{v} - m\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) \end{aligned}$$

où l'on a utilisé la propriété $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$. Au final, l'équation du mouvement s'écrit

$$m\dot{\mathbf{v}} = -m\dot{\boldsymbol{\Omega}} \times \mathbf{r} - 2m\boldsymbol{\Omega} \times \mathbf{v} - m\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) \quad (2.68)$$

On reconnaît ici l'expression des forces d'Euler, de Coriolis et centrifuge.

IV. Covariance : invariance de la forme des équations de Lagrange

On considère la transformation des équations de Lagrange sous un changement de coordonnées généralisées défini par la transformation ponctuelle $(q_\alpha, \dot{q}_\alpha) \longrightarrow (u_\alpha, \dot{u}_\alpha)$

$$u_\alpha = u_\alpha(q_1, \dots, q_n, t), \quad \alpha = 1, \dots, n \quad (2.69)$$

qu'on suppose inversible en tout point sous la forme

$$q_\alpha = q_\alpha(u_1, \dots, u_n, t), \quad \alpha = 1, \dots, n. \quad (2.70)$$

On notera $\Lambda_{\alpha\beta}$ la matrice jacobienne de cette dernière transformation⁵

$$\Lambda_{\alpha\beta}(u_1, \dots, u_n, t) = \frac{\partial q_\alpha}{\partial u_\beta}(u_1, \dots, u_n, t). \quad (2.71)$$

En dérivant (2.70), on peut écrire la vitesse généralisée \dot{q}_α selon

$$\dot{q}_\alpha = \frac{\partial q_\alpha}{\partial t} + \sum_{\beta} \frac{\partial q_\alpha}{\partial u_\beta} \dot{u}_\beta = \frac{\partial q_\alpha}{\partial t} + \sum_{\beta} \Lambda_{\alpha\beta} \dot{u}_\beta. \quad (2.72)$$

5. L'inversibilité exige $\det \Lambda \neq 0$.

d'où, comme dans la démonstration des équations de Lagrange, les relations

$$\frac{\partial \dot{q}_\alpha}{\partial \dot{u}_\beta} = \Lambda_{\alpha\beta}, \quad (2.73)$$

$$\frac{\partial \dot{q}_\alpha}{\partial u_\beta} = \frac{\partial}{\partial u_\beta} \left(\frac{\partial q_\alpha}{\partial t} + \sum_{\beta'} \dot{u}_{\beta'} \frac{\partial q_\alpha}{\partial u_{\beta'}} \right) = \frac{\partial}{\partial t} \frac{\partial q_\alpha}{\partial u_\beta} + \sum_{\beta'} \dot{u}_{\beta'} \frac{\partial}{\partial u_{\beta'}} \frac{\partial q_\alpha}{\partial u_\beta} = \frac{d}{dt} \Lambda_{\alpha\beta}. \quad (2.74)$$

La fonction de Lagrange se transforme en $\mathcal{L}(u_\alpha, \dot{u}_\alpha, t) = \mathcal{L}(q_\beta(u_\alpha, t), \dot{q}_\beta(u_\alpha, \dot{u}_\alpha, t), t)$, avec

$$\frac{\partial \mathcal{L}}{\partial u_\alpha} = \sum_\beta \frac{\partial \mathcal{L}}{\partial q_\beta} \frac{\partial q_\beta}{\partial u_\alpha} + \sum_\beta \frac{\partial \mathcal{L}}{\partial \dot{q}_\beta} \frac{\partial \dot{q}_\beta}{\partial u_\alpha} = \sum_\beta \frac{\partial \mathcal{L}}{\partial q_\beta} \Lambda_{\beta\alpha} + \sum_\beta \frac{\partial \mathcal{L}}{\partial \dot{q}_\beta} \frac{d}{dt} \Lambda_{\beta\alpha} \quad (2.75)$$

$$\frac{\partial \mathcal{L}}{\partial \dot{u}_\alpha} = \sum_\beta \frac{\partial \mathcal{L}}{\partial \dot{q}_\beta} \Lambda_{\beta\alpha} \quad (2.76)$$

En dérivant l'équation (2.76) par rapport au temps, il vient

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{u}_\alpha} \right) = \sum_\beta \Lambda_{\beta\alpha} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\beta} \right) + \sum_\beta \frac{\partial \mathcal{L}}{\partial \dot{q}_\beta} \frac{d\Lambda_{\beta\alpha}}{dt} = \sum_\beta \Lambda_{\beta\alpha} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\beta} \right) + \frac{\partial \mathcal{L}}{\partial u_\alpha} - \sum_\beta \Lambda_{\beta\alpha} \frac{\partial \mathcal{L}}{\partial q_\beta}$$

puis l'identité⁶

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{u}_\alpha} \right) - \frac{\partial \mathcal{L}}{\partial u_\alpha} = \sum_\beta \Lambda_{\beta\alpha} \left(\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\beta} \right) - \frac{\partial \mathcal{L}}{\partial q_\beta} \right) \quad (2.77)$$

Il est alors clair que si \mathcal{L} satisfait aux équations de Lagrange, alors \mathcal{L} y satisfait aussi. Ces équations sont donc invariantes par changement de coordonnées, on dit qu'elles ont une forme covariante. C'est un résultat important et qui aurait pu être la motivation de la recherche des équations de Lagrange sans passer par le principe de D'Alembert. On aurait pu rechercher la formulation de la mécanique newtonienne qui reste covariante. Ainsi, dans la formulation lagrangienne, toute la difficulté est déplacée dans la recherche du choix optimale de coordonnées pour décrire le système et l'écriture du lagrangien, les équations du mouvement s'obtiennent toujours de façon très directe.

Exemple : particule à 2D dans un potentiel U . – Changement de coordonnées en polaires $(x, \dot{x}, y, \dot{y}) \longleftrightarrow (r, \dot{r}, \theta, \dot{\theta})$ avec $x = r \cos \theta$ et $y = r \sin \theta$: les équations de Lagrange donne l'équation radiale selon r et le théorème du moment cinétique selon θ . En particulier, dans le cas d'une force centrale ($F_\theta = 0$), elles permettent d'identifier immédiatement une quantité conservée ($L_z = mr^2\dot{\theta}$).

Cartésiennes	Polaires
$\mathcal{L}(x, \dot{x}, y, \dot{y}) = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - U(x, y)$	$\mathcal{L}(r, \dot{r}, \theta, \dot{\theta}) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) - U(r, \theta)$
$\mathbf{F} = F_x \mathbf{e}_x + F_y \mathbf{e}_y = -\frac{\partial U}{\partial x} \mathbf{e}_x - \frac{\partial U}{\partial y} \mathbf{e}_y$	$\mathbf{F} = F_r \mathbf{e}_r + F_\theta \mathbf{e}_\theta = -\frac{\partial U}{\partial r} \mathbf{e}_r - \frac{1}{r} \frac{\partial U}{\partial \theta} \mathbf{e}_\theta$
$Q_x = F_x ; Q_y = F_y$	$Q_r = F_r ; Q_\theta = rF_\theta$
Lagrange selon $x : m\ddot{x} = F_x$	Lagrange selon $r : m(\ddot{r} - r\dot{\theta}^2) = F_r$
Lagrange selon $y : m\ddot{y} = F_y$	Lagrange selon $\theta : m \frac{d}{dt}(mr^2\dot{\theta}) = rF_\theta$

6. On pourra remarquer qu'il s'agit de la formule de changement de base pour un opérateur différentiel, d'où l'apparition de la matrice jacobienne et la nécessité que son déterminant ne soit pas nul.

3. Principe de moindre action

Nous avons dérivé les équations de Euler-Lagrange à partir du principe d'Alembert (c'est-à-dire les équations de Newton). Il est également possible de le dériver à partir d'un principe variationnel, le principe de Hamilton, c'est-à-dire sous la forme d'une recherche d'un extremum d'une fonctionnelle. Une fonctionnelle généralise le concept de fonction qui dépend d'une ou plusieurs variables. Une fonctionnelle va dépendre d'une ou plusieurs fonctions. L'avantage d'un principe variationnel réside dans le fait qu'il s'affranchit de tout système de coordonnées.

I. Calcul des variations

1. Position du problème : le brachistochrone

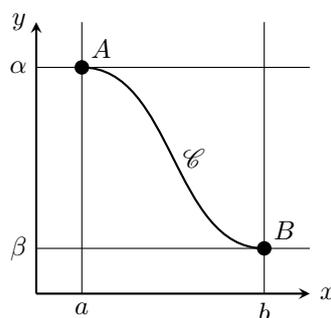


FIGURE 3.1. – Le problème du brachistochrone : Quelle est la courbe \mathcal{C} passant par les points A et B tel que si un mobile glisse sans frottement le long de cette courbe le temps de parcours est minimum ?

Pour illustrer le concept de fonctionnelle nous allons introduire le problème classique de la brachistochrone (Fig. 3.1), c'est-à-dire la courbe de descente la plus rapide. Ce problème fut résolu pour la première fois par Jean Bernoulli en 1696. Soit une courbe $y(x)$ passant par les points A et B de coordonnées (a, α) et (b, β) si bien que $y(a) = \alpha$ et $y(b) = \beta$. Une masse initialement en A et sans vitesse initiale glisse sans frottement le long de la courbe. Quelle courbe minimise le temps de parcours ? On formalise ce problème en écrivant le théorème de l'énergie cinétique

$$\frac{1}{2}mv^2 = mg(\alpha - y), \tag{3.1}$$

si bien que la vitesse v s'écrit

$$v = \sqrt{\dot{x}^2 + \dot{y}^2} = \left| \frac{dx}{dt} \right| \sqrt{1 + y'(x)^2} = \sqrt{2g} \sqrt{\alpha - y(x)}, \quad (3.2)$$

où $y' = dy/dx$. En supposant que $dx/dt > 0 \forall t$, le temps mis pour parcourir la courbe s'écrit

$$T = \int_0^T dt = \int_a^b \frac{dt}{dx} dx = \frac{1}{\sqrt{2g}} \int_a^b \sqrt{\frac{1 + y'(x)^2}{\alpha - y(x)}} dx \quad (3.3)$$

Le temps de parcours $T = T[y(x)]$ est une fonctionnelle car sa valeur dépend de la courbe $y(x)$. La solution du brachistochrone n'est pas la ligne droite mais la courbe cycloïde. Cette solution est détaillée par l'utilisation de l'équation d'Euler dans l'annexe C1. .

2. Fonctionnelle

Définition : soit $y : x \in [a, b] \mapsto y(x) \in \mathbb{R}$ une notation pour des fonctions réelles, \mathcal{Y} un espace de fonctions satisfaisant aux mêmes propriétés (domaine, classe de différentiabilité, ...) alors

$$F[y] : y \in \mathcal{Y} \mapsto F[y] \in \mathbb{R} \quad (3.4)$$

est une fonctionnelle de y que l'on note aussi $F[y(x)]$. En physique, la très grande majorité des fonctionnelles sont sous une forme intégrale

$$F[y] = \int_a^b dx f(y(x), y'(x), \dots, x) \quad (3.5)$$

dans laquelle, la "densité" f est exprimée comme fonction de plusieurs variables, y et ses dérivées, ainsi que x pour une dépendance explicite. Ces variables seront traitées comme indépendantes dans les dérivées partielles de f .

Interprétation intuitive mais non rigoureuse : Si l'on imagine une discrétisation de l'axe des x en points $\{x_i\}_{i=1, N}$, séparés de Δ qui conduit à une discrétisation de la fonction $y(x)$ en variables y_i , la fonctionnelle peut être vue comme le passage à la limite $N \rightarrow \infty$ et $\Delta \rightarrow 0$ de la fonction à plusieurs variables $F(y_1, \dots, y_N) \rightarrow F[y(x)]$ à une infinité de variables. L'indice discret i passant à l'indice continu x .

Par exemple, on peut imaginer le passage d'une somme de Riemann à une fonctionnelle selon

$$F(y_1, \dots, y_N) = \Delta \sum_{i=1}^N \left(\frac{y_{i+1} - y_i}{\Delta} \right)^2 \rightarrow F[y] = \int_a^b dx (y'(x))^2 \quad (3.6)$$

3. Extrémalisation et dérivée fonctionnelle

Rappelons le principe de recherche des extremums d'une fonction

Problème 1D : extrémaliser $F(y)$ équivaut à résoudre l'équation algébrique $\frac{dF}{dy} = 0$ puis étudier le signe de la dérivée seconde pour les solutions.

Problème ND : extrémaliser $F(\mathbf{y})$ équivaut à résoudre les N équations algébriques $\frac{\partial F}{\partial y_i} = 0$ puis étudier le signe des valeurs propres de matrice hessienne $\frac{\partial^2 F}{\partial x_i \partial x_j}$ pour les solutions.

Problème ∞ D : extrémaliser $F[y]$ équivaut à résoudre l'équation différentielle $\frac{\delta F}{\delta y(x)} = 0$ et on discute rarement le signe de la dérivée seconde.

Nous devons donc donner un sens à cette dérivée fonctionnelle $\frac{\delta F}{\delta y(x)}$.

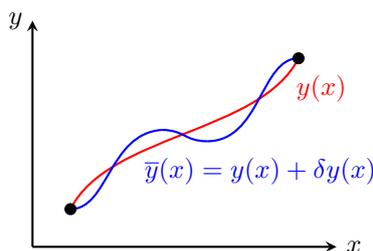


FIGURE 3.2. – La courbe $y(x)$ en noir représente la courbe qui donne un extremum de la fonctionnelle $F[y(x)]$. On effectue une petite variation de cette courbe en définissant la courbe $\bar{y}(x) = y(x) + \delta y(x)$

Définition de la dérivée fonctionnelle : On recherche les fonctions y qui rendent stationnaire $F[y]$. On effectue une variation autour de $y \in \mathcal{Y}$ définie par $y(x) + \alpha\eta(x) \in \mathcal{Y}$ avec α un réel qu'on fait tendre vers zéro et $\eta(x)$ une fonction telle que $\eta(a) = \eta(b) = 0$ pour que la variation ne change pas les conditions aux limites. Alors, on définit la dérivée fonctionnelle $\frac{\delta F}{\delta y(x)} : x \in \mathbb{R} \mapsto \mathbb{R}$ (fonction réelle de x) si l'on peut écrire pour $\alpha \rightarrow 0^+$

$$F[y + \alpha\eta] - F[y] = \alpha \int_a^b dx \frac{\delta F}{\delta y(x)} \eta(x) + \mathcal{O}(\alpha^2) \quad (3.7)$$

En physique, on utilise une notation plus compacte en introduisant $\delta y(x) = \alpha\eta(x)$ et $\delta F = F[y + \alpha\eta] - F[y]$

$$\boxed{\delta F = \int_a^b dx \frac{\delta F}{\delta y(x)} \delta y(x)} \quad (3.8)$$

Cette définition permet de calculer en pratique la dérivée fonctionnelle en cherchant à factoriser δy dans l'intégrale et identifier le reste de l'intégrand avec la dérivée fonctionnelle. On remarque que du point de vue dimensionnel, $\frac{\delta F}{\delta y(x)}$ n'a pas la dimension $[F]/[y]$ comme la notation pourrait le laisser penser. Le problème d'extrémalisation $\forall \delta y, \delta F = 0$ équivaut donc à résoudre l'équation différentielle

$$\boxed{\frac{\delta F}{\delta y(x)} = 0} . \quad (3.9)$$

Interprétation intuitive mais non rigoureuse : Une manière simple de comprendre la formule comme un passage d'un indice discret i à l'indice continu x est de considérer le développement de Taylor au premier ordre en α

$$F(y_1 + \alpha\eta_1, \dots, y_N + \alpha\eta_N) \simeq F(y_1, \dots, y_N) + \alpha \sum_{i=1}^N \frac{\partial F}{\partial y_i} \eta_i \quad (3.10)$$

$$\longrightarrow F[y + \alpha\eta] \simeq F[y] + \alpha \int_a^b dx \frac{\delta F}{\delta y(x)} \eta(x) \quad (3.11)$$

et voir ainsi la dérivée fonctionnelle comme généralisation des dérivées partielles.

4. Formule d'Euler

Dans le cas où la fonctionnelle s'écrit sous la forme

$$F[y] = \int_a^b dx f(y(x), y'(x), x), \quad (3.12)$$

c'est-à-dire que f ne fait intervenir qu'au plus y' dans son expression, et avec conditions aux limites fixées en a et b , alors la dérivée fonctionnelle se calcule avec

$$\boxed{\frac{\delta F}{\delta y(x)} = \frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right)}. \quad (3.13)$$

en pensant à traiter y , y' et x comme trois variables indépendantes dans les dérivées partielles.

La démonstration part de la définition

$$\delta F = \int_a^b dx \{ f(y(x) + \delta y(x), y'(x) + \delta y'(x), x) - f(y, y', x) \} \quad (3.14)$$

$$= \int_a^b dx \left\{ \frac{\partial f}{\partial y}(x) \delta y(x) + \frac{\partial f}{\partial y'}(x) \delta y'(x) \right\} = \int_a^b dx \left\{ \frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial y'} \frac{d}{dx} \delta y \right\}, \quad (3.15)$$

où nous avons utilisé la commutativité des variations

$$\delta y' = \delta \frac{dy}{dx} = \frac{d}{dx} \delta y. \quad (3.16)$$

En intégrant par partie cette dernière équation on obtient

$$\delta F = \int_a^b dx \left\{ \frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} \right\} \delta y + \left[\frac{\partial f}{\partial y'} \delta y \right]_a^b. \quad (3.17)$$

Le terme de bord s'annule car on a $\delta y(a) = \delta y(b) = 0$. D'où le résultat.

Dès lors, la condition de stationnarité de la fonctionnelle (3.9) peut se réécrire

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) = 0. \quad (3.18)$$

C'est l'équation d'Euler dont on remarque qu'elle a exactement la même forme que les équations de Lagrange! Nous les appellerons désormais **les équations d'Euler-Lagrange**.

Généralisations : on généralise sans difficulté

- (i) au cas d'une dépendance fonctionnelle en plusieurs fonctions. Si $F[y(x), z(x)]$ et $f(y, y', z, z', x)$, on aura deux dérivées fonctionnelles par rapport à y et z à annuler pour assurer la condition de stationnarité $\delta F = 0$:

$$\frac{\delta F}{\delta y(x)} = \frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) = 0 \quad (3.19)$$

$$\frac{\delta F}{\delta z(x)} = \frac{\partial f}{\partial z} - \frac{d}{dx} \left(\frac{\partial f}{\partial z'} \right) = 0 \quad (3.20)$$

et ainsi de suite si l'on augmente le nombre de "fonctions variables".

- (ii) aux cas d'ordres supérieurs en dérivées. Si l'on a $f(y, y', y'', x)$, il vient par le même raisonnement

$$\frac{\delta F}{\delta y(x)} = \frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) + \frac{d^2}{dx^2} \left(\frac{\partial f}{\partial y''} \right). \quad (3.21)$$

5. Principe de moindre action de Hamilton

Nous devinons maintenant que nous pouvons re-dériver les équations de la mécanique d'Euler-Lagrange à partir d'un principe variationnel, c'est le principe de Hamilton. Soit un système décrit par n coordonnées généralisées $\{q_\alpha, \dot{q}_\alpha\}_{\alpha=1, \dots, n}$ pour lesquelles on omettra les accolades dans les variables pour alléger les notations comme précédemment. Le principe stipule que la trajectoire réalisée physiquement est celle qui minimise l'**action** $S[q_\alpha(t)]$ définie par

$$S[q_\alpha] = \int_{t_1}^{t_2} dt \mathcal{L}(q_\alpha(t), \dot{q}_\alpha(t), t). \quad (3.22)$$

La minimisation se fait sur l'ensemble des chemins $q_\alpha(t)$ passant aux mêmes positions aux instants t_1 et t_2 . Il est clair qu'il redonne les équations d'Euler-Lagrange

$$\forall \alpha, \quad \frac{\delta S}{\delta q_\alpha(t)} = 0 \quad \Leftrightarrow \quad \frac{\partial \mathcal{L}}{\partial q_\alpha} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) = 0 \quad (3.23)$$

C'est un principe global qui intègre l'ensemble de la trajectoire dont on tire des équations locales pour le mouvement. Une fois ces équations obtenues, elles sont résolues comme d'habitude à partir des conditions initiales plutôt que des conditions aux bords en t_1 et t_2 dont on a fait l'usage par commodité. Le point important est le changement de point de vue pour recherche plus généralement les lois de la physique : plutôt que de rechercher les équations du mouvement satisfaites à tout temps, on recherche sur la base de grands principes quel lagrangien pourrait décrire le système. Les équations du mouvement sont ensuite obtenues directement par le principe de Hamilton.

II. Prise en compte des contraintes par des multiplicateurs de Lagrange

1. Principe : théorème des extrêma liés

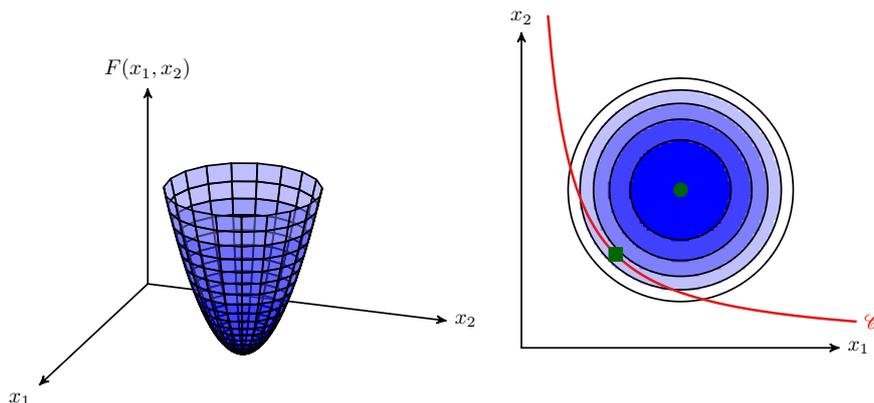


FIGURE 3.3. – Représentations d'une fonction à plusieurs variable possédant un extremum. Le cercle vert indique la position de l'extremum libre de la fonction et le carré vert indique la position du l'extremum lié de la fonction le long de la courbe \mathcal{C} représentée en rouge.

Comme représenté sur la figure 3.3, on cherche à minimiser une fonction à deux variables $F(x_1, x_2)$ sous une contrainte holonome $f(x_1, x_2) = 0$ qui détermine une courbe \mathcal{C} dans le plan (x_1, x_2) . Nous allons montrer que cette minimisation (extrémalisation) sous contrainte peut être résolue en introduisant **un multiplicateur de Lagrange** $\lambda \in \mathbb{R}$ comme variable de la nouvelle fonction

$$\bar{F}(x_1, x_2, \lambda) = F(x_1, x_2) + \lambda f(x_1, x_2) \quad (3.24)$$

puis en minimisant cette fonction \bar{F} comme un problème libre, c'est-à-dire en annulant ses trois dérivées partielles.

Pour le montrer, écrivons les trois équations d'extrémalisation de $\bar{F}(x_1, x_2, \lambda)$

$$\frac{\partial \bar{F}}{\partial x_1} = \frac{\partial F}{\partial x_1}(x_1, x_2) + \lambda \frac{\partial f}{\partial x_1}(x_1, x_2) = 0 \quad (3.25)$$

$$\frac{\partial \bar{F}}{\partial x_2} = \frac{\partial F}{\partial x_2}(x_1, x_2) + \lambda \frac{\partial f}{\partial x_2}(x_1, x_2) = 0 \quad (3.26)$$

$$\frac{\partial \bar{F}}{\partial \lambda} = f(x_1, x_2) = 0 \quad (3.27)$$

qui donnent trois équations pour les trois inconnues (x_1, x_2, λ) .

Montrons qu'elles sont équivalentes aux équations tirées de la position du problème. Soit $(\delta x_1, \delta x_2)$ un déplacement virtuel qui satisfait donc la contrainte $f = 0$. On a donc d'après (2.3)

$$\frac{\partial f}{\partial x_1} \delta x_1 + \frac{\partial f}{\partial x_2} \delta x_2 = 0 \quad (3.28)$$

De plus, F doit être extrême au voisinage de l'extremum contraint (x_1^*, x_2^*) de sorte que la différentielle, écrite avec les δx_i , prend la forme

$$\delta F = \frac{\partial F}{\partial x_1} \delta x_1 + \frac{\partial F}{\partial x_2} \delta x_2 = 0 \quad (3.29)$$

En considérant le ratio $\delta x_2/\delta x_1$ ou le déterminant du système de deux équations qui doit s'annuler pour une solution non nulle, on trouve que les coordonnées de l'extremum contraint doivent satisfaire, d'une part $f(x_1^*, x_2^*) = 0$ (donnée par (3.27)) et d'autre part

$$\left(\frac{\partial F}{\partial x_1} \right) \left(\frac{\partial f}{\partial x_2} \right) = \left(\frac{\partial F}{\partial x_2} \right) \left(\frac{\partial f}{\partial x_1} \right). \quad (3.30)$$

On retrouve bien cette équation en éliminant λ des équations (3.25) et (3.26). Ces équations, tout comme (3.28) et (3.29) ont une interprétation physique intuitive : l'extremum contraint se trouve aux endroits où les « forces » associées aux « potentiels » F et f sont alignées et se compensent. Il existe un λ tel que

$$\nabla F + \lambda \nabla f = \mathbf{0} \quad (3.31)$$

Une autre manière d'obtenir ce résultat et qui se généralise plus facilement correspond à additionner (3.28) et (3.29) en multipliant (3.28) par λ arbitraire. Il vient

$$\forall \delta x_1, \delta x_2, \quad \left(\frac{\partial F}{\partial x_1} + \lambda \frac{\partial f}{\partial x_1} \right) \delta x_1 + \left(\frac{\partial F}{\partial x_2} + \lambda \frac{\partial f}{\partial x_2} \right) \delta x_2 = 0 \quad (3.32)$$

mais δx_1 et δx_2 sont liés par la contrainte. On choisit alors λ de sorte que $\frac{\partial F}{\partial x_1} + \lambda \frac{\partial f}{\partial x_1} = 0$, soit l'équation (3.25), puis il reste

$$\forall \delta x_2, \quad \left(\frac{\partial F}{\partial x_2} + \lambda \frac{\partial f}{\partial x_2} \right) \delta x_2 = 0 \quad (3.33)$$

qui redonne (3.26) comme δx_2 est un paramètre libre.

Cas général : On minimise une fonction $F(x_1, \dots, x_n)$ sous les $K < n$ contraintes holonomes $f_\beta(x_1, \dots, x_n)$ en introduisant K multiplicateurs de Lagrange λ_β (un par contrainte) et en minimisant « sans » contrainte la fonction

$$\bar{F}(x_1, \dots, x_n, \lambda_1, \dots, \lambda_K) = F(x_1, \dots, x_n) + \sum_{\beta=1}^K \lambda_\beta f_\beta(x_1, \dots, x_n) \quad (3.34)$$

par rapport à ses $n + K$ variables.

Complément : preuve du cas général – Soit $F(x_1, \dots, x_n)$ une fonction continue et dérivable de n variables $x_i, i = 1, \dots, n$. La recherche d'un point stationnaire le long d'un déplacement virtuel donne

$$\delta F = \sum_i \frac{\partial F}{\partial x_i} \delta x_i = 0, \quad \forall \delta x_i. \quad (3.35)$$

En présence d'une contrainte $f_\beta(x_1, x_2, \dots, x_n) = 0$, nous aurons de plus

$$\delta f_\beta = \sum_i \frac{\partial f_\beta}{\partial x_i} \delta x_i = 0, \quad (3.36)$$

Nous combinons les conditions d'extremum et les contraintes (3.35) et (3.36) en introduisant K multiplicateurs de Lagrange λ_β

$$\delta F + \sum_\beta \lambda_\beta \delta f_\beta = \sum_{i=1}^n \left(\frac{\partial F}{\partial x_i} + \sum_\beta \lambda_\beta \frac{\partial f_\beta}{\partial x_i} \right) \delta x_i = \sum_{i=1}^n \frac{\partial \bar{F}}{\partial x_i} \delta x_i = 0, \quad (3.37)$$

Cela est vrai pour tous λ_β mais les δx_i ne sont pas indépendants. On choisit de fixer les valeurs des λ_β de sorte d'annuler successivement

$$\frac{\partial \bar{F}}{\partial x_n} = 0; \quad \frac{\partial \bar{F}}{\partial x_{n-1}} = 0; \quad \dots \quad \frac{\partial \bar{F}}{\partial x_{n-K+1}} = 0; \quad (3.38)$$

Restent alors dans la somme (3.37)

$$\sum_{i=1}^{n-K} \frac{\partial \bar{F}}{\partial x_i} \delta x_i = 0, \quad \forall \delta x_i \quad (3.39)$$

dans laquelle les $n - K$ variables δx_i sont maintenant indépendantes de sorte que l'on doit avoir

$$\frac{\partial \bar{F}}{\partial x_i} = 0, \quad i = 1, \dots, n - K. \quad (3.40)$$

La combinaison des équations (3.38) et (3.40) correspond bien à la minimisation de \bar{F} selon l'ensemble de variables x_i tandis que sa minimisation par rapport aux λ_β redonne les contraintes qui doivent être également satisfaites. On a donc $n + K$ équations pour $n + K$ inconnues et un problème bien posé.

2. Principe de moindre action de Hamilton en présence de contraintes

Nous avons introduit les coordonnées généralisées telles qu'elles permettent de décrire le système en incluant directement l'ensemble des contraintes holonomes. Toutefois il est parfois souhaitable pour des raisons notamment pratiques de ne pas inclure certaines contraintes dans le jeu de coordonnées choisies et de les inclure via des multiplicateurs de Lagrange. Pour comprendre la nature du multiplicateur de Lagrange dans ce cas, prenons un exemple à une coordonnées q et une contrainte holonome $f(q, t) = 0$. La passage de la minimisation d'une fonction de plusieurs variables F à la minimisation d'une fonctionnelle S correspond à passer d'une équation algébrique en x à une équation différentielle en temps. Schématiquement, l'équation de minimisation de transforme selon

$$\frac{\partial F}{\partial x} + \lambda \frac{\partial f}{\partial x} = 0 \quad \longrightarrow \quad \frac{\delta S}{\delta q(t)} + \lambda(t) \frac{\delta S_f}{\delta q(t)} = 0 \quad (3.41)$$

avec $S_f[q] = \int_{t_1}^{t_2} f(q, t) dt$ un terme d'action qui insert la contrainte. Comme l'équation différentielle est paramétrée par le temps t , le multiplicateur de Lagrange doit, dans ce cas, être une fonction du temps $\lambda(t)$.

Généralisons en écrivant q_α un jeu de n coordonnées généralisées et $f_\beta(q_\alpha, t) = 0$ un ensemble de K contraintes holonomes. Le nombre de degrés de liberté est donc $g = n - K$. Le système est décrit par la fonction de Lagrange $\mathcal{L}(q_\alpha, \dot{q}_\alpha, t)$ associée à l'action

$$S[q_\alpha] = \int_{t_1}^{t_2} dt \mathcal{L}(q_\alpha, \dot{q}_\alpha, t) \quad (3.42)$$

Le principe de moindre action sous contrainte se définit en considérant que la trajectoire suivie doit minimiser l'action \bar{S} définie par

$$\bar{S}[q_\alpha, \lambda_\beta] = \int_{t_1}^{t_2} dt \bar{\mathcal{L}}(q_\alpha, \dot{q}_\alpha, t, \lambda_\beta), \quad (3.43)$$

avec

$$\bar{\mathcal{L}}(q_\alpha, \dot{q}_\alpha, t, \lambda_\beta) = \mathcal{L}(q_\alpha, \dot{q}_\alpha, t) + \sum_{\beta=1}^K \lambda_\beta(t) f_\beta(q_\alpha, t). \quad (3.44)$$

Au final, il y a $n + K$ fonctions inconnues $(q_\alpha, \lambda_\beta)$ pour $n + K$ équations

$$\frac{d}{dt} \left(\frac{\partial \bar{\mathcal{L}}}{\partial \dot{q}_\alpha} \right) = \frac{\partial \bar{\mathcal{L}}}{\partial q_\alpha} \quad \alpha = 1, \dots, n, \quad (3.45)$$

$$f_\beta(q_\alpha, t) = 0, \quad \beta = 1, \dots, K. \quad (3.46)$$

Interprétation physique des multiplicateurs de Lagrange – Plus qu'un simple intermédiaire de calcul, ils s'interprètent physiquement comme les coefficients de forces de réactions généralisées associées aux contraintes. En notant $S_{f_\beta}[q_\alpha] = \int_{t_1}^{t_2} dt f_\beta(q_\alpha, t)$ et en remarquant que pour des contraintes holonomes

$$\frac{\delta S_{f_\beta}}{\delta q_\alpha(t)} = \frac{\partial f_\beta}{\partial q_\alpha} \quad (3.47)$$

par la formule d'Euler, on peut mettre les équations (3.45) sous la forme

$$\frac{d}{dt} \left(\frac{\partial \bar{\mathcal{L}}}{\partial \dot{q}_\alpha} \right) - \frac{\partial \bar{\mathcal{L}}}{\partial q_\alpha} = \sum_{\beta} Q_{\alpha\beta} \quad (3.48)$$

où les $Q_{\alpha\beta}$ sont des forces de réactions généralisées

$$Q_{\alpha\beta} = \lambda_\beta(t) \frac{\partial f_\beta}{\partial q_\alpha}(q_\alpha, t). \quad (3.49)$$

qui généralisent la discussion de la section I du chapitre 1.

3. Exemple : Le pendule plan

Nous allons chercher ici à calculer la force de tension du fil dans le cas du pendule plan (Fig. 2.2a). Pour cela, nous prenons un système de coordonnées ne respectant pas la contrainte

et nous allons l'imposer via un multiplicateur de Lagrange. En utilisant les coordonnées polaires, la fonction de Lagrange s'écrit

$$\mathcal{L}(r, \dot{r}, \theta, \dot{\theta}) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + mgr \cos \theta. \quad (3.50)$$

et la contrainte $f(r, \theta) = r - \ell = 0$. La fonction de Lagrange modifiée prend le forme $\overline{\mathcal{L}} = \mathcal{L} + \lambda(r - \ell)$ et sa minimisation donne les trois équations

$$m(\ddot{r} - r\dot{\theta}^2) - mg \cos \theta = \lambda(t) \quad (3.51)$$

$$\frac{d}{dt}(r^2\dot{\theta}) + gr \sin \theta = 0 \quad (3.52)$$

$$r - \ell = 0 \quad (3.53)$$

Les Eqs. (3.52) et (3.53) donnent l'équation du pendule

$$\ddot{\theta} + \frac{g}{\ell} \sin \theta = 0. \quad (3.54)$$

qui s'intègrent pour donner la conservation de l'énergie mécanique

$$E_m = \frac{1}{2}m\ell^2\dot{\theta}^2 - mg\ell \cos \theta. \quad (3.55)$$

Les Eqs. 3.51 et (3.53) donnent l'expression du multiplicateur de Lagrange

$$\lambda(t) = -mg \cos \theta(t) - m\ell\dot{\theta}^2(t) = -\frac{2E}{\ell} - 3mg \cos \theta(t). \quad (3.56)$$

La tension du fil est simplement donnée par le gradient de la contrainte.

$$\mathbf{T} = \lambda(t)\nabla(r - \ell) = \lambda(t)\mathbf{e}_r. \quad (3.57)$$

Le multiplicateur de Lagrange décrit donc bien l'intensité de la force qui est orientée le long du fil. Si le fil est rigide, $\lambda(t)$ peut prendre des valeurs positives ou négatives. Si le fil est non-rigide, la contrainte est non-holonome $r \leq \ell$ et le multiplicateur de Lagrange ne peut pas prendre de valeur positive (on suppose ici que le fil ne transmet de tension que s'il est tendu, c'est-à-dire $r = \ell$). On peut donc déduire de ce raisonnement l'angle maximum du pendule $\cos \theta_0 = -2E/(3mg\ell)$ au-delà duquel la masse va entrer en chute libre (avant de remettre le fil en tension).

C1. . Solution du problème du brachistochrone

Pour simplifier le problème du brachistochrone, on introduit le repère Oxy tel que la courbe $y(x)$ commence en O et tel que l'axe Oy soit orienté vers le bas. À un facteur près, le temps de parcours s'écrit

$$T[y] = \int_0^a dx \sqrt{\frac{1 + y'^2}{y}}. \quad (3.58)$$

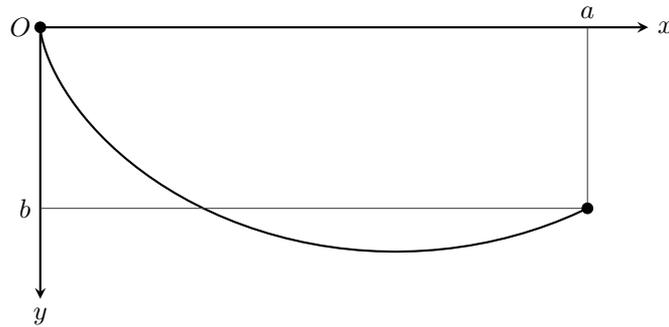


FIGURE 3.4. – La solution du problème du brachistochrone est la courbe cycloïde.

La densité de la fonctionnelle est donc $f(y, y') = \sqrt{\frac{1+y'^2}{y}}$. On peut en principe résoudre directement les équations d'Euler-Lagrange pour ce système. Toutefois, puisque l'intégrand est indépendant de la position x , on peut montrer que la quantité

$$E = y' \frac{\partial f}{\partial y'} - f. \quad (3.59)$$

reste constante en montrant que $\frac{dE}{dx} = 0$. Il s'agit de **l'identité de Beltrami** que l'on retrouvera dans la définition de l'énergie lagrangienne au chapitre 3. Après quelques dérivations simples on obtient

$$y(1 + y'^2) = D. \quad (3.60)$$

où $D = 1/E^2$. On introduit maintenant le changement de variable $y \rightarrow \theta$ défini par

$$y' = \frac{dy}{dx} = \cot\left(\frac{\theta}{2}\right), \quad (3.61)$$

ce qui donne

$$y(\theta) = \frac{D}{2}(1 - \cos \theta). \quad (3.62)$$

On montre également facilement que la différentielle dx s'écrit

$$dx = \frac{D}{2}(1 - \cos \theta) d\theta \quad (3.63)$$

que l'on peut facilement intégrer. Pour finir on trouve que la courbe brachistochrone sous une forme paramétrée est la courbe cycloïde

$$x(\theta) = \frac{D}{2}(\theta - \sin \theta) \quad (3.64)$$

$$y(\theta) = \frac{D}{2}(1 - \cos \theta) \quad (3.65)$$

La figure 3.4 représente la solution cycloïde du brachistochrone pour une courbe passant par l'origine et passant par le point (a, b) du plan Oxy .

4. Formulation lagrangienne - Grandeurs physiques

Ce chapitre discute de grandeurs physiques apparaissant dans le formalisme lagrangien et généralisant ceux de la mécanique newtonienne (impulsion, énergie,...). On y étudie les invariances du lagrangien sous des transformations (jauge, changement de référentiels) et enfin le rôle des symétries sur la structure et la conservation de grandeurs physiques.

I. Moment généralisé

1. Définition

Jusqu'ici nous avons introduit le concept de coordonnées et de vitesses généralisés. Le moment généralisé associé à la coordonnée q_α est défini par

$$p_\alpha = \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \quad (4.1)$$

Exemples

Particule dans un potentiel conservatif. – Si l'on a $\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) = \frac{1}{2}m\dot{\mathbf{r}}^2 - U(\mathbf{r})$, on retrouve $\mathbf{p} = m\dot{\mathbf{r}}$ l'impulsion ou quantité de mouvement de la mécanique newtonienne.

Coordonnées polaires – on trouve $p_r = m\dot{r}$ et $p_\theta = mr^2\dot{\theta}$ qui correspond au moment cinétique selon z (voir page 34).

Particule chargée – Cette fois, le potentiel généralisé fait apparaître une vitesse généralisée. On a montré $\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) = \frac{1}{2}m\dot{\mathbf{r}}^2 + q(\dot{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r}, t) - V(\mathbf{r}, t))$, qui conduit à

$$\mathbf{p} = m\dot{\mathbf{r}} + q\mathbf{A} \neq m\dot{\mathbf{r}}. \quad (4.2)$$

On voit ici, qu'en raison du couplage avec le champ, le moment n'est pas égal à la quantité de mouvement.

2. Coordonnées cycliques

Lorsque la fonction de Lagrange ne dépend pas explicitement d'une coordonnée q_{α_0} , c'est-à-dire

$$\frac{\partial \mathcal{L}}{\partial q_{\alpha_0}} = 0, \quad (4.3)$$

on parle alors de coordonnée cyclique. Dans ce cas, la dynamique le long de cette coordonnée est résolue simplement. Les équations d'Euler-Lagrange donnent immédiatement que $\dot{p}_{\alpha_0} = 0$, soit

$$p_{\alpha_0}(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t) = \text{cste} \quad (4.4)$$

qui est une équation du premier ordre par rapport au temps. Le moment généralisé p_{α_0} est une grandeur conservée et cette équation fournit une **intégrale première** du mouvement.

Exemple : problème à force centrale – Pour un lagrangien en coordonnées polaires de la forme

$$\mathcal{L}(r, \dot{r}, \theta, \dot{\theta}) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - U(r), \quad (4.5)$$

la coordonnée θ est une coordonnée cyclique. Le moment conjugué est conservé

$$p_{\theta} = mr^2\dot{\theta} = \text{cste} \equiv L_z \quad (4.6)$$

Cette constante du mouvement correspond à la conservation du moment cinétique. On peut réinjecter ce résultat (la seconde équation du mouvement) dans l'équation du mouvement selon r

$$m(\ddot{r} - r\dot{\theta}^2) = -\frac{\partial U}{\partial r} \quad \longrightarrow \quad m\ddot{r} = -\frac{\partial U_{\text{eff}}}{\partial r} \quad \text{avec} \quad U_{\text{eff}}(r) = U(r) + \frac{L_z^2}{2mr^2} \quad (4.7)$$

le second terme étant le potentiel centrifuge.

Remarque : attention à ne pas réinjecter cette loi de conservation directement dans le lagrangien, cela ne conduirait pas aux bonnes équations du mouvement (voir complément ci-dessous). Transformer le lagrangien pour qu'il dépende de p_{θ} et plus de $\dot{\theta}$ se fera à l'aide d'une transformée de Legendre dans le formalisme hamiltonien.

Complément : réécriture du lagrangien avec une coordonnée cyclique – On choisit l'indice n pour la coordonnée cyclique par simplicité. Puisque la fonction de Lagrange ne dépend pas de q_n on a

$$p_n = p_n(q_1, \dots, q_{n-1}, \dot{q}_1, \dots, \dot{q}_{n-1}, \dot{q}_n, t) = c_n. \quad (4.8)$$

À partir de l'expression de p_n on peut donc en principe exprimer \dot{q}_n en fonction des autres coordonnées et vitesses.

$$\dot{q}_n = f(q_1, \dots, q_{n-1}, \dot{q}_1, \dots, \dot{q}_{n-1}, c_n, t). \quad (4.9)$$

Cette équation s'intègre formellement en

$$q_n = \int dt f(q_1, \dots, q_{n-1}, \dot{q}_1, \dots, \dot{q}_{n-1}, c_n, t). \quad (4.10)$$

Ainsi q_n peut-être éliminé du problème. Il suffit alors de résoudre les équations d'Euler-Lagrange pour les autres degrés de liberté en remplaçant \dot{q}_n par son expression formelle (Eq. 4.9) et on en déduit alors la coordonnée q_n par intégration. On peut toutefois se demander s'il ne serait pas possible d'éliminer la coordonnée q_n avant de dériver les équations d'Euler-Lagrange directement dans un principe variationnel. Cette approche est possible mais il faut toutefois modifier ce principe. En effet, dans le principe variationnel, nous avons supposé que les déplacements virtuels étaient indépendants. Le fait que le moment conjugué soit constant impose une contrainte entre q_n et les autres coordonnées Eq. (4.10). On définit ainsi la variation δ' comme la variation élémentaire restreinte aux trajectoires telles que $p_n = c_n = \text{cste}$. La variation de l'action donne

$$\delta' S = \int_{t_1}^{t_2} \sum_{\alpha=1}^n \left(\frac{\partial \mathcal{L}}{\partial q_\alpha} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) \right) \delta' q_\alpha dt + \sum_{\alpha=1}^n \left[\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \delta' q_\alpha \right]_{t_1}^{t_2}. \quad (4.11)$$

Le premier terme est nul car la solution du problème doit nécessairement respecter les équations d'Euler-Lagrange. Le second terme est le terme de bord. Pour $\alpha \neq n$, on impose que les variations aux temps t_1 et t_2 est nul $\delta' q_\alpha(t_1) = \delta' q_\alpha(t_2) = 0$. Ceci n'est cependant pas nécessairement vrai puisqu'il existe une relation entre q_n et les autres coordonnées. Ainsi, comme $p_n = \text{cste}$, la variation de l'action s'écrit

$$\delta' S = \delta' [p_n q_n]_{t_1}^{t_2} = \delta' \int_{t_1}^{t_2} p_n \dot{q}_n dt. \quad (4.12)$$

Nous allons faire passer le terme de droite de l'équation du coté gauche et on obtient

$$\delta' \left[\int_{t_1}^{t_2} \mathcal{L} - p_n \dot{q}_n \right] = 0 \quad (4.13)$$

Dans le cas d'une coordonnée cyclique, on peut donc utiliser un principe variationnel modifié et basé sur la fonction de Lagrange modifiée $\mathcal{L}' = \mathcal{L} - p_n \dot{q}_n$, telle que $\delta' S' = 0$, avec $S' = \int_{t_1}^{t_2} \mathcal{L}' dt$. Il suffit alors de résoudre les équations d'Euler-Lagrange avec cette fonction de Lagrange modifiée en prenant soin d'avoir éliminer la variable \dot{q}_n de l'expression. Ce problème est plus simple car la coordonnée q_n a été complètement éliminée avant de dériver les équations du mouvement.

Retour sur le problème à force centrale – Une erreur serait d'exprimer directement θ en fonction de p_θ dans l'expression de la fonction de Lagrange et de dériver les équations d'Euler-Lagrange. On peut éliminer la coordonnée θ en utilisant la fonction de Lagrange modifiée puis en remplaçant $\dot{\theta}$ en fonction de p_θ

$$\mathcal{L}'(r, \dot{r}) = \frac{1}{2} m \dot{r}^2 + \frac{p_\theta^2}{2mr^2} - U(r) - p_\theta \frac{p_\theta}{mr^2} = \frac{1}{2} m \dot{r}^2 - U_{\text{eff}}(r) \quad (4.14)$$

On est ramené à un problème à 1 degré de liberté.

II. Énergie

1. Théorème d'Euler sur les fonctions homogènes

Pour une fonction de n variables $F(x_1, \dots, x_n)$, on dit qu'elle est homogène de degré k si elle satisfait à la relation

$$\forall \lambda > 0, \quad F(\lambda x_1, \dots, \lambda x_n) = \lambda^k F(x_1, \dots, x_n). \quad (4.15)$$

En dérivant cette propriété par rapport à λ et en faisant $\lambda = 1$ dans le résultat, on trouve facilement la relation d'Euler

$$\sum_{i=1}^n x_i \frac{\partial F}{\partial x_i}(x_1, \dots, x_n) = k F(x_1, \dots, x_n) \quad (4.16)$$

2. Définitions

Énergie cinétique – On rappelle la forme générale $T = \sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i^2$ ainsi que le cas où la transformation $\mathbf{r}_i \rightarrow q_\alpha$ est indépendante du temps (2.44)

$$T = \sum_{\alpha, \beta} \frac{1}{2} M_{\alpha\beta}(q_\alpha) \dot{q}_\alpha \dot{q}_\beta \quad \text{telle que} \quad \sum_\alpha \dot{q}_\alpha \frac{\partial T}{\partial \dot{q}_\alpha} = 2T \quad (4.17)$$

pour T est homogène de degré 2 par rapport aux vitesses généralisées.

Énergie mécanique – On suit la définition utilisée en mécanique newtonienne pour un système conservatif :

$$E_m(q_\alpha, \dot{q}_\alpha, t) = T(q_\alpha, \dot{q}_\alpha, t) + U(q_\alpha, \dot{q}_\alpha, t) . \quad (4.18)$$

Énergie lagrangienne – on introduit la définition

$$\boxed{E(q_\alpha, \dot{q}_\alpha, t) = \sum_\alpha \dot{q}_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} - \mathcal{L}} \quad (4.19)$$

3. Dépendance en temps et conservation de l'énergie

La définition de l'énergie lagrangienne est E motivée par la considération suivante :

$$\frac{d\mathcal{L}}{dt} = \frac{\partial \mathcal{L}}{\partial t} + \sum_\alpha \dot{q}_\alpha \frac{\partial \mathcal{L}}{\partial q_\alpha} + \sum_\alpha \ddot{q}_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} = \frac{\partial \mathcal{L}}{\partial t} + \sum_\alpha \dot{q}_\alpha \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) + \sum_\alpha \ddot{q}_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \quad (4.20)$$

$$= \frac{\partial \mathcal{L}}{\partial t} + \frac{d}{dt} \left(\sum_\alpha \dot{q}_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) \quad (4.21)$$

dont on tire immédiatement la relation générale

$$\boxed{\frac{dE}{dt} = - \frac{\partial \mathcal{L}}{\partial t}} \quad (4.22)$$

Ainsi pour que l'énergie E soit conservée il faut et il suffit que la fonction de Lagrange ne dépende pas explicitement du temps. Il est important de constater qu'en général $\boxed{E \neq E_m}$.

Système conservatif et indépendant du temps – Dans ce cas particulier mais courant, on retrouve en combinant (4.17) et (4.19) :

$$E = \sum_\alpha \dot{q}_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} - \mathcal{L} = 2T - (T - U) = T + U = E_m . \quad (4.23)$$

Exemple : perle glissant le long d'un cerceau en rotation (Fig. 2.5) – Nous avons trouvé le lagrangien $\mathcal{L}(\theta, \dot{\theta}, t) = \frac{m\ell^2}{2}(\dot{\theta}^2 + \dot{\varphi}^2(t) \sin^2 \theta) - mgl \cos \theta$. L'énergie mécanique vaut donc

$$E_m(\theta, \dot{\theta}, t) = \frac{m\ell^2}{2}(\dot{\theta}^2 + \dot{\varphi}^2(t) \sin^2 \theta) + mgl \cos \theta . \quad (4.24)$$

L'énergie lagrangienne se calcule en prenant la définition

$$E(\theta, \dot{\theta}, t) = \dot{\theta} \frac{\partial \mathcal{L}}{\partial \dot{\theta}} - \mathcal{L} = E_m - m\ell^2 \dot{\varphi}^2(t) \sin^2 \theta \quad (4.25)$$

Il n'y a donc égalité qu'en l'absence de mouvement $\dot{\varphi} = 0$. Concernant la conservation de l'énergie, on trouve

$$\frac{dE}{dt} = -\frac{\partial \mathcal{L}}{\partial t} = -m\ell^2 \sin^2 \theta \dot{\varphi}(t) \ddot{\varphi}(t) \quad (4.26)$$

de sorte qu'il suffit qu'il y ait $\ddot{\varphi} = 0$ et une vitesse angulaire constante pour avoir conservation de l'énergie E avec $E \neq E_m$ dans ce cas.

III. Invariance de Jauge du Lagrangien

Toute l'information sur la dynamique est contenue dans le lagrangien \mathcal{L} . Cependant, peut-il exister plusieurs Lagrangien conduisant à la même dynamique, c'est-à-dire aux mêmes équations d'Euler-Lagrange? Par exemple, multiplier le lagrangien par une constante ne les change pas. Par convention et par simplicité, le choix pertinent est $\mathcal{L} = T - U$ pour avoir la dimension de l'énergie et un lien direct avec elle lorsque l'on fera la formulation Hamiltonienne pour laquelle ce choix est également le plus pertinent physiquement. Enfin, ce choix d'unité donne à l'action S la même dimension que la constante de Planck réduite \hbar ce qui est également le plus pertinent pour étudier le lien avec la mécanique quantique.

1. Invariance de jauge du lagrangien

Montrons maintenant que si le lagrangien diffère d'une dérivée totale par rapport au temps d'une fonction ne dépendant que des positions et du temps conduit aux mêmes équations du mouvement. Mathématiquement, le lagrangien construit selon

$$\mathcal{L}'(q_\alpha, \dot{q}_\alpha, t) = \mathcal{L}(q_\alpha, \dot{q}_\alpha, t) + \frac{d}{dt} \chi(q_\alpha, t), \quad (4.27)$$

est donc équivalent à \mathcal{L} , avec $\chi(q_\alpha, t)$ une fonction quelconque des coordonnées généralisées et du temps mais indépendantes des vitesses généralisées. Cela se démontre facilement en calculant l'action associée

$$S'[q_\alpha] = \int_{t_1}^{t_2} dt \mathcal{L}' = S[q_\alpha] + \int_{t_1}^{t_2} dt \frac{d}{dt} \chi(q_\alpha, t) = S[q_\alpha] + \chi(q_\alpha(t_2), t_2) - \chi(q_\alpha(t_1), t_1) .$$

Pour une variation $\delta q_\alpha(t)$ telle que $\delta q_\alpha(t_i) = 0$ aux bords, on aura donc $\delta\chi(q_\alpha(t_i), t_i) = \chi(q_\alpha(t_i) + \delta q_\alpha(t_i), t_i) - \chi(q_\alpha(t_i), t_i) = 0$ si bien que $\delta S' = \delta S$ conduisant aux mêmes équations du mouvement. Une démonstration directe en calculant les équations d'Euler-Lagrange sur \mathcal{L}' conduit au même résultat.

Cette « liberté de jauge » permet de simplifier l'écriture des lagrangiens mais aussi voir que deux sont équivalents ou d'identifier des invariances du lagrangien sous des transformations.

2. Transformation de jauge électromagnétique

Rappelons que la transformation de jauge

$$\mathbf{A} \longrightarrow \mathbf{A}' = \mathbf{A} + \nabla\chi \quad \text{et} \quad V \longrightarrow V' = V - \frac{\partial\chi}{\partial t} \quad (4.28)$$

laisse invariante les équations de Maxwell et conduit aux mêmes champs \mathbf{E} et \mathbf{B} . Sous cette transformation, le lagrangien d'une particule chargée $\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) = T + q(\dot{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r}, t) - V(\mathbf{r}, t))$, devient

$$\begin{aligned} \mathcal{L}'(\mathbf{r}, \mathbf{v}, t) &= T + q(\dot{\mathbf{r}} \cdot \mathbf{A}'(\mathbf{r}, t) - V'(\mathbf{r}, t)) \\ &= T + q(\dot{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r}, t) - V(\mathbf{r}, t)) + q\left(\dot{\mathbf{r}} \cdot \nabla\chi(\mathbf{r}, t) + \frac{\partial\chi}{\partial t}(\mathbf{r}, t)\right) \\ &= \mathcal{L}(\mathbf{r}, \mathbf{v}, t) + q\frac{d}{dt}\chi(\mathbf{r}, t) \end{aligned}$$

si bien que l'on trouve que la formulation lagrangienne est invariante sous cette transformation de jauge d'après le résultat de la section précédente.

3. Particule dans un référentiel en translation

Soit \mathcal{R}_0 un référentiel Galiléen et \mathcal{R} un référentiel en translation à la vitesse $\mathbf{V}(t)$, a priori quelconque, par rapport à \mathcal{R}_0 . Soit une particule de masse m et de vitesses $\dot{\mathbf{r}}_0 = \dot{\mathbf{r}} + \mathbf{V}$ par rapport à \mathcal{R}_0 et $\dot{\mathbf{r}}$ par rapport à \mathcal{R} , soumise au potentiel $U(\mathbf{r}) = U(\mathbf{r}_0)$ identique au même point dans les deux référentiels. Pour étudier l'expression du lagrangien sous le changement de repère, on effectue le changement de variables $(\mathbf{r}_0, \dot{\mathbf{r}}_0) \rightarrow (\mathbf{r}, \dot{\mathbf{r}})$, de sorte que

$$\mathcal{L} = \frac{1}{2}m\dot{\mathbf{r}}_0^2 - U(\mathbf{r}_0) = \frac{1}{2}m(\dot{\mathbf{r}} + \mathbf{V})^2 - U(\mathbf{r}) = \frac{1}{2}m\dot{\mathbf{r}}^2 - \mathbf{U}(\mathbf{r}) + m\dot{\mathbf{r}} \cdot \mathbf{V}(t) + \frac{1}{2}m\mathbf{V}^2(t) \quad (4.29)$$

Le dernier terme en \mathbf{V}^2 est indépendant de \mathbf{r} et $\dot{\mathbf{r}}$ et ne modifiera donc pas les équations d'Euler-Lagrange¹. Il peut donc être éliminé et une dernière réécriture permet d'avoir

$$\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) = \frac{1}{2}m\dot{\mathbf{r}}^2 - U(\mathbf{r}) - m\dot{\mathbf{r}} \cdot \frac{d\mathbf{V}}{dt}(t) + \frac{d}{dt}(m\dot{\mathbf{r}} \cdot \mathbf{V}(t)) \quad (4.30)$$

1. Il peut sinon être éliminé par liberté de jauge en l'écrivant $\frac{d}{dt} \int^t du \frac{1}{2}m\mathbf{V}^2(u)$.

Le dernier terme disparaît en invoquant la liberté du choix de jauge et il reste finalement

$$\mathcal{L}(\mathbf{r}, \mathbf{v}, t) = \frac{1}{2}m\mathbf{v}^2 - U_{\text{eff}}(\mathbf{r}, t), \quad \text{avec} \quad U_{\text{eff}}(\mathbf{r}, t) = U(\mathbf{r}) + m\mathbf{r} \cdot \frac{d\mathbf{V}}{dt}(t) \quad (4.31)$$

où le potentiel effectif inclut la contribution de la force d'inertie $-m\frac{d\mathbf{V}}{dt}(t)$ associée à l'accélération du référentiel \mathcal{R} par rapport à \mathcal{R}_0 . Au passage, dans le cas \mathbf{V} constante, on obtient que ce lagrangien non-relativiste est invariant sous les transformations de Galilée, passages d'un référentiel inertiel à un autre, ce qui est cohérent avec les lois fondamentales de la dynamique.

IV. Rôle des symétries dans la formulation lagrangienne

1. Forme du lagrangien d'une particule libre

On illustre ici la démarche consistant à proposer un lagrangien non pas construit pour retrouver des équations du mouvement obtenues par d'autres approches mais construit sur la base d'arguments généraux qui contiennent souvent les grands principes et les symétries. C'est en général l'approche féconde pour proposer de nouvelles théories. On prend deux exemples dont le premier peut paraître un peu artificiel mais illustratifs pour une particule libre, ce qui revient à discuter la forme du terme d'énergie cinétique.

Mécanique galiléenne – L'espace est homogène et isotrope et le temps absolu. Les lois sont invariantes sous les transformations galiléennes. Comme le temps est absolu, c'est un paramètre indépendant des référentiels et l'action s'écrit naturellement comme une intégrale temporelle, comme fait jusqu'à présent :

$$S[\mathbf{r}] = \int_{t_1}^{t_2} dt \mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) \quad (4.32)$$

Les symétries spatiales et temporelles font que \mathcal{L} ne peut dépendre que du module de la vitesse et on écrira $\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) = L(\dot{\mathbf{r}}^2)$. Comme dans le paragraphe précédent, l'invariance galiléenne entre deux référentiel \mathcal{R} et \mathcal{R}' avec une vitesse relative \mathbf{v} qui tend vers 0 conduit à

$$L(\dot{\mathbf{r}}^2) \simeq L(\dot{\mathbf{r}}'^2) + 2\dot{\mathbf{r}}' \cdot \mathbf{v} \frac{\partial L}{\partial \dot{\mathbf{r}}^2}(\dot{\mathbf{r}}'^2) \quad (4.33)$$

L'invariance est assurée si le dernier terme est une dérivée totale $\frac{d\chi}{dt}(\mathbf{r}', t) = \frac{\partial \chi}{\partial t} + \dot{\mathbf{r}}' \cdot \frac{\partial \chi}{\partial \mathbf{r}}(\mathbf{r}', t)$. Il s'ensuit que

$$\frac{\partial L}{\partial \dot{\mathbf{r}}^2}(\dot{\mathbf{r}}'^2) = \text{cste} \quad \Rightarrow \quad L(\dot{\mathbf{r}}^2) = \frac{1}{2}m\dot{\mathbf{r}}^2 \quad (4.34)$$

en choisissant de noter $\frac{1}{2}m$ la constante, positive pour assurer le minimum du principe de moindre action, et donc $\chi(\mathbf{r}', t) = m\mathbf{r}' \cdot \mathbf{v}$.

Mécanique relativiste – Vous verrez qu’avec les principes relativistes, le temps n’est plus absolu et se transforme avec l’espace. Il faut paramétrer l’action selon une quantité invariante par transformation de Lorentz. Le choix le plus naturel est l’intervalle d’espace-temps $ds = \sqrt{c^2 dt^2 - d\mathbf{r}^2}$ où la trajectoire (\mathbf{r}, ct) est associée à un référentiel donné associé à (\mathbf{r}, t) dans laquelle la particule a une vitesse $\dot{\mathbf{r}}$ et c représente la vitesse de la lumière. Un choix d’action simple est donc

$$S[\mathbf{r}] = -a \int_{s_1}^{s_2} ds = -ac \int_{t_1}^{t_2} dt \sqrt{1 - \dot{\mathbf{r}}^2/c^2} \quad (4.35)$$

avec a une constante que l’on détermine dans la limite non-relativiste $\|\dot{\mathbf{r}}\|^2 \ll c^2$ du lagrangien

$$\mathcal{L} \simeq -ac + \frac{ac}{2c^2} \dot{\mathbf{r}}^2 \quad \Rightarrow \quad a = mc. \quad (4.36)$$

Le choix de lagrangien relativiste

$$\mathcal{L} = -mc^2 \sqrt{1 - \frac{\dot{\mathbf{r}}^2}{c^2}} \quad (4.37)$$

reproduit bien les équations de la dynamique obtenues par Einstein et intègre par construction les bonnes invariances. On trouve en particulier

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}} = \frac{m\dot{\mathbf{r}}}{\sqrt{1 - \dot{\mathbf{r}}^2/c^2}} \quad \text{et} \quad E = \dot{\mathbf{r}} \cdot \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}} - \mathcal{L} = \frac{mc^2}{\sqrt{1 - \dot{\mathbf{r}}^2/c^2}} \quad (4.38)$$

dont on tire l’invariant relativiste $m^2 c^4 = E^2 - \mathbf{p}^2 c^2$. On notera la différence $\mathbf{p} \neq m\dot{\mathbf{r}}$. La généralisation de l’intervalle à des espaces courbes permet d’aborder le lagrangien de relativité générale.

2. Théorème de Noether

La physique moderne est fondée sur la notion de symétrie. Toute la physique fondamentale et en particulier la physique des particules se base sur la théorie des groupes notamment les groupes de Lie décrivant les symétries continues. Le théorème de Noether est un théorème important qui stipule que, pour toute invariance par symétrie, il est associée une quantité conservée. Considérons une transformation ponctuelle des coordonnées

$$Q_\alpha = f_\alpha(q_\beta, t; \epsilon) \quad (4.39)$$

paramétrée par un réel ϵ et telle que pour $\epsilon = 0$ on retrouve l’identité

$$f_\alpha(q_\beta, t; \epsilon = 0) = q_\alpha. \quad (4.40)$$

Le lagrangien est invariant sous f s’il satisfait à la relation

$$\forall \epsilon, \quad \mathcal{L}(f_\alpha(q_\beta, t; \epsilon); \dot{f}_\alpha(q_\beta, t; \epsilon), t) = \mathcal{L}(q_\alpha, \dot{q}_\alpha, t) \quad (4.41)$$

Dans la limite $\epsilon \rightarrow 0$, on obtient

$$\begin{aligned} \frac{d}{d\epsilon} \left(\mathcal{L}(f_\alpha(q_\beta, t; \epsilon); \dot{f}_\alpha(q_\beta, t; \epsilon), t) \right) &= 0 = \sum_\alpha \frac{\partial \mathcal{L}}{\partial q_\alpha} \frac{d}{d\epsilon} f_\alpha + \sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \frac{d}{d\epsilon} \frac{d}{dt} f_\alpha \\ \sum_\alpha \frac{\partial \mathcal{L}}{\partial q_\alpha} \frac{d f_\alpha}{d\epsilon} + \sum_\alpha \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \frac{d f_\alpha}{d\epsilon} \right) - \sum_\alpha \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) \frac{d f_\alpha}{d\epsilon} &= 0 \\ \sum_\alpha \left(\frac{\partial \mathcal{L}}{\partial q_\alpha} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) \right) \frac{d f_\alpha}{d\epsilon} + \frac{d}{dt} \left(\sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \frac{d f_\alpha}{d\epsilon} \right) &= 0 \end{aligned}$$

Le premier terme s'annule en vertu des équations d'Euler-Lagrange si bien que la quantité

$$\boxed{\Gamma = \sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \frac{d f_\alpha}{d\epsilon} \Big|_{\epsilon=0} = \text{cste}} \quad (4.42)$$

est une constante du mouvement.

On peut également généraliser le théorème de Noether dans le cas où la fonction de Lagrange n'est pas invariante sous la transformation ponctuelle mais peut s'écrire sous la forme d'une transformation de Jauge, qui laisse donc par conséquent les équations d'Euler-Lagrange invariantes

$$\mathcal{L}(f_\alpha(q_\beta, t; \epsilon); \dot{f}_\alpha(q_\beta, t; \epsilon), t) = \mathcal{L}(q_\alpha, \dot{q}_\alpha, t) + \frac{d}{dt} \Phi(q_\alpha, t; \epsilon). \quad (4.43)$$

Dans ce cas la quantité conservée s'écrit

$$\Gamma = \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \frac{d f_\alpha}{d\epsilon} \Big|_{\epsilon=0} - \frac{d\Phi}{d\epsilon} \Big|_{\epsilon=0} = \text{cste}. \quad (4.44)$$

Invariance par translation On considère un système de N particules de positions $\mathbf{r}_i = x_i \mathbf{e}_x + y_i \mathbf{e}_y + z_i \mathbf{e}_z$. On suppose que le système est invariant par une translation globale du système par exemple le long de l'axe Ox $\mathbf{r}_i \rightarrow \mathbf{f}_i(\{\mathbf{r}_i\}; \epsilon) = \mathbf{r}_i + \epsilon \mathbf{e}_x$. Si le lagrangien est invariant sous cette transformation la quantité conservée s'écrit

$$P_x = \sum_{i=1}^N \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i} \cdot \frac{d \mathbf{f}_i}{d\epsilon} \Big|_{\epsilon=0} = \sum_i \mathbf{p}_i \cdot \mathbf{e}_x = \sum_i p_{i,x} \quad (4.45)$$

La quantité conservée associée à l'invariance par translation le long de l'axe Ox est le moment linéaire total dans la direction Ox . On peut généraliser aux autres directions si bien que la quantité conservée associée à un système invariant par translation dans l'espace est le moment linéaire total $\mathbf{P} = \sum_i \mathbf{p}_i$.

Invariance par rotation On considère un système de N particules de positions \mathbf{r}_i . La rotation infinitésimale peut toujours s'écrire

$$\mathbf{r}_i \rightarrow \mathbf{r}_i + \delta \boldsymbol{\theta} \times \mathbf{r}_i, \quad (4.46)$$

où $\delta\boldsymbol{\theta} = \delta\theta\mathbf{n}$ est le vecteur rotation infinitésimale d'un angle $\delta\theta$ autour de l'axe \mathbf{n} . La dérivée de la transformation va donc s'écrire

$$\left. \frac{d\mathbf{f}_i}{d\theta} \right|_{\theta=0} = \mathbf{n} \times \mathbf{r}_i \quad (4.47)$$

Si le système est invariant par rotation autour de l'axe \mathbf{n} alors le théorème de Noether nous indique que la quantité conservée s'écrit

$$L_{\mathbf{n}} = \sum_{i=1}^N \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i} \cdot \left. \frac{d\mathbf{f}_i}{d\theta} \right|_{\theta=0} = \sum_{i=1}^N \mathbf{p}_i \cdot (\mathbf{n} \times \mathbf{r}_i) = \sum_{i=1}^N \mathbf{n} \cdot (\mathbf{r}_i \times \mathbf{p}_i) = \text{cste.} \quad (4.48)$$

Il s'agit ici du moment angulaire le long de l'axe \mathbf{n} . Si le système est complètement invariant par rotation alors ceci est vrai quelque soit \mathbf{n} ainsi on a nécessairement la conservation de la quantité

$$\mathbf{L} = \sum_{i=1}^N \mathbf{r}_i \times \mathbf{p}_i \quad (4.49)$$

qui est le moment angulaire total du système.

Invariance par translation dans le temps On considère la transformation de translation dans le temps et on suppose la fonction de Lagrange invariante sous cette transformation

$$\mathcal{L}(q_\alpha, \dot{q}_\alpha, t + \epsilon) = \mathcal{L}(q_\alpha, \dot{q}_\alpha, t), \quad \forall \epsilon, \quad (4.50)$$

c'est-à-dire $\frac{\partial \mathcal{L}}{\partial t} = 0$. On a déjà vu que dans un tel cas alors l'énergie E est une constante du mouvement. On peut simplement voir que si l'on considère l'espace des configurations étendu, le moment associé au temps et l'énergie $p_t = -E$ est donc une quantité conservée. Ainsi la conservation de l'énergie, de l'impulsion totale et du moment cinétique totale est une conséquence des symétries de l'espace-temps.

5. Oscillations

Comme application du formalisme lagrangien, nous étudions l'exemple de petites oscillations au voisinage d'un point d'équilibre. Si cela peut être traité à partir des équations de Newton, le formalisme lagrangien permet une approche fondée sur les termes énergétiques et les variables pertinentes, particulièrement utile pour la physique statistique et la physique quantique.

I. Linéarisation des équations du mouvement

1. Exemples et motivations

Nous avons en tête différents types de situations présentant des systèmes couplés : pendule double, systèmes masses-ressorts complexes, vibrations de molécules complexes et vibrations de réseaux cristallins.

2. Voisinage de l'équilibre

On adopte une notation $\alpha \rightarrow i = 1, \dots, n$ pour mettre en évidence une structure matricielle des équations. Le potentiel, noté $V(q_i)$, est supposé indépendant des vitesses généralisées. Autour d'un point d'équilibre q_i^* , caractérisé par l'annulation des forces $-\frac{\partial V}{\partial q_i}(q_k^*) = 0$, il se développe selon

$$V(q_i) \simeq V(q_i^*) + \sum_{i,j} \frac{1}{2} V_{ij} \frac{\partial^2 V}{\partial q_i \partial q_j}(q_k^*) (q_i - q_i^*) (q_j - q_j^*) = \text{cste} + \frac{1}{2} \mathbf{x}^t \mathbf{V} \mathbf{x} \quad (5.1)$$

avec $V_{ij} = \frac{\partial^2 V}{\partial q_i \partial q_j}(q_k^*)$ les éléments d'une matrice symétrique \mathbf{V} et $x_i = q_i - q_i^*$ les écarts à l'équilibre. De part la forme générale de l'énergie cinétique lorsque la transformation $\mathbf{r}_i \rightarrow q_\alpha$ est indépendante du temps, on obtient au second ordre en \mathbf{x} le développement

$$T(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \sum_{i,j} M_{ij}(q_k) \dot{q}_i \dot{q}_j \simeq \frac{1}{2} \sum_{i,j} M_{ij}(q_k^*) \dot{x}_i \dot{x}_j = \frac{1}{2} \dot{\mathbf{x}}^t \mathbf{M} \dot{\mathbf{x}} \quad (5.2)$$

avec $M_{ij} = M_{ij}(q_k^*)$ les éléments d'une matrice \mathbf{M} symétrique et définie positive ($\forall \mathbf{y}, \mathbf{y}^t \mathbf{M} \mathbf{y} \geq 0$) car elle représente l'énergie cinétique, ses valeurs propres sont réelles et positives ou nulles.

Le lagrangien prend donc une forme harmonique ou quadratique

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} \dot{\mathbf{x}}^t \mathbf{M} \dot{\mathbf{x}} - \frac{1}{2} \mathbf{x}^t \mathbf{V} \mathbf{x} = \frac{1}{2} \sum_{ij} M_{ij} \dot{x}_i \dot{x}_j - V_{ij} x_i x_j. \quad (5.3)$$

3. Équations du mouvement

L'application des équations d'Euler-Lagrange conduisent à

$$\frac{\partial \mathcal{L}}{\partial x_k} = -\frac{1}{2} \sum_j V_{kj} x_j - \frac{1}{2} \sum_j V_{jk} x_j = -\sum_j V_{kj} x_j \quad \text{et} \quad \frac{\partial \mathcal{L}}{\partial \dot{x}_k} = \sum_j M_{kj} \dot{x}_j \quad (5.4)$$

soit le système linéaire à résoudre

$$\mathbf{M} \ddot{\mathbf{x}} + \mathbf{V} \mathbf{x} = \mathbf{0}. \quad (5.5)$$

II. Pulsations propres et modes propres

1. Systèmes aux valeurs propres généralisé

On cherche à résoudre le système précédent en recherchant à découpler les équations. La linéarité en \mathbf{x} et ses dérivées suggère de recherche des solutions sous la forme $\mathbf{x} = \mathbf{a} e^{i\omega t}$ qui conduit au système $(-\omega^2 \mathbf{M} + \mathbf{V}) \mathbf{a} = \mathbf{0}$. Pour qu'il existe une solution $\mathbf{a} \neq \mathbf{0}$, le déterminant de ce système linéaire doit être nul, ce qui donne l'équation caractéristique

$$\det(\mathbf{V} - \omega^2 \mathbf{M}) = 0. \quad (5.6)$$

Cette équation généralise une équation aux valeurs propres $\lambda = \omega^2$ où, habituellement, $\mathbf{M} = \mathbb{I}$. On recherche donc les couples vecteurs propres / valeurs propres $\{\mathbf{a}_\alpha, \lambda_\alpha\}$ solutions du système aux valeurs propres généralisé

$$\mathbf{V} \mathbf{a}_\alpha = \lambda_\alpha \mathbf{M} \mathbf{a}_\alpha. \quad (5.7)$$

On utilise maintenant les hypothèses que \mathbf{V} est réelle symétrique et \mathbf{M} réelle, symétrique et définie positive. Dans ce cas, on démontre en annexe que le problème aux valeurs propres généralisé (5.7) admet n valeurs propres réelles λ_α associées aux \mathbf{a}_α qui peuvent être choisis réels et orthogonaux vis-a-vis du produit scalaire défini par la matrice \mathbf{M} , ie.

$$\mathbf{a}_\alpha^T \mathbf{M} \mathbf{a}_\beta = \delta_{\alpha\beta} \quad \text{ou} \quad \mathbf{P}^t \mathbf{M} \mathbf{P} = \mathbb{I} \quad (5.8)$$

avec \mathbf{P} la matrice de passage qui contient les vecteurs \mathbf{a}_α en colonnes

$$\mathbf{P} = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \quad (5.9)$$

En général, $\mathbf{P}^t \neq \mathbf{P}^{-1}$ et la matrice n'est pas orthogonale. En réécrivant (5.7) comme $\mathbf{VP} = \mathbf{MP}\mathbf{\Lambda}$ avec $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_n)$ et en multipliant par \mathbf{P}^t , il vient que \mathbf{P} diagonalise la matrice \mathbf{V}

$$\boxed{\mathbf{P}^t \mathbf{VP} = \mathbf{\Lambda}} \quad (5.10)$$

2. Modes propres ou modes normaux

Le paragraphe précédent suggère de se placer dans la base associée aux vecteurs propres. En posant $\mathbf{x} = \mathbf{P}\mathbf{Q}$, on voit que

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{V}\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{MP}\ddot{\mathbf{Q}} + \mathbf{VP}\mathbf{Q} = \mathbf{0} \Rightarrow \mathbf{P}^t \mathbf{MP}\ddot{\mathbf{Q}} + \mathbf{P}^t \mathbf{VP}\mathbf{Q} = \mathbf{0} \quad (5.11)$$

$$\Rightarrow \ddot{\mathbf{Q}} + \mathbf{\Lambda}\mathbf{Q} = \mathbf{0} \Rightarrow \boxed{\ddot{Q}_\alpha + \omega_\alpha^2 Q_\alpha = 0} \quad (5.12)$$

On obtient donc n équations découplées pour les variables Q_α qui sont appelées modes propres ou modes normaux. Les solutions sont de la forme

$$Q_\alpha(t) = \begin{cases} c_\alpha \cos(\omega_\alpha t + \varphi_\alpha) & \text{si } \lambda_\alpha > 0 \\ c_\alpha t + b_\alpha & \text{si } \lambda_\alpha = 0 \\ c_\alpha e^{|\omega_\alpha|t} + b_\alpha e^{-|\omega_\alpha|t} & \text{si } \lambda_\alpha < 0 \end{cases} \quad (5.13)$$

où les constantes c_α, b_α sont à déterminer en fonction des conditions initiales.

On trouve que le lagrangien se simplifie en

$$\boxed{\mathcal{L}(\mathbf{Q}, \dot{\mathbf{Q}}) = \frac{1}{2} \sum_\alpha \dot{Q}_\alpha^2 - \omega_\alpha^2 Q_\alpha^2} \quad (5.14)$$

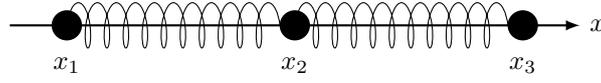
Connaissant la matrice \mathbf{P} , on aurait d'ailleurs pu simplifier directement le lagrangien duquel les équations du mouvement auraient été déduites trivialement.

Si \mathbf{M} est diagonale avec $M_{ii} > 0$, on peut simplement se ramener à un problème aux valeurs propres standard en simplifiant par $\mathbf{M}^{-1/2}$ le problème aux valeurs propres. Si \mathbf{V} est également une matrice définie positive alors toutes les valeurs propres sont positives $\lambda_\alpha > 0$.

III. Exemple : Chaîne linéaire de 3 particules

On considère ici le cas d'une chaîne linéaire de 3 particules de masse m qui interagissent via un potentiel qui s'apparente à des ressorts entre chaque particules (Fig. 5.1). Typiquement ce type de potentiel intervient lors que l'on considère des interactions entre atomes dans un solide ou dans une molécule. Le potentiel s'écrit donc

$$V = \frac{k}{2} (x_2 - x_1 - a)^2 + \frac{k}{2} (x_3 - x_2 - a)^2, \quad (5.15)$$

FIGURE 5.1. – Chaîne linéaire de 3 particules de masse m .

où a est la distance d'équilibre entre les masses et où k est une constante de raideur. L'énergie cinétique du système s'écrit

$$T = \frac{1}{2}m (\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2). \quad (5.16)$$

Le système est invariant par translation, c'est-à-dire qu'il est possible de faire une translation globale du système le long de l'axe Ox sans changer les propriétés du système. On peut cependant définir des positions d'équilibre de manière relative. Ces positions sont données par

$$x_2^{(0)} - x_1^{(0)} = a, \quad (5.17)$$

$$x_3^{(0)} - x_2^{(0)} = a. \quad (5.18)$$

On définit l'écart à la position d'équilibre $q_i = x_i - x_i^{(0)}$. Le Lagrangien s'écrit donc

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{m}{2} (\dot{q}_1^2 + \dot{q}_2^2 + \dot{q}_3^2) - \frac{k}{2} (q_1^2 + 2q_2^2 + q_3^2 - 2q_1q_2 - 2q_2q_3). \quad (5.19)$$

Les équations de Euler-Lagrange donnent

$$m\ddot{q}_1 + kq_1 - kq_2 = 0, \quad (5.20)$$

$$m\ddot{q}_2 + 2kq_2 - kq_1 - kq_3 = 0, \quad (5.21)$$

$$m\ddot{q}_3 + kq_3 - kq_2 = 0. \quad (5.22)$$

On recherche une solution sous la forme $q_i = \text{Re}(a_i e^{i\omega t})$

$$\begin{pmatrix} \omega_0^2 - \omega^2 & -\omega_0^2 & 0 \\ -\omega_0^2 & 2\omega_0^2 - \omega^2 & -\omega_0^2 \\ 0 & -\omega_0^2 & \omega_0^2 - \omega^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = 0, \quad (5.23)$$

où $\omega_0 = \sqrt{k/m}$. Une solution non-triviale est donc donnée par l'équation caractéristique

$$\begin{vmatrix} \omega_0^2 - \lambda & -\omega_0^2 & 0 \\ -\omega_0^2 & 2\omega_0^2 - \lambda & -\omega_0^2 \\ 0 & -\omega_0^2 & \omega_0^2 - \lambda \end{vmatrix} = 0, \quad (5.24)$$

où $\lambda = \omega^2$. Cette équation se factorise sous la forme

$$\lambda (\omega_0^2 - \lambda) (\lambda - 3\omega_0^2) = 0 \quad (5.25)$$

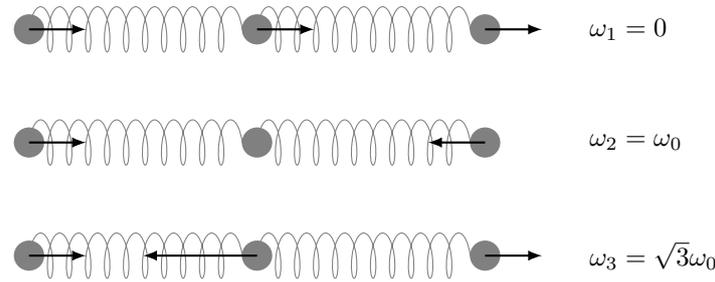


FIGURE 5.2. – Représentation des 3 modes propres de la chaîne linéaire à 3 masses.

La solution générale est donc une somme d'oscillateurs harmonique de fréquences propres $\omega_1 = 0$, $\omega_2 = \omega_0$ et $\omega_3 = \sqrt{3}\omega_0$. Chaque oscillateur harmonique est un mode propre du système. Pour comprendre la nature de chacun des modes propres il est nécessaire de résoudre le problème aux valeurs propres généralisé.

Mode propre $\omega_1 = 0$ En reprenant la définition des matrices \mathbf{V} et \mathbf{M} introduites précédemment celle-ci s'écrivent dans la cas de la chaîne à 3 masses sous la forme

$$\mathbf{V} = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{pmatrix} \quad (5.26)$$

Le vecteur propre de valeur propre 0 est donné par l'équation $\mathbf{V}\mathbf{a} = 0$, ce qui donne

$$a_1 = a_2 = a_3. \quad (5.27)$$

Ce mode propre correspond donc à une translation globale du système. On retrouve l'invariance par translation le long de l'axe Ox . Ce mode propre est représenté sur la figure 5.2.

Mode propre $\omega_2 = \omega_0$ Le vecteur propre correspondant est donné par l'équation $\mathbf{V}\mathbf{a} = \omega_0^2\mathbf{M}\mathbf{a}$, ce qui donne

$$a_2 = 0 \quad a_1 = -a_3. \quad (5.28)$$

C'est à dire que c'est un mode de vibration symétrique par rapport à la symétrie d'inversion du système.¹ Ce mode propre est représenté sur la figure 5.2.

Mode propre $\omega_3 = \sqrt{3}\omega_0$ Le vecteur propre correspondant est donné par l'équation $\mathbf{V}\mathbf{a} = 3\omega_0^2\mathbf{M}\mathbf{a}$. Ce qui donne

$$a_2 = -2a_1 = -2a_3. \quad (5.29)$$

1. En général les propriétés de symétrie d'un système permet de simplifier la détermination des fréquences propres et modes normaux par l'utilisation de la théorie des groupes.

C'est un mode de vibration antisymétrique par rapport à la symétrie d'inversion. Ce mode propre est représenté sur la figure 5.2. Les modes normaux peuvent être normalisé par rapport au produit scalaire définie par la matrice \mathbf{M} , $\mathbf{a}^t \mathbf{M} \mathbf{a} = 1$, ce qui donne la matrice de passage

$$\mathbf{P} = \frac{1}{\sqrt{m}} \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{pmatrix} \quad (5.30)$$

La solution générale s'écrit donc $\mathbf{q}(t) = \mathbf{P}\mathbf{Q}(t)$ où

$$Q_1(t) = A_1 t + B_1, \quad (5.31)$$

$$Q_2(t) = C_2 \cos(\omega_2 t + \varphi_2), \quad (5.32)$$

$$Q_3(t) = C_3 \cos(\omega_3 t + \varphi_3). \quad (5.33)$$

C1. . Problème aux valeurs propres généralisé

Soit \mathbf{A} et \mathbf{B} deux matrices hermitiennes de dimension $n \times n$. On suppose également que \mathbf{B} est définie positive, c'est-à-dire

$$\forall x \in \mathbb{C}^n, \quad x^\dagger \mathbf{B} x > 0 \quad (5.34)$$

Le problème aux valeurs propres généralisé s'écrit

$$\mathbf{A}x = \lambda \mathbf{B}x. \quad (5.35)$$

On peut réécrire cette équation sous la forme

$$(\mathbf{A} - \lambda \mathbf{B})x = 0 \quad (5.36)$$

Une solution non triviale $x \neq 0$ implique que la matrice $\mathbf{A} - \lambda \mathbf{B}$ est une matrice singulière, c'est-à-dire que l'on obtient l'équation caractéristique

$$\det\{\mathbf{A} - \lambda \mathbf{B}\} = 0. \quad (5.37)$$

C'est une équation algébrique de degré n en λ , elle possède donc n solutions notées λ_i . Les vecteurs propres x_i sont alors déterminés par l'équation

$$\mathbf{A}x_i = \lambda_i \mathbf{B}x_i \quad (5.38)$$

En multipliant cette équation par x_j^\dagger on obtient

$$x_j^\dagger \mathbf{A}x_i = \lambda_i x_j^\dagger \mathbf{B}x_i. \quad (5.39)$$

Prenons le complexe conjugué de cet équation et inversons les indices i et j on obtient

$$x_j^\dagger \mathbf{A}x_i = \lambda_j^* x_j^\dagger \mathbf{B}x_i \quad (5.40)$$

En effectuant une soustraction les deux dernières équations on obtient

$$(\lambda_j^* - \lambda_i) x_j^\dagger \mathbf{B} x_i = 0. \quad (5.41)$$

Pour $i = j$, on en déduit que $\lambda_i = \lambda_i^*$, c'est-à-dire que l'ensemble des valeurs propres sont réelles. Pour $i \neq j$ on a donc

$$(\lambda_j - \lambda_i) x_j^\dagger \mathbf{B} x_i = 0. \quad (5.42)$$

Si toutes les valeurs propres sont distinctes $\lambda_i \neq \lambda_j$ alors

$$x_j^\dagger \mathbf{B} x_i = 0. \quad (5.43)$$

D'autre part, on peut choisir arbitrairement la normalisation de chaque vecteur propre ainsi on peut choisir cette normalisation de sorte que $x_i^\dagger \mathbf{B} x_i = 1$ ce qui est toujours possible si \mathbf{B} est définie positive. Ainsi on a

$$x_i^\dagger \mathbf{B} x_j = \delta_{ij}. \quad (5.44)$$

Si il existe une valeur propre λ_0 de multiplicité g ($1 < g \leq n$). Dans ce cas, il existe un ensemble de g vecteurs propres distincts y_k tel que

$$\mathbf{A} y_k = \lambda_0 \mathbf{B} y_k \quad (5.45)$$

Toute superposition de ces vecteurs propres est également solution de cette équation. Ainsi, cet ensemble de vecteurs propres associés forme un sous espace vectoriel de dimension g et tout vecteur appartenant à ce sous-espace est un vecteur propre de valeur propre λ_0 . On peut donc choisir de manière arbitraire ces vecteurs : on peut les choisir orthogonaux entre-eux vis-à-vis du produit scalaire défini par la matrice \mathbf{B} , c'est-à-dire

$$x_i^\dagger \mathbf{B} x_j = \delta_{ij}. \quad (5.46)$$

Ainsi, qu'il y ait des valeurs propres dégénérées ou non il est toujours possible de choisir un ensemble de n vecteurs propres x_i orthogonaux entre eux vis-a-vis du produit scalaire généré par la matrice \mathbf{B} . Ce dernier résultat permet de montrer que

$$\lambda_i = x_i^\dagger \mathbf{A} x_i. \quad (5.47)$$

Ainsi si \mathbf{A} est une matrice définie positive alors l'ensemble des valeurs propres sont positives $\lambda_i > 0 \forall i$.

On peut introduire la matrice de passage \mathbf{P} dont les colonnes correspondent aux vecteurs propres x_i ainsi

$$\mathbf{P} = (x_1, x_2, \dots, x_n) = \begin{pmatrix} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & \dots & x_{n2} \\ \vdots & \vdots & \dots & \vdots \\ x_{1n} & x_{2n} & \dots & x_{nn} \end{pmatrix} \quad (5.48)$$

En général cette matrice n'est pas unitaire² $\mathbf{P}^{-1} \neq \mathbf{P}^\dagger$. On montre aisément les deux équations suivantes :

$$\mathbf{P}^\dagger \mathbf{B} \mathbf{P} = \mathbb{I} \quad (5.49)$$

$$\mathbf{P}^\dagger \mathbf{A} \mathbf{P} = \mathbf{\Lambda} \quad (5.50)$$

où $\mathbf{\Lambda}$ est la matrice diagonale dont les éléments sont les valeurs propres λ_i . Ainsi dans la base des vecteurs propres les deux matrices \mathbf{A} et \mathbf{B} sont diagonales.

Si les matrices \mathbf{A} et \mathbf{B} sont également réelles (et par conséquent symétriques) alors il est toujours possible de choisir les vecteurs propres réels.

2. Dans le cas du problème aux valeurs propres classique où la matrice \mathbf{B} est la matrice identité alors \mathbf{P} est une matrice unitaire

Deuxième partie

Formulation Hamiltonienne

6. Formalisme hamiltonien

L'objectif du formalisme hamiltonien est de passer d'un lagrangien dépendant des q_α et \dot{q}_α à une fonction dépendante des q_α et p_α . Cette reformulation est préférable dans de nombreuses situations et approches de la mécanique. En particulier, les équations présente une structure qui offre une meilleure interprétation qualitative et quantitative de la dynamique. Ce sera le jeu de variables naturellement utilisé pour la mécanique quantique et la physique statistique. Plusieurs méthodes efficaces de résolution des équations du mouvement s'inscrivent naturellement dans cette formulation.

I. Transformée de Legendre

1. Approche intuitive

Soit $\phi : x \rightarrow \phi(x)$ une fonction réelle, $x \in \mathbb{R}$, telle que $\phi'(x)$ existe et inversible. L'idée est de construire une fonction ψ qui dépende de la variable $p = \phi'(x)$ et qui permette de reconstruire ϕ si besoin, qui "transforme" l'information contenue dans ϕ . La méthode intuitive utilisée en thermodynamique est de raisonner sur les différentielles. En choisissant $\psi = xp - \phi(x)$, on voit que

$$d\phi = \phi'(x)dx = pdx \quad \text{et} \quad d\psi = pdx + xdp - d\phi = xdp \quad (6.1)$$

de laquelle on retient à la fois que la variable de ψ est p , soit $\psi(p)$ et que l'on a la relation réciproque $x = \psi'(p)$. On en retient le schéma de transformée de Legendre et sa réciproque :

$$\begin{array}{ccc}
 & \begin{array}{c} p=\phi'(x) \\ \psi=xp-\phi(x) \end{array} & \\
 & \curvearrowright & \\
 (x, \phi(x)) & & (p, \psi(p)) \\
 & \curvearrowleft & \\
 & \begin{array}{c} x=\psi'(p) \\ \phi=xp-\psi(p) \end{array} &
 \end{array}$$

On dit que x et p sont des variables conjuguées par rapport à ϕ ou ψ .

Exemple : En thermodynamique, [*Attention, les conventions de signe sont généralement différentes* $x = V, \phi = U, p = -P, \psi = -H$]. La transformée de $U(S, V)$ telle que $dU = TdS - PdV$ avec $P = -\frac{\partial U}{\partial V}$ par rapport au volume V se fait en introduisant $H = U + PV$ qui vérifie $dH = TdS + VdP$ et $V = \frac{\partial H}{\partial P}$.

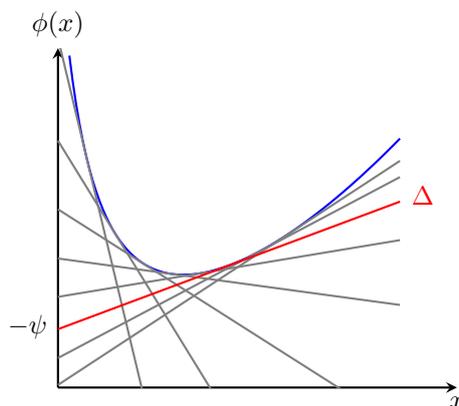


FIGURE 6.1. – Représentation géométrique de la transformée de Legendre.

2. Approche plus rigoureuse

On définit $p(x) = \phi'(x)$ que l'on peut inverser en $x(p) = (\phi')^{-1}(p)$ et on recherche la forme de la fonction $\psi(p)$ qui doit satisfaire à la relation symétrique $x(p) = \psi'(p)$. Raisonnons avec la variable p :

$$\phi'(x(p)) = p \Leftrightarrow \phi'(x(p))x'(p) = px'(p) \Leftrightarrow \frac{d}{dp}(\phi(x(p))) = \frac{d}{dp}(px(p)) - x(p) \quad (6.2)$$

$$\Leftrightarrow x(p) = \frac{d}{dp}(-\phi(x(p)) + px(p)) \quad (6.3)$$

de sorte que le choix $\psi(p) = px(p) - \phi(x(p))$ assure bien que $x(p) = \psi'(p)$.

Exemple : $\phi(x) = \frac{1}{2}\alpha x^2 + \beta x + \gamma$ avec $\alpha > 0$ donne $p(x) = \alpha x + \beta$ puis $x(p) = (p - \beta)/\alpha$ et finalement

$$\psi(p) = x(p)p - \phi(x(p)) = \frac{(p - \beta)^2}{2\alpha} - \gamma. \quad (6.4)$$

Interprétation géométrique : On peut représenter la fonction ϕ par l'ensemble des lignes tangentes à la courbe (Fig. 6.1). Chacune de ces lignes est caractérisée par sa pente p et son ordonnée à l'origine $-\psi$ si bien que l'on a $\phi = px - \psi$. La fonction ϕ peut donc être également représentée par l'ensemble des couples (p, ψ) .

La figure 6.2 représente une fonction concave $\phi(x)$ ainsi que sa transformée de Legendre $\psi(p)$. Le point $p = 0$ correspond au minimum de $\phi(x)$. De même, le point $x = 0$ correspond au minimum de $\psi(p)$. Au point x , la pente de la droite tangente à la courbe $\phi(x)$ est donnée par p et la hauteur du triangle est alors donnée par xp , la transformée de Legendre est ainsi définie par $\psi(p) = xp - \phi(x)$.

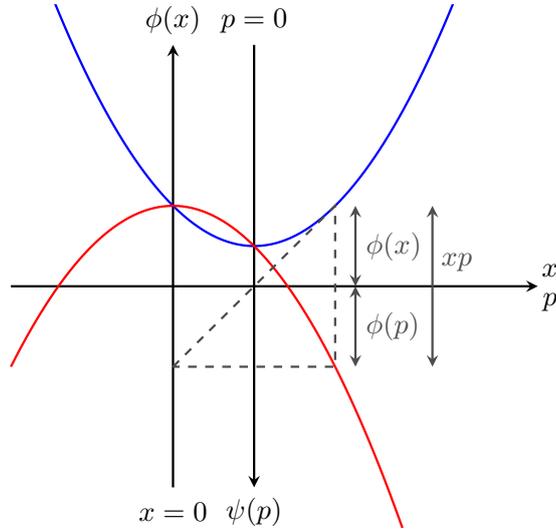


FIGURE 6.2. – Interprétation géométrique de la transformation de Legendre.

3. Généralisation à plusieurs variables

Soit $\mathbf{x} = (x_1, \dots, x_n)$ un ensemble de n variables et $\phi(\mathbf{x})$ une fonction dérivable. On définit les variables conjuguées

$$p_i(x_1, \dots, x_n) = \frac{\partial \phi}{\partial x_i}, \tag{6.5}$$

si bien que la différentielle de ϕ s'écrit $d\phi = \sum_{i=1}^n p_i(x_1, \dots, x_n) dx_i$.

La transformée de Legendre de la fonction ϕ par rapport aux k premières variables s'écrit

$$\psi(p_1, \dots, p_k, x_{k+1}, \dots, x_n) = \sum_{i=1}^k x_i p_i - \phi(\mathbf{x}), \tag{6.6}$$

et sa différentielle s'écrit $d\psi = \sum_{i=1}^k p_i dx_i + x_i dp_i - d\phi = \sum_{i=1}^k x_i dp_i - \sum_{i=k+1}^n p_i dx_i$,

si bien qu'on a la relation conjuguée à l'Eq. (6.5) pour $i = 1, \dots, k$

$$x_i(p_1, \dots, p_k, x_{k+1}, \dots, x_n) = \frac{\partial \psi}{\partial p_i} \tag{6.7}$$

Exemple : Soit $\phi(\mathbf{x}) = \frac{1}{2} \sum_{i,j} A_{ij} x_i x_j$, où A_{ij} est une matrice symétrique et définie positive. On définit les variables conjuguées $p_i = \frac{\partial \phi}{\partial x_i} = \sum_j A_{ij} x_j$, si bien que $x_i = \sum_j (A^{-1})_{ij} p_j$. La transformée de Legendre de $\phi(\mathbf{x})$ s'écrit

$$\psi(\mathbf{p}) = \sum_{i=1}^n x_i(\mathbf{p}) p_i - \phi(\mathbf{x}(\mathbf{p})) = \frac{1}{2} \sum_{i,j} (A^{-1})_{ij} p_i p_j \tag{6.8}$$

Remarque : Pour obtenir les $x_i(p_1, \dots, p_k, x_{k+1}, \dots, x_n)$, il faut pouvoir inverser le changement de variables $x_i \rightarrow p_i = \frac{\partial \phi}{\partial x_i}$ donc le jacobien est $J_{ij} = \frac{\partial p_i}{\partial x_j} = \frac{\partial^2 \phi}{\partial x_i \partial x_j}$, c'est-à-dire la Hessienne de ϕ . Il faut donc que $\det J \neq 0$ et à 1D, $\phi''(x) \neq 0$, ce qui n'est pas toujours vérifié.

II. Équations de Hamilton

1. Fonction de Hamilton ou hamiltonien

On passe à une description en fonction du jeu de variables (q_α, p_α) en effectuant une transformation de Legendre sur les variables \dot{q}_α uniquement, avec $p_\alpha = \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}$. On définit ainsi le Hamiltonien, ou fonction de Hamilton :

$$H(q_\alpha, p_\alpha, t) = \sum_{\alpha=1}^n p_\alpha \dot{q}_\alpha - \mathcal{L}(q_\alpha, \dot{q}_\alpha, t) . \quad (6.9)$$

Remarque : L'expression du Hamiltonien coïncide avec celle de l'énergie lagrangienne mais ce ne sont pas les mêmes fonctions car elles n'ont pas les mêmes variables et la même expression en fonction de celles-ci. Dans de nombreux cas pratiques simples, H représente bien l'énergie du système.

2. Équations du mouvement

On cherche maintenant à décrire les équations du mouvement en termes des variables (q_α, p_α) . Une première approche est de partir des équations d'Euler-Lagrange :

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) - \frac{\partial \mathcal{L}}{\partial q_\alpha} = 0 \Leftrightarrow \frac{dp_\alpha}{dt} = \frac{\partial \mathcal{L}}{\partial q_\alpha} = - \frac{\partial H}{\partial q_\alpha} . \quad (6.10)$$

Une seconde équation est fournie par la relation de conjugaison entre les variables \dot{q}_α et p_α par rapport à H dans la transformé de Legendre, il vient

$$\dot{q}_\alpha = \frac{\partial H}{\partial p_\alpha} . \quad (6.11)$$

Cet ensemble de $2n$ équations du *premier ordre* par rapport au temps constitue les **équations de Hamilton** :

$$\begin{cases} \dot{q}_\alpha = + \frac{\partial H}{\partial p_\alpha} \\ \dot{p}_\alpha = - \frac{\partial H}{\partial q_\alpha} \end{cases} . \quad (6.12)$$

Concernant la dépendance temporelle, il vient immédiatement que

$$\frac{\partial H}{\partial t} = - \frac{\partial \mathcal{L}}{\partial t} . \quad (6.13)$$

3. Principe de Hamilton

Ces équations peuvent être dérivées dans le cadre un peu plus général du principe de Hamilton. Sans supposer qu'il existe un lagrangien et que l'on puisse lui appliquer une transformée de Legendre, on peut utiliser le principe de moindre action en considérant que les trajectoires sont décrites dans un jeu de variables indépendantes étendu $(q_\alpha, \dot{q}_\alpha, p_\alpha, \dot{p}_\alpha)$ par l'action

$$\mathcal{S}[q_\alpha, p_\alpha] = \int_{t_1}^{t_2} dt \left(\sum_\alpha p_\alpha \dot{q}_\alpha - H(q_\alpha, p_\alpha, t) \right) \quad (6.14)$$

qui est une fonctionnelle de q_α et de p_α . On suppose que les variables (q_α, p_α) sont fixées aux bords, i.e. en t_1 et t_2 . La minimisation fonctionnelle de \mathcal{S} conduit alors, en posant $s(q_\alpha, \dot{q}_\alpha, p_\alpha, \dot{p}_\alpha) = \sum_\alpha p_\alpha \dot{q}_\alpha - H(q_\alpha, p_\alpha, t)$ et en utilisant la formule d'Euler lagrange pour la dérivée fonctionnelle

$$\frac{\delta \mathcal{S}}{\delta p_\alpha} = \frac{d}{dt} \left(\frac{\partial s}{\partial \dot{p}_\alpha} \right) - \frac{\partial s}{\partial p_\alpha} = 0 \Rightarrow \dot{q}_\alpha = \frac{\partial H}{\partial p_\alpha}, \quad (6.15)$$

$$\frac{\delta \mathcal{S}}{\delta q_\alpha} = \frac{d}{dt} \left(\frac{\partial s}{\partial \dot{q}_\alpha} \right) - \frac{\partial s}{\partial q_\alpha} = 0 \Rightarrow \dot{p}_\alpha = -\frac{\partial H}{\partial q_\alpha}. \quad (6.16)$$

On conclue donc que la minimisation d'une action écrite sous la forme (6.14) supposant seulement l'existence d'un Hamiltonien H conduit génériquement aux équations de Hamilton.

III. Exemples (voir TDs)

1. Le pendule simple

Avec $p_\theta = m\ell^2\dot{\theta}$, on obtient

$$\mathcal{L}(\theta, \dot{\theta}) = \frac{1}{2}m\ell^2\dot{\theta}^2 + mg\ell \cos \theta \quad \Rightarrow \quad H(\theta, p_\theta) = \frac{p_\theta^2}{2m\ell^2} - mg\ell \cos \theta. \quad (6.17)$$

2. Particule chargée dans un champ électromagnétique

À partir de la forme du lagrangien et du moment conjugué

$$\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) = \frac{1}{2}m\dot{\mathbf{r}}^2 + q(\dot{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r}, t) - V(\mathbf{r}, t)) \quad (6.18)$$

$$\mathbf{p} = m\dot{\mathbf{r}} + q\mathbf{A}(\mathbf{r}, t) \quad (6.19)$$

on en tire la fonction de Hamilton sous la forme

$$H(\mathbf{r}, \mathbf{p}, t) = \frac{1}{2m} (\mathbf{p} - q\mathbf{A}(\mathbf{r}, t))^2 + qV(\mathbf{r}, t). \quad (6.20)$$

3. Hamiltonien relativiste

Nous avons obtenu le lagrangien relativiste $\mathcal{L} = -mc^2 \sqrt{1 - \dot{\mathbf{r}}^2/c^2}$ et le moment conjugué $\mathbf{p} = \frac{m\dot{\mathbf{r}}}{\sqrt{1 - \dot{\mathbf{r}}^2/c^2}}$ dont on tire, après calculs, l'Hamiltonien

$$H(\mathbf{r}, \mathbf{p}) = \sqrt{m^2c^4 + (\mathbf{p}c)^2} . \quad (6.21)$$

4. Coordonnées cylindriques et sphériques

Il s'agit d'exemple de changement de coordonnées qui conduisent à une forme non triviale de l'Hamiltonien. Pour une particule dans un potentiel externe $U(\mathbf{r})$.

En cylindrique (r, θ, z) , la partie cinétique $T = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2)$ donne $p_r = m\dot{r}$, $p_\theta = mr^2\dot{\theta}$, $p_z = m\dot{z}$ puis

$$H(r, \theta, z, p_r, p_\theta, p_z) = \frac{1}{2m} \left(p_r^2 + \frac{p_\theta^2}{r^2} + p_z^2 \right) + U(r, \theta, z) . \quad (6.22)$$

En sphériques (r, θ, φ) , la partie cinétique $T = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2 \theta \dot{\varphi}^2)$ donne $p_r = m\dot{r}$, $p_\theta = mr^2\dot{\theta}$, $p_\varphi = mr^2 \sin^2 \theta \dot{\varphi}$ puis

$$H(r, \theta, \varphi, p_r, p_\theta, p_\varphi) = \frac{1}{2m} \left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2 \sin^2 \theta} \right) + U(r, \theta, \varphi) . \quad (6.23)$$

Ces formes seront par exemple utiles en physique statistique.

7. Dynamique hamiltonienne

Le but de ce chapitre est de montrer l'utilité de la formulation hamiltonienne pour la description de l'évolution temporelle des systèmes. On y introduit notamment l'espace des phases, des résultats importants physiquement ainsi que les crochets de Poisson.

I. Évolution temporelle et espace des phases

1. Système dynamique

En mathématiques, un système dynamique est décrit par un vecteur réel $\mathbf{X}(t)$ de dimension N et gouverné par une équation différentielle de la forme

$$\dot{\mathbf{X}} = \mathbf{F}(\mathbf{X}, t) \quad (7.1)$$

dans laquelle le champ de vecteurs \mathbf{F} caractérise la vitesse locale dans l'espace des phases qui est l'espace dans lequel évolue $\mathbf{X}(t)$.

Si \mathbf{F} ne dépend pas explicitement du temps t , on le qualifie de **système autonome**. Dans ce cas, $\mathbf{F}(\mathbf{X})$ étant fixé. S'il ne présente pas de trop fortes singularités, les trajectoires suivent de proche en proche ce champ de vecteur en lui étant localement tangentes, conduisant à des résultats importants.

On appelle **points fixes** les points de l'espace des phases tels que $\dot{\mathbf{X}} = 0$ pour des systèmes autonomes. Leur détermination revient donc à chercher les solutions de $\mathbf{F}(\mathbf{X}) = \mathbf{0}$. Il s'agit des "positions d'équilibre" habituellement étudiées en mécanique.

Il est clair que les équations de Hamilton entre dans le cadre des systèmes dynamiques avec $\mathbf{X} = (q_1, \dots, q_n, p_1, \dots, p_n)$ un vecteur de dimension $2n$. L'espace des phases est donc l'espace des (q_α, p_α) dans ce cas, plutôt difficile à représenter dès que $n > 1$.

2. Déterminisme

Pour un système autonome, le théorème de Cauchy-Lifshitz propose des conditions de régularité suffisantes pour assurer l'existence et l'unicité de la solution pour conditions initiales données, autrement appelé le **déterminisme**. Par exemple si \mathbf{F} est continûment

différentiable sur un ouvert de \mathbb{R}^N , ce qui est très souvent le cas en physique. Ce résultat, dont la démonstration relève plus du cours de mathématiques, est cependant intuitif lorsque l'on comprend que fixer le champ de vecteur \mathbf{F} avec localement des variations continues détermine de façon univoque la suite de la trajectoire.

Une conséquence importante physiquement est que, dans ce cas, **deux trajectoires ne peuvent se couper dans l'espace des phases**. En effet, si c'était le cas, le croisement constituerait des conditions initiales identiques pour deux trajectoires ensuite distinctes, en contradiction avec le théorème. En zoomant sur le point de croisement, cela signifierait que le champ \mathbf{F} est localement discontinu ou bivalué. Notons que le cas des solutions périodiques est un peu particulier dans le sens où tous les points de l'orbite sont des conditions initiales qui partagent la même trajectoire sans qu'il y ait de croisement. Enfin, notons qu'aux points fixes peuvent converger ou diverger des trajectoires "issues" de voisinage du point fixe (voir le cas des séparatrices pour le pendule simple ci-après).

Remarque : Dans le cas de systèmes non-autonomes, on a souvent déterminisme mais une trajectoire peut se recouper à des temps différents car le champ \mathbf{F} n'est plus fixé.

3. Système hamiltonien conservatif et portrait de phase

Dans le cas d'un hamiltonien indépendant du temps $\frac{\partial H}{\partial t} = 0$ (système conservatif), on a donc un système dynamique autonome. Il s'ensuit que la fonction de Hamilton est une constante du mouvement :

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_{\alpha} \frac{\partial H}{\partial q_{\alpha}} \dot{q}_{\alpha} + \frac{\partial H}{\partial p_{\alpha}} \dot{p}_{\alpha} = \sum_{\alpha} \frac{\partial H}{\partial q_{\alpha}} \frac{\partial H}{\partial p_{\alpha}} - \frac{\partial H}{\partial p_{\alpha}} \frac{\partial H}{\partial q_{\alpha}} = 0 \Rightarrow \boxed{H(q_{\alpha}, p_{\alpha}) = \text{const.}} \quad (7.2)$$

L'évolution dynamique du système se situe donc sur l'hypersurface de dimension $2n - 1$ d'équation $H(q_{\alpha}, p_{\alpha}) = \text{const.}$ On appelle **portrait de phase** l'analyse géométrique des trajectoires typiques possibles dans l'espace des phases. Pour un système à $n = 1$ degré de liberté, cela permet une résolution graphique du mouvement comme illustré en TD et dans l'exemple ci-dessous.

Exemple : le pendule simple – Le pendule simple est un oscillateur non-linéaire dont la dynamique, en omettant les constantes physiques, est gouvernée par le hamiltonien

$$H(q, p) = \frac{p^2}{2} - \cos q, \quad (7.3)$$

q est donc l'angle θ du pendule. Le champ de vitesse dans l'espace des phases s'écrit

$$\dot{\mathbf{X}} = \mathbf{F}(\mathbf{X}) = \begin{pmatrix} p \\ -\sin q \end{pmatrix} \quad (7.4)$$

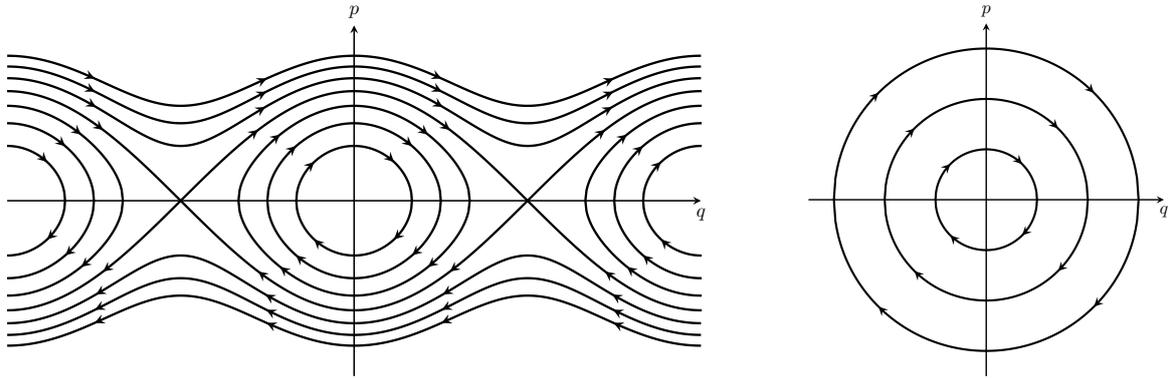


FIGURE 7.1. – Portraits de phase de l'oscillateur harmonique et de l'oscillateur non-linéaire.

Sur la figure 7.1 sont représentées les trajectoires, ou orbites, possibles correspondant à $H(q, p) = \text{const.}$ Les points fixes sont en $(0, 0)$ (équilibre stable) et $(\pm\pi, 0)$ (équilibre instable, pendule à l'envers).

Les trajectoires concentriques qui s'enroulent autour de $(0, 0)$ sont des mouvements d'oscillations du pendule avec aller-retours, c'est-à-dire changement de signe de la vitesse angulaire. Au voisinage de $(0, 0)$, on retrouve le portrait de phase d'un oscillateur harmonique $H(q, p) = \frac{p^2}{2} + \frac{q^2}{2}$ qui sont ici des cercles (figure de droite).

Les trajectoires qui passent au-dessus et en-dessous des points instables $(\pm\pi, 0)$ sont une succession de rotations complètes du pendule sans changement de signe de la vitesse. Les deux ensembles de trajectoires sont séparés par une ligne séparatrice qui correspond au cas limite d'un mouvement qui part de $q = -\pi$ sans vitesse initiale pour aller à $q = +\pi$ en un temps infini.

4. Théorèmes généraux

Ces théorèmes concernent les systèmes hamiltoniens autonomes. Ils sont essentiels aux fondements de la description classique de la physique statistique.

Théorème de Liouville : Il peut se retenir sous la forme « Le volume de l'espace des phases constitué d'un ensemble de points qui évoluent selon la dynamique hamiltonienne est conservé au cours du temps. » Pour le montrer, considérons un l'ensemble de points appartenant à un volume connexe $\Omega(t_1)$ qui va évoluer dans le temps vers le domaine noté $\Omega(t)$. Sous forme d'intégrales, les volumes entre les temps t et $t + dt$ sont notés

$$\Omega(t) = \int_{(q_\alpha, p_\alpha) \in \Omega(t)} dq_1 \cdots dq_n dp_1 \cdots dp_n \quad \text{et} \quad \Omega(t + dt) = \int_{(q'_\alpha, p'_\alpha) \in \Omega(t+dt)} dq'_1 \cdots dq'_n dp'_1 \cdots dp'_n \quad (7.5)$$

Les points évoluent selon les trajectoires physiques de sorte qu'au premier ordre en dt , on

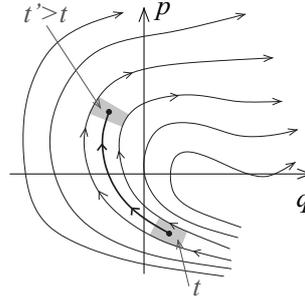


FIGURE 7.2. – **Théorème de Liouville.** On a représenté des trajectoires physiques dans l'espace des phases. L'évolution temporelle conserve le volume dans l'espace des phases.

peut écrire le lien entre les variables $\{q_\alpha, p_\alpha\}$ et $\{q'_\alpha, p'_\alpha\}$

$$\begin{cases} q'_\alpha = q_\alpha(t + dt) \simeq q_\alpha(t) + \dot{q}_\alpha(t) dt = q_\alpha(t) + \frac{\partial H}{\partial p_\alpha} dt \\ p'_\alpha = p_\alpha(t + dt) \simeq p_\alpha(t) + \dot{p}_\alpha(t) dt = p_\alpha(t) - \frac{\partial H}{\partial q_\alpha} dt \end{cases} \quad (7.6)$$

qui peut être vu comme un changement de variables $\{q_\alpha, p_\alpha\} \rightarrow \{q'_\alpha(\{q_\beta, p_\beta\}), p'_\alpha(\{q_\beta, p_\beta\})\}$. Or, ce changement de variables peut être utilisé dans l'intégrale $\Omega(t + dt)$ en introduisant le jacobien J de la transformation $dq'_1 \cdots dq'_n dp'_1 \cdots dp'_n = J dq_1 \cdots dq_n dp_1 \cdots dp_n$ avec

$$J = \left| \begin{array}{cc} \begin{pmatrix} \vdots \\ \dots & \frac{\partial q'_\alpha}{\partial q_\beta} & \dots \\ \vdots \end{pmatrix} & \begin{pmatrix} \vdots \\ \dots & \frac{\partial p'_\alpha}{\partial q_\beta} & \dots \\ \vdots \end{pmatrix} \\ \begin{pmatrix} \vdots \\ \dots & \frac{\partial q'_\alpha}{\partial p_\beta} & \dots \\ \vdots \end{pmatrix} & \begin{pmatrix} \vdots \\ \dots & \frac{\partial p'_\alpha}{\partial p_\beta} & \dots \\ \vdots \end{pmatrix} \end{array} \right|. \quad (7.7)$$

Avec (7.6) et en posant $M = \begin{pmatrix} \left(\frac{\partial^2 H}{\partial q_\beta \partial p_\alpha} \right) & \left(-\frac{\partial^2 H}{\partial q_\beta \partial q_\alpha} \right) \\ \left(\frac{\partial^2 H}{\partial p_\beta \partial p_\alpha} \right) & \left(-\frac{\partial^2 H}{\partial p_\beta \partial q_\alpha} \right) \end{pmatrix}$, on obtient, en développant le déterminant au premier ordre en dt :

$$J = |\det(\mathbf{1} + Mdt)| \stackrel{dt \rightarrow 0}{\simeq} |\mathbf{1} + \text{Tr}(M)dt| = 1, \text{ car } \text{Tr}(M) = \sum_{i=1}^n \frac{\partial^2 H}{\partial q_\alpha \partial p_\alpha} - \frac{\partial^2 H}{\partial p_\alpha \partial q_\alpha} = 0.$$

On a donc bien $\Omega(t + dt) = \Omega(t)$.

Théorème dit "du retour" de Poincaré (admis) :

Si l'évolution du système se fait dans un domaine borné de l'espace des phases, alors toute trajectoire repassera arbitrairement près de sa position initiale si on attend suffisamment longtemps.

II. Crochets de Poisson

Les crochets de Poisson sont un outil mathématique, associé à la géométrie symplectique, qui permettent de réécrire et de comprendre la structure géométrique de la dynamique hamiltonienne dans l'espace des phases.

1. Définition

Soit \mathcal{F} l'espace des fonctions différentiables sur l'espace des phases. Soit $A(q_\alpha, p_\alpha, t)$ et $B(q_\alpha, p_\alpha, t) \in \mathcal{F}$ deux fonctions, ou **observables**, de l'espace des phases. Les **crochets de Poisson** sont définis par

$$\boxed{\{A, B\}_{q,p} = \sum_{\alpha} \frac{\partial A}{\partial q_{\alpha}} \frac{\partial B}{\partial p_{\alpha}} - \frac{\partial B}{\partial q_{\alpha}} \frac{\partial A}{\partial p_{\alpha}}}. \quad (7.8)$$

En pratique, on omet souvent de préciser les indices q, p pour n'écrire que $\{A, B\}$; nous verrons pourquoi au prochain chapitre. Il s'agit d'une application antisymétrique et bilinéaire de A et B

$$\{.,.\} : \mathcal{F} \times \mathcal{F} \rightarrow \mathcal{F}. \quad (7.9)$$

L'antisymétrie signifie

$$\{A, B\} = -\{B, A\}, \quad (7.10)$$

si bien que l'on a $\{A, A\} = 0$. La bilinéarité est la linéarité selon chaque argument :

$$\{A_1 + A_2, B_1 + B_2\} = \{A_1, B_1\} + \{A_2, B_1\} + \{A_1, B_2\} + \{A_2, B_2\}. \quad (7.11)$$

Remarque : Il existe un principe de correspondance entre mécanique classique et quantique qui permet de "quantifier" les théories classiques et qui fait correspondre les commutateurs de la physique quantique aux crochets de Poisson selon la règle :

$$\{A, B\} \rightarrow \frac{1}{i\hbar} [\hat{A}, \hat{B}]. \quad (7.12)$$

Vous verrez que cela fonctionne bien dans de nombreux cas simples que vous connaissez mais cette approche présente des limites.

2. Propriétés mathématiques (TDs)

- Relation de Leibniz :

$$\{A, BC\} = \{A, B\}C + B\{A, C\}. \quad (7.13)$$

- Dérivées

$$\{q_{\alpha}, A\} = \frac{\partial A}{\partial p_{\alpha}} ; \quad \{p_{\alpha}, A\} = -\frac{\partial A}{\partial q_{\alpha}} \quad (7.14)$$

- Crochets fondamentaux

$$\boxed{\{q_\alpha, q_\beta\} = 0; \quad \{p_\alpha, p_\beta\} = 0; \quad \{q_\alpha, p_\beta\} = \delta_{\alpha\beta}}. \quad (7.15)$$

- Identité de Jacobi :

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0. \quad (7.16)$$

- Moment cinétique : $\mathbf{L} = \mathbf{r} \wedge \mathbf{p}$

$$\{L_\alpha, L_\beta\} = \varepsilon_{\alpha\beta\gamma} L_\gamma \quad ; \quad \{L_\alpha, \mathbf{L}^2\} = 0. \quad (7.17)$$

3. Évolution temporelle avec les crochets de Poisson

Ils permettent une réécriture symétrique des équations de Hamilton

$$\boxed{\begin{aligned} \dot{q}_\alpha &= \{q_\alpha, H\} \\ \dot{p}_\alpha &= \{p_\alpha, H\} \end{aligned}}. \quad (7.18)$$

ou sous la forme $\dot{\mathbf{X}} = \{\mathbf{X}, H\}$.

La dérivée totale d'une observable $A(q_\alpha, p_\alpha, t) \in \mathcal{F}$ prend la forme condensée

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + \sum_i \left(\frac{\partial A}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial A}{\partial p_i} \dot{p}_i \right) = \frac{\partial A}{\partial t} + \{A, H\}. \quad (7.19)$$

On notera l'analogie avec le théorème d'Ehrenfest, ou plus encore avec la représentation de Heisenberg, en mécanique quantique.

Exemple : On considère une densité de probabilité $\rho(q_\alpha, p_\alpha, t)$ définie dans l'espace des phases pour une dynamique autonome. Le nombre de points $\delta N(t)$ contenus dans un volume élémentaire $\delta V(t)$ définit $\rho(t) = \delta N(t)/\delta V(t)$. Mais $\delta N(t)$ étant constant au cours de l'évolution et également $\delta V(t)$ par le théorème de Liouville, il vient $\dot{\rho} = 0$, ou l'équation d'évolution locale

$$\frac{\partial \rho}{\partial t} = -\{\rho, H\}. \quad (7.20)$$

En particulier, une distribution stationnaire doit satisfaire à $\{\rho, H\} = 0$ pour laquelle une solution possible simple est de prendre $\rho(q_\alpha, p_\alpha) = f(H(q_\alpha, p_\alpha))$ avec f une fonction à un argument, c'est-à-dire que ρ dépende de la position dans l'espace des phase à travers la fonction de Hamilton. Cet argument se généralise au cas quantique et justifie que les distributions utilisées en physique statistique soient des fonctions de l'Hamiltonien.

4. Liouvillien et opérateur d'évolution

Toujours dans le cas d'un système autonome avec H indépendant du temps, on peut définir un opérateur d'évolution pour écrire une solution formelle de la solution. Cette approche permet également de construire des schémas d'intégration numérique appelés intégrateurs symplectiques qui permettent d'assurer exactement la conservation de l'énergie et sont utilisés en mécanique céleste, physique des plasma ou dynamique moléculaire. Cette approche permet également de mettre en place des méthodes d'approximation perturbatives. La structure symplectique sera présentée au prochain chapitre.

On définit d'abord un opérateur différentiel linéaire appelé liouvillien selon

$$\hat{\mathcal{L}} : A \in \mathcal{F} \rightarrow \hat{\mathcal{L}}A = \{H, A\} \in \mathcal{F} \quad (7.21)$$

Pour une observable sans dépendance explicite en temps, l'équation d'évolution s'intègre alors selon

$$\dot{A} = \{A, H\} = -\hat{\mathcal{L}}A \quad \Rightarrow \quad A(q_\alpha(t), p_\alpha(t)) = e^{-t\hat{\mathcal{L}}}A(q_\alpha(0), p_\alpha(0)) . \quad (7.22)$$

avec $e^{-t\hat{\mathcal{L}}}$ l'opérateur évolution tout à fait analogue à celui de la mécanique quantique $e^{-it\hat{H}/\hbar}$.

Exponentielle avec crochets de Poisson : elle se définit à l'aide de la série de l'exponentielle en considérant les puissances de $\hat{\mathcal{L}}$ comme des compositions / applications successives des crochets de Poisson :

$$e^{-t\hat{\mathcal{L}}} = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} \hat{\mathcal{L}}^n = \mathbf{1} - t\{H, \cdot\} + \frac{t^2}{2!}\{H, \{H, \cdot\}\} - \frac{t^3}{3!}\{H, \{H, \{H, \cdot\}\}\} + \dots \quad (7.23)$$

On voit en dérivant par rapport au temps cette expression que l'opérateur évolution fournit bien la solution à l'équation d'évolution de A

$$\frac{d}{dt}e^{-t\hat{\mathcal{L}}} = -\{H, \cdot\} + t\{H, \{H, \cdot\}\} - \frac{t^2}{2!}\{H, \{H, \{H, \cdot\}\}\} + \dots \quad (7.24)$$

$$= -\left\{H, \left(\mathbf{1} - t\{H, \cdot\} + \frac{t^2}{2!}\{H, \{H, \cdot\}\} + \dots\right)\right\} \quad (7.25)$$

$$= -\hat{\mathcal{L}} e^{-t\hat{\mathcal{L}}} \quad (7.26)$$

Exemple : La chute libre. Nous allons considérer ici le cas de la chute libre. Le Hamiltonien s'écrit $H(z, p) = \frac{p^2}{2m} + mgz$. L'évolution temporelle du système est donc donnée

$$z(t) = e^{-\hat{\mathcal{L}}t}z(0) = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} \hat{\mathcal{L}}^n z(0) \quad ; \quad p(t) = e^{-\hat{\mathcal{L}}t}p(0) = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} \hat{\mathcal{L}}^n p(0) \quad (7.27)$$

On peut calculer de manière successive l'effet du Liouvillien sur les variables z et p

$$\hat{\mathcal{L}}z = \{H, z\} = -\frac{p}{m} \quad (7.28)$$

$$\hat{\mathcal{L}}p = \{H, p\} = mg \quad (7.29)$$

$$\hat{\mathcal{L}}^2 z = -g \quad (7.30)$$

$$\hat{\mathcal{L}}^n z = 0, \quad \forall n \geq 3 \quad (7.31)$$

$$\hat{\mathcal{L}}^n p = 0, \quad \forall n \geq 2 \quad (7.32)$$

Si bien que les développements de $z(t)$ et $p(t)$ donnent

$$z(t) = z(0) + \frac{p(0)}{m}t - \frac{g}{2}t^2 \quad (7.33)$$

$$p(t) = p(0) - mgt \quad (7.34)$$

5. Intégrales premières et systèmes intégrables

Une intégrale première est une observable $K(q_\alpha, p_\alpha, t)$ telle que $\dot{K} = 0$, ie. qui est une constante du mouvement. Si K ne dépend pas explicitement du temps $\dot{K} = 0 \Leftrightarrow \{K, H\} = 0$. Une intégrale première toujours présente pour un système autonome est H . Il peut y avoir des intégrales premières liées à des symétries globales (invariance par translation, rotation, ...) ou par des symétries spécifiques de l'hamiltonien (vecteur de Rung-Lenz pour le problème de Kepler).

Théorème de Poisson : si K_1 et K_2 sont deux intégrales premières, alors $\{K_1, K_2\}$ en est également une.

On montre ce théorème par le calcul, en utilisant l'identité de Jacobi et les propriétés des crochets de Poisson :

$$\begin{aligned} \frac{d}{dt}\{K_1, K_2\} &= \{\{K_1, K_2\}, H\} + \frac{\partial}{\partial t}\{K_1, K_2\} \\ &= \{K_1, \{K_2, H\}\} + \{K_2, \{H, K_1\}\} + \left\{\frac{\partial K_1}{\partial t}, K_2\right\} + \left\{K_1, \frac{\partial K_2}{\partial t}\right\} \\ &= \left\{K_1, \frac{\partial K_2}{\partial t} + \{K_2, H\}\right\} - \left\{K_2, \{K_1, H\} + \frac{\partial K_1}{\partial t}\right\} \\ &= \{\dot{K}_1, K_2\} + \{K_1, \dot{K}_2\} \end{aligned}$$

Cela peut être 0, une combinaison de K_1 et K_2 ou bien une nouvelle intégrale première fonctionnellement indépendante de K_1 et K_2 .

Système intégrable au sens de Liouville : un système à n degrés de liberté est dit intégrable s'il possède n intégrales premières en involution, ie. $\forall i, j, \{K_i, K_j\} = 0$.

Discussion qualitative : Pour les systèmes à 1 degré de liberté conservatif, ils sont toujours intégrables. On retrouve la discussion sur les portraits de phase et la possibilité d'intégrer les équations à partir de l'intégrale première.

En général, il y a peu d'intégrales premières par rapport au nombre de degré de liberté, les trajectoires sont peu contraintes dans l'espace des phases et, malgré le déterminisme, l'évolution temporelle est très difficile à prédire aux temps longs. On a généralement un régime chaotique dans lequel deux trajectoires initialement voisines finissent par "diverger" entre elles de façon exponentielle. Cette discussion sera développée dans les cours de M1.

Lorsque le système est intégrable, l'évolution temporelle est fortement contrainte par les intégrales premières et le mouvement devient alors beaucoup plus régulier et prédictible aux temps longs si le mouvement est borné. Ces systèmes, bien que non « génériques », sont souvent présents en physique. Une description possible peut parfois se faire par l'introduction de transformations canoniques adaptées.

8. Transformations canoniques

I. Définition et fonctions génératrices

1. Définition

On appelle transformation canonique tout changement de variables $(q_\alpha, p_\alpha) \rightarrow (Q_\alpha, P_\alpha)$ avec les dépendances $Q_\alpha(q_\beta, p_\beta, t)$ et $P_\alpha(q_\beta, p_\beta, t)$ tel qu'il préserve la structure "canonique" des équations de la dynamique, c'est-à-dire qu'il existe un nouvel hamiltonien $K(Q_\alpha, P_\alpha, t)$ qui décrit le système selon

$$\boxed{\begin{aligned} \dot{Q}_\alpha &= + \frac{\partial K}{\partial P_\alpha} \\ \dot{P}_\alpha &= - \frac{\partial K}{\partial Q_\alpha} \end{aligned}} \quad (8.1)$$

Cela inclue les transformations dites ponctuelles, ou changements de coordonnées de type $Q_\alpha(q_\beta, t)$ comme ceux utilisés en dynamique lagrangienne, mais aussi des combinaisons beaucoup plus générales des variables q_α et p_α , ce que permet la formulation hamiltonienne qui est beaucoup plus "symétrique" dans ses variables.

Exemples : Considérons la transformation

$$\begin{cases} Q = \frac{1}{\sqrt{2}}(q + p) \\ P = \frac{1}{\sqrt{2}}(p - q) \end{cases} \Leftrightarrow \begin{cases} q = \frac{1}{\sqrt{2}}(Q - P) \\ p = \frac{1}{\sqrt{2}}(Q + P) \end{cases} \Rightarrow \begin{cases} \frac{\partial}{\partial q} = \frac{1}{\sqrt{2}}(-\frac{\partial}{\partial P} + \frac{\partial}{\partial Q}) \\ \frac{\partial}{\partial p} = \frac{1}{\sqrt{2}}(\frac{\partial}{\partial P} + \frac{\partial}{\partial Q}) \end{cases} \quad (8.2)$$

Il est facile de montrer que l'on a les équations

$$\begin{cases} \dot{Q} = \frac{\partial K}{\partial P} \\ \dot{P} = - \frac{\partial K}{\partial Q} \end{cases} \quad (8.3)$$

en prenant simplement le nouvel Hamiltonien comme $K(Q, P) = H(q(Q, P), p(Q, P))$.

Par contre, la transformation $(q_1, p_1, q_2, p_2) \rightarrow (Q_1 = q_2, P_1 = p_1, Q_2 = q_1, P_2 = p_2)$ permet de décrire de manière identique l'espace des phases mais il est impossible de définir un Hamiltonien K tel que la dynamique est donnée par les équations de Hamilton. Elle n'est donc pas canonique.

2. Fonctions génératrices

Les équations de Hamilton dans leur plus grande généralité découlent du principe de moindre action. Ainsi, toute transformation canonique doit être telle que l'action, exprimée dans l'ancien ou le nouveau jeu de variables

$$S[q_\alpha, p_\alpha] = \int_{t_1}^{t_2} dt \left(\sum_\alpha p_\alpha \dot{q}_\alpha - H(q_\alpha, p_\alpha, t) \right) \text{ et } \tilde{S}[Q_\alpha, P_\alpha] = \int_{t_1}^{t_2} dt \left(\sum_\alpha P_\alpha \dot{Q}_\alpha - K(Q_\alpha, P_\alpha, t) \right)$$

soit minimale. D'après l'invariance de jauge du lagrangien, cela signifie que les intégrands ne doivent différer que d'une dérivée totale par rapport au temps

$$\sum_\alpha p_\alpha \dot{q}_\alpha - H(q_\alpha, p_\alpha, t) - \sum_\alpha P_\alpha \dot{Q}_\alpha - K(Q_\alpha, P_\alpha, t) = \frac{dF}{dt}. \quad (8.4)$$

On aura en effet $S - \tilde{S} = F|_{t_2} - F|_{t_1}$ et $\delta S = \delta \tilde{S}$ pour des conditions aux bords fixées (fixer les (q_α, p_α) revient à fixer les (Q_α, P_α)). L'idée est de montrer que l'on peut extraire la transformation canonique et le nouvel hamiltonien de l'existence et de la connaissance de F . Une fonction génératrice est une fonction dont on extrait l'information utile par dérivation.

De façon générale, on s'attend à avoir une fonction F qui dépend des variables $(q_\alpha, p_\alpha, Q_\alpha, P_\alpha, t)$. Cependant, sa différentielle doit s'écrire

$$dF = [K(Q_\alpha, P_\alpha, t) - H(q_\alpha, p_\alpha, t)]dt + \sum_\alpha p_\alpha dq_\alpha - \sum_\alpha P_\alpha dQ_\alpha, \quad (8.5)$$

qui laisse penser qu'elle ne dépend que des q_α, Q_α et du temps t . Par ailleurs, nous recherchons deux ensembles d'équations reliant les anciennes aux nouvelles variables, si bien qu'une fonction à deux variables type position / impulsion suffit. On introduit donc une fonction génératrice notée $F_1(q_\alpha, Q_\alpha, t)$ dont la différentielle sera (8.5) et donc telle que

$$p_\alpha = \frac{\partial F_1}{\partial q_\alpha}(q_\alpha, Q_\alpha, t) \quad (8.6)$$

$$P_\alpha = -\frac{\partial F_1}{\partial Q_\alpha}(q_\alpha, Q_\alpha, t) \quad (8.7)$$

$$K(Q_\alpha, P_\alpha, t) = \tilde{H}(Q_\alpha, P_\alpha, t) + \frac{\partial F_1}{\partial t} \quad (8.8)$$

où $\tilde{H}(Q_\alpha, P_\alpha, t) = H(q_\beta(Q_\alpha, P_\alpha, t), p_\beta(Q_\alpha, P_\alpha, t), t)$ est l'ancien hamiltonien ré-exprimé dans les nouvelles variables. Pour le choix des deux variables de la fonction génératrice, nous pouvons prendre une autre paire que (q_α, Q_α) en considérant d'autres *types* de fonction génératrice. Formellement, les 4 types de fonctions génératrices seront reliés par transformation de Legendre, ce qu'on résumer

$$F_2(q_\alpha, P_\alpha, t) = F_1(q_\alpha, Q_\alpha, t) + \sum_\alpha P_\alpha Q_\alpha \quad (8.9)$$

$$F_3(p_\alpha, Q_\alpha, t) = F_1(q_\alpha, Q_\alpha, t) - \sum_\alpha p_\alpha q_\alpha \quad (8.10)$$

$$F_4(p_\alpha, P_\alpha, t) = F_1(q_\alpha, Q_\alpha, t) + \sum_\alpha P_\alpha Q_\alpha - \sum_\alpha p_\alpha q_\alpha \quad (8.11)$$

Fonction génératrice	Règles de transformation	
$F_1(q_\alpha, Q_\alpha, t)$	$p_\alpha = + \frac{\partial F_1}{\partial q_\alpha}$	$P_\alpha = - \frac{\partial F_1}{\partial Q_\alpha}$
$F_2(q_\alpha, P_\alpha, t)$	$p_\alpha = + \frac{\partial F_2}{\partial q_\alpha}$	$Q_\alpha = + \frac{\partial F_2}{\partial P_\alpha}$
$F_3(p_\alpha, Q_\alpha, t)$	$q_\alpha = - \frac{\partial F_3}{\partial p_\alpha}$	$P_\alpha = - \frac{\partial F_3}{\partial Q_\alpha}$
$F_4(p_\alpha, P_\alpha, t)$	$q_\alpha = - \frac{\partial F_4}{\partial p_\alpha}$	$Q_\alpha = + \frac{\partial F_4}{\partial P_\alpha}$

TABLE 8.1. – Résumé des règles de définition des transformations canoniques à partir des différents types de fonctions génératrices

pour lesquelles les règles de génération des transformations correspondantes sont résumées dans la table. 8.1. Dans tous les cas, le nouvel Hamiltonien s'exprime selon

$$K(Q_\alpha, P_\alpha, t) = \tilde{H}(Q_\alpha, P_\alpha, t) + \frac{\partial F_i}{\partial t}. \quad (8.12)$$

De façon plus général, il est possible de ne faire une transformée de Legendre que selon certaines coordonnées α et pas d'autres, conduisant à une fonction génératrice hybride. Notons que pour obtenir l'expression de la transformation canonique, il faut pouvoir substituer et éliminer des variables. Par exemple, à partir de F_1 , on obtient les expressions de $p_\alpha(q_\beta, Q_\beta, t)$ et $P_\alpha(q_\beta, Q_\beta, t)$ qu'il faut réécrire comme $Q_\alpha(q_\beta, p_\beta, t)$ et $P_\alpha(q_\beta, p_\beta, t)$. De même, les substitutions sont à effectuer pour obtenir K dans les bonnes variables.

II. Structure symplectique et invariance des crochets de Poisson

1. Formulation symplectique de la dynamique hamiltonienne

On passe à des notations compactes $\mathbf{X} = (q_1, \dots, q_n, p_1, \dots, p_n)$ pour un point de l'espace des phases de dimension $2n$. Pour le gradient d'une observable $A \in \mathcal{F}$ et les équations de Hamilton, on peut écrire

$$\frac{\partial A}{\partial \mathbf{X}} = \begin{pmatrix} \vdots \\ \frac{\partial A}{\partial q_\alpha} \\ \vdots \\ \frac{\partial A}{\partial p_\alpha} \\ \vdots \end{pmatrix} \quad \text{et} \quad \dot{\mathbf{X}} = \underbrace{\begin{pmatrix} 0 & 0 & 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & +1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix}}_{\mathbf{J}} \begin{pmatrix} \vdots \\ \frac{\partial H}{\partial q_\alpha} \\ \vdots \\ \frac{\partial H}{\partial p_\alpha} \\ \vdots \end{pmatrix} \Rightarrow \dot{\mathbf{X}} = \mathbf{J} \frac{\partial H}{\partial \mathbf{X}} \quad (8.13)$$

De même, les crochets de Poisson sous forme matricielle prennent la forme

$$\{A, B\}_X = \left(\cdots \quad \frac{\partial A}{\partial q_\alpha} \quad \cdots \quad \cdots \quad \frac{\partial A}{\partial p_\alpha} \quad \cdots \right) \underbrace{\left(\begin{array}{ccc|ccc} 0 & 0 & 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & +1 \\ \hline -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{array} \right)}_{\mathbf{J}} \begin{pmatrix} \vdots \\ \frac{\partial B}{\partial q_\alpha} \\ \vdots \\ \vdots \\ \frac{\partial B}{\partial p_\alpha} \\ \vdots \end{pmatrix} \quad (8.14)$$

$$\Rightarrow \boxed{\{A, B\}_X = \left(\frac{\partial A}{\partial \mathbf{X}} \right)^T \mathbf{J} \left(\frac{\partial B}{\partial \mathbf{X}} \right)}$$

Propriétés de la matrice $\mathbf{J} = \begin{pmatrix} 0 & \mathbb{I}_n \\ -\mathbb{I}_n & 0 \end{pmatrix}$

- $\mathbf{J}^2 = -\mathbb{I}$
- $\mathbf{J}^T = -\mathbf{J} = \mathbf{J}^{-1}$ antisymétrique et orthogonale.
- $\det \mathbf{J} = 1$

Groupe des matrices symplectiques : $\mathbf{M} \in \text{Sp}(2n) \Leftrightarrow \boxed{\mathbf{M}^T \mathbf{J} \mathbf{M} = \mathbf{J}}$. On peut montrer que les matrices réelles symplectiques forment un groupe pour la multiplication des matrices (elles sont donc inversibles), et avec $\det \mathbf{M} = 1$ et $\boxed{\mathbf{M} \mathbf{J} \mathbf{M}^T = \mathbf{J}}$.

2. Application aux transformations canoniques

Soit $\mathcal{T} : \mathbf{X} \rightarrow \mathbf{Y} = (Q_1, \dots, Q_n, P_1, \dots, P_n)$ une transformation inversible dont on note \mathbf{M} la matrice jacobienne, $M_{ij} = \frac{\partial Y_i}{\partial X_j}$. Pour les observables, on note $A(\mathbf{X})$ et $\tilde{A}(\mathbf{Y}) = A(\mathbf{X}(\mathbf{Y}))$, de

sorte que le changement de variables sur le gradient s'écrit de façon compacte $\boxed{\frac{\partial A}{\partial \mathbf{X}} = \mathbf{M}^T \frac{\partial \tilde{A}}{\partial \mathbf{Y}}}$.

On dira que \mathcal{T} est symplectique si sa jacobienne \mathbf{M} est une matrice symplectique.

- \mathcal{T} est symplectique $\Leftrightarrow \mathcal{T}$ laisse les crochets de Poisson invariants.

$$\{A, B\}_X = \left(\frac{\partial A}{\partial \mathbf{X}} \right)^T \mathbf{J} \left(\frac{\partial B}{\partial \mathbf{X}} \right) = \left(\frac{\partial \tilde{A}}{\partial \mathbf{Y}} \right)^T \mathbf{M} \mathbf{J} \mathbf{M}^T \left(\frac{\partial \tilde{B}}{\partial \mathbf{Y}} \right) = \left(\frac{\partial \tilde{A}}{\partial \mathbf{Y}} \right)^T \mathbf{J} \left(\frac{\partial \tilde{B}}{\partial \mathbf{Y}} \right) = \{\tilde{A}, \tilde{B}\}_Y$$

Le plus souvent, on omet donc de préciser le jeu de variable en indices des crochets de Poisson que l'on note plus simplement $\{.,.\}$.

- \mathcal{T} laisse les crochets de Poisson invariants $\Leftrightarrow \mathcal{T}$ laisse les crochets fondamentaux invariants.

Il est clair que le sens \Rightarrow est immédiat :

$$\{Q_\alpha, Q_\beta\}_X = 0, \quad \{P_\alpha, P_\beta\}_X = 0, \quad \{Q_\alpha, P_\beta\}_X = \delta_{\alpha\beta}. \quad (8.15)$$

Pour une transformation générale, qui ne préserve pas forcément les crochets de Poisson, on peut montrer la relation

$$\begin{aligned} \{A, B\}_X = \sum_{\alpha, \beta} \frac{\partial \tilde{A}}{\partial Q_\alpha} \frac{\partial \tilde{B}}{\partial Q_\beta} \{Q_\alpha, Q_\beta\}_X + \frac{\partial \tilde{A}}{\partial P_\alpha} \frac{\partial \tilde{B}}{\partial P_\beta} \{P_\alpha, P_\beta\}_X \\ + \left(\frac{\partial \tilde{A}}{\partial Q_\alpha} \frac{\partial \tilde{B}}{\partial P_\beta} - \frac{\partial \tilde{A}}{\partial P_\beta} \frac{\partial \tilde{B}}{\partial Q_\alpha} \right) \{Q_\alpha, P_\beta\}_X, \end{aligned}$$

qui implique naturellement la réciproque.

- \mathcal{T} est symplectique $\Leftrightarrow \mathcal{T}$ est canonique.

La dérivée totale de \mathbf{Y} s'écrit

$$\dot{\mathbf{Y}} = \mathbf{M}\dot{\mathbf{X}} + \frac{\partial \mathbf{Y}}{\partial t} = \mathbf{M}\mathbf{J} \frac{\partial H}{\partial \mathbf{X}} + \frac{\partial \mathbf{Y}}{\partial t} = \underbrace{\mathbf{M}\mathbf{J}\mathbf{M}^T}_{=\mathbf{J}} \frac{\partial \tilde{H}}{\partial \mathbf{Y}} + \frac{\partial \mathbf{Y}}{\partial t} = \mathbf{J} \frac{\partial K}{\partial \mathbf{Y}}$$

où le dernier terme est l'équation canonique de Hamilton avec K à déterminer. En utilisant la dernière égalité et en l'intégrant partiellement selon \mathbf{Y} , on peut écrire

$$K(\mathbf{Y}) = \tilde{H}(\mathbf{Y}) - \int d\mathbf{Y}' \cdot \left(\mathbf{J} \frac{\partial \mathbf{Y}}{\partial t} \right)_{\mathbf{X}} (\mathbf{X}(\mathbf{Y}'))$$

De façon plus explicite, le nouvel hamiltonien se déduit par la formule

$$\begin{aligned} K(Q_\alpha, P_\alpha) = \tilde{H}(Q_\alpha, P_\alpha) + \sum_\alpha \int \left(\frac{\partial Q_\alpha}{\partial t} \right)_{q,p} (q_\beta(Q', P', t), p_\beta(Q', P', t), t) dP'_\alpha \\ - \int \left(\frac{\partial P_\alpha}{\partial t} \right)_{q,p} (q_\beta(Q', P', t), p_\beta(Q', P', t), t) dQ'_\alpha \end{aligned}$$

qui signifie qu'il faut d'abord prendre la dérivée partielle par rapport au temps des fonctions $Q(q, p, t)$ puis réexprimer le résultat en fonction des (Q, P) et enfin trouver une primitive.

En conclusion, pour vérifier qu'une transformation est canonique, on peut

- s'assurer qu'elle dérive d'une fonction génératrice (canonique par construction).
- vérifier qu'elle laisse invariants les crochets fondamentaux (probablement le plus simple si on vous donne la transformation).
- vérifier que M est symplectique (le plus pédestre).

Exemple : Soit la transformation $(q, p) \rightarrow (Q, P)$ définie par

$$Q = \ln \left(\frac{1}{q} \sin p \right), \quad P = q \cot p. \quad (8.16)$$

On vérifie aisément que cette transformation est bien canonique en calculant les crochets de Poisson fondamentaux.

III. Exemples de transformations canoniques

1. Transformations d'échange

Les transformations les plus simples, qui sont l'identité ou les échanges de variables, s'écrivent facilement à l'aide de fonctions génératrices ayant des produits de variables, comme illustré dans le tableau ci-dessous. On voit en particulier que l'on peut générer l'identité à partir d'une fonction de type 2, ce qui s'avère particulièrement utile dans la théorie des transformations ou la considération de transformations infinitésimales.

Fonction génératrice	Transformation canonique	
$F_1(q_\alpha, Q_\alpha) = \sum_\alpha q_\alpha Q_\alpha$	$p_\alpha = Q_\alpha$	$P_\alpha = -q_\alpha$
$F_2(q_\alpha, P_\alpha) = \sum_\alpha q_\alpha P_\alpha$	$p_\alpha = P_\alpha$	$Q_\alpha = q_\alpha$
$F_3(p_\alpha, Q_\alpha) = \sum_\alpha p_\alpha Q_\alpha$	$q_\alpha = -Q_\alpha$	$P_\alpha = -p_\alpha$
$F_4(p_\alpha, P_\alpha) = \sum_\alpha p_\alpha P_\alpha$	$q_\alpha = -P_\alpha$	$Q_\alpha = p_\alpha$

2. Oscillateur harmonique

On cherche à résoudre l'oscillateur harmonique par une transformation canonique qui rende triviale les équations de la dynamique. Cette stratégie sera développée plus systématiquement au prochain chapitre. Comme il s'agit d'un système conservatif d'Hamiltonien

$$H(q, p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 \quad (8.17)$$

dont le portrait de phase est une ellipse d'équation $E = H(q, p)$ dans l'espace des phases. Il est donc intuitif de rechercher une transformation sous la paramétrisation

$$\begin{cases} q = \frac{1}{m\omega} f(P) \sin Q \\ p = f(P) \cos Q \end{cases} \Rightarrow \tilde{H} = \frac{f(P)^2}{2m} \quad (8.18)$$

En cherchant une fonction génératrice de type $F_1(q, Q)$, on peut intégrer les règles de transformation pour trouver la forme

$$F_1(q, Q) = \frac{m\omega}{2} q^2 \cot Q \quad (8.19)$$

dont on tire $f(P) = \sqrt{2m\omega P}$ et $\tilde{H} = \omega P$. Comme Q est cyclique, P est constante (c'est E/ω !) et on a $\dot{Q} = \omega$, soit $Q(t) = \omega t + Q_0$ puis

$$\begin{cases} q = \sqrt{\frac{2P}{m\omega}} \sin(\omega t + Q_0) \\ p = \sqrt{2m\omega P} \cos(\omega t + Q_0) \end{cases} \quad (8.20)$$

Bien qu'on n'ait rien appris de nouveau sur l'oscillateur harmonique, on retient l'idée de la stratégie qui est de construire une transformation canonique dans laquelle les nouvelles équations de la dynamique sont triviales à résoudre (ici, $\dot{P} = 0$ et $\dot{Q} = \text{cste}$). La difficulté étant reportée dans la recherche de la transformation canonique.

3. Transformations infinitésimales (voir TD)

Jusque là nous avons considéré une vue *passive* des transformations canoniques $(q, p) \rightarrow (Q, P)$. On passe d'une représentation de l'espace des phases à une autre. Il est possible d'avoir un point de vue *actif* des transformations canoniques, c'est-à-dire en opérant des transformations à partir et vers le même espace des phases. C'est le cas de la rotation : on peut voir cette transformation de manière passive (introduction d'un nouveau système de coordonnées) ou de manière active via le déplacement d'un point dans l'espace.

Introduisons le concept de transformation canonique infinitésimale via une vue active. Pour cela nous allons considérer une transformation dépendant continûment d'un paramètre ϵ .

$$Q_i = Q_i(q, p, \epsilon), \quad P_i = P_i(q, p, \epsilon), \quad (8.21)$$

tel que l'on trouve la transformation identité pour $\epsilon = 0$, c'est-à-dire $Q_i(q, p, 0) = q_i$ et $P_i(q, p, 0) = p_i$. Le paramètre ϵ contrôlant la transformation, nous allons considérer un transformation infinitésimale c'est-à-dire pour $\epsilon = \delta\epsilon$ un paramètre infinitésimal. Pour définir la transformation, cherchons une fonction génératrice $G_2(q, P)$ de type 2. Comme la transformation est infinitésimale, cette fonction génératrice est proche de la transformation identité

$$G_2(q, P) = \sum_i q_i P_i + \delta\epsilon G(q, P). \quad (8.22)$$

Les règles de transformations donnent

$$\begin{aligned} Q_i &= \frac{\partial G_2}{\partial P_i} = q_i + \delta\epsilon \frac{\partial G}{\partial P_i}(q, P) \\ p_i &= \frac{\partial G_2}{\partial q_i} = P_i + \delta\epsilon \frac{\partial G}{\partial q_i}(q, P) \end{aligned} \quad (8.23)$$

Ces deux équations permettent d'écrire

$$\begin{aligned} Q_i &= q_i + \delta\epsilon \frac{\partial G}{\partial P_i}(q, P) = q_i + \delta q_i \\ P_i &= p_i - \delta\epsilon \frac{\partial G}{\partial q_i}(q, P) = p_i + \delta p_i \end{aligned} \quad (8.24)$$

En ne gardant que les termes linéaires en $\delta\epsilon$ on peut finalement exprimer les variations δq_i et δp_i en fonction de (q, p) uniquement

$$\begin{aligned} \delta q_i &= \delta\epsilon \frac{\partial G}{\partial p_i}(q, p) \\ \delta p_i &= -\delta\epsilon \frac{\partial G}{\partial q_i}(q, p) \end{aligned} \quad (8.25)$$

Lors d'une transformation active la valeur d'une fonction de l'espace des phases $F(q, p)$ change

$$\begin{aligned}\delta F &= F(q + \delta q, p + \delta p) - F(q, p) = \sum_i \frac{\partial F}{\partial q_i} \delta q_i + \frac{\partial F}{\partial p_i} \delta p_i \\ &= \delta \epsilon \sum_i \left(\frac{\partial F}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q_i} \right) = \delta \epsilon \{F, G\}\end{aligned}\quad (8.26)$$

Ainsi la fonction G permet de définir de manière unique la transformation infinitésimale, on parle de générateur d'une transformation.

Exemples : Nous allons considérer l'opération de translation dans le temps $\delta \epsilon = \delta t$. A partir des équations de Hamilton on a

$$\delta q_i = q_i(t + \delta t) - q_i = \dot{q}_i \delta t = \frac{\partial H}{\partial p_i} \delta t \quad (8.27)$$

$$\delta p_i = p_i(t + \delta t) - p_i = \dot{p}_i \delta t = -\frac{\partial H}{\partial q_i} \delta t \quad (8.28)$$

On voit par le biais de ces relations que le Hamiltonien H est le générateur d'opération de translation dans le temps. Ainsi l'évolution temporelle peut-être vu comme une succession continue de transformations canoniques infinitésimales. D'autre part l'évolution temporelle d'une fonction F de l'espace des phases indépendante du temps $F(q, p)$ est bien donnée par

$$\delta F = \delta t \{F, H\}. \quad (8.29)$$

On remarquera que l'on retrouve la conservation de l'énergie lorsque le Hamiltonien est indépendant du temps $\delta H = \delta t \{H, H\} = 0$.

Nous allons également considérer le cas de la rotation autour de l'axe Oz . Pour toute rotation d'angle θ , les coordonnées x et y s'écrivent

$$\begin{aligned}x' &= x \cos \theta - y \sin \theta \\ y' &= x \sin \theta + y \cos \theta\end{aligned}\quad (8.30)$$

Une relation similaire permet de relier les moments p_x, p_y aux moments $p_{x'}, p_{y'}$. si bien que pour une variation infinitésimale $\delta \theta$ on a

$$\begin{aligned}\delta x &= -y \delta \theta, & \delta y &= x \delta \theta \\ \delta p_x &= -p_y \delta \theta, & \delta p_y &= p_x \delta \theta\end{aligned}\quad (8.31)$$

À partir de ces relations, on peut vérifier que le générateur de la transformation est donnée par le moment cinétique le long de l'axe Oz .

$$G = xp_y - yp_x = L_z \quad (8.32)$$

si bien que la variation d'une fonction de l'espace des phases sous l'effet d'une rotation autour de l'axe Oz est donnée par

$$\delta F = \delta \theta \{F, L_z\} \quad (8.33)$$

9. Équation de Hamilton-Jacobi

I. La fonction principale de Hamilton

Les équations d'Euler-Lagrange peuvent être dérivés du principe de Hamilton pour lequel on recherche un extremum de l'action vue comme une fonctionnelle du chemin suivi

$$\mathcal{S}[q(t)] = \int_{t_1}^{t_2} \mathcal{L}(q(t), \dot{q}(t), t) dt \quad (9.1)$$

La solution des équations d'Euler-Lagrange noté $\bar{q}(t)$ donne donc une valeur extrême à l'action $S = \mathcal{S}[\bar{q}(t)]$. Puisque la variation de l'action aux temps initiaux t_1 et finaux t_2 est nulle cette action est une fonction des temps t_1 et t_2 et des positions initiales et finales $q_i(t_1)$ et $q_i(t_2)$ que l'on notera simplement ici $q_i^{(1)}$ et $q_i^{(2)}$. Ceci définit la fonction principale de Hamilton $S(q_i^{(1)}, q_i^{(2)}, t_1, t_2)$. On remarquera que par définition pour $t_1 = t_2$ on a $S = 0$.

Pour déterminer la dépendance de la fonction principale de Hamilton en fonction des coordonnées, on effectue une petite variation de S par rapport aux positions finales et positions initiales

$$\delta' S = \sum_i \left[\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta' q_i \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \sum_i \left(\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \right) \delta' q_i(t) dt \quad (9.2)$$

La variation δ' est définie par rapport à une variation des positions initiales et finales en gardant le chemin $q(t)$ toujours solution des équations d'Euler-Lagrange si bien que le second terme est nul on a donc

$$\delta' S = \sum_i p_i(t_2) \delta q_i(t_2) - p_i(t_1) \delta q_i(t_1) \quad (9.3)$$

Ceci permet d'obtenir les moments aux instants initiaux et finaux à partir de la fonction principale de Hamilton.

$$p_i^{(2)} = \frac{\partial S}{\partial q_i^{(2)}}(q_i^{(1)}, q_i^{(2)}, t_1, t_2), \quad p_i^{(1)} = -\frac{\partial S}{\partial q_i^{(1)}}(q_i^{(1)}, q_i^{(2)}, t_1, t_2) \quad (9.4)$$

Ainsi la connaissance de la fonction principale de Hamilton permet en principe de déterminer complètement le mouvement. En effet pour un instant t donné on peut trouver une relation entre les positions q à l'instant t , les positions initiales q_0 et moments initiaux p_0 .

$$p_0 = -\frac{\partial S}{\partial q_0}(q, q_0, t, t_0), \quad (9.5)$$

que l'on peut inverser pour trouver $q(t, t_0, q_0, p_0)$. De même l'équation

$$p = \frac{\partial S}{\partial q}(q, q_0, t, t_0), \quad (9.6)$$

permet d'obtenir $p(t, t_0, q_0, p_0)$ si l'on y injecte la solution $q(t, t_0, q_0, p_0)$.

Pour déterminer l'évolution temporelle de la fonction principale de Hamilton, nous allons tout d'abord calculer la dérivée totale par rapport au temps de S ,

$$\frac{dS}{dt_2} = \frac{\partial S}{\partial t_2} + \sum_i \frac{\partial S}{\partial q_i^{(2)}} \dot{q}_i^{(2)}. \quad (9.7)$$

Au cours de la variation temporelle nous avons supposé que les variations des conditions finales $q_i^{(2)} = q_i(t_2)$ s'effectuaient dans la continuité de la trajectoire reliant $q_i(t_1)$ à $q_i(t_2)$. Ainsi la dérivée par rapport au temps de S donne

$$\frac{dS}{dt_2} = \frac{\partial S}{\partial t_2} + \sum_i p_i^{(2)} \dot{q}_i^{(2)} = \mathcal{L}(q_i^{(2)}, \dot{q}_i^{(2)}, t_2) \quad (9.8)$$

Au final grâce à la définition de la fonction de Hamilton et en effectuant une variation de la fonction principale de Hamilton par rapport au temps initial on obtient les deux équations.

$$\begin{aligned} \frac{\partial S}{\partial t_2} + H(q_i^{(2)}, p_i^{(2)}, t_2) &= 0 \\ \frac{\partial S}{\partial t_1} - H(q_i^{(1)}, p_i^{(1)}, t_1) &= 0 \end{aligned} \quad (9.9)$$

Ce résultat se résume grâce à l'expression de la différentielle totale exacte de S

$$dS = \sum_i p_i^{(2)} dq_i^{(2)} - \sum_i p_i^{(1)} dq_i^{(1)} - H^{(2)} dt_2 + H^{(1)} dt_1 \quad (9.10)$$

En principe il est possible de résoudre le système d'équations (9.9) et (9.4) pour obtenir la fonction principale de Hamilton. Cette approche constitue une nouvelle méthode de résolution d'un problème mécanique. Toutefois la recherche directe de $S(q_i^{(1)}, q_i^{(2)}, t_1, t_2)$ n'est pas une méthode pratique. Nous allons voir qu'une approche similaire introduite par Jacobi, beaucoup plus pratique, peut-être mise en place : c'est la méthode de Hamilton-Jacobi.

II. Équation de Hamilton-Jacobi

Dans l'approche de Hamilton-Jacobi, nous allons considérer une seule des équations aux dérivées partielles (EDP) précédentes et ne considérer que la dépendance selon le temps final t et la position finale q de la fonction principale de Hamilton $S(q, t)$. En remarquant qu'au temps final le moment p dérive de la fonction principale de Hamilton on en déduit l'équation de Hamilton-Jacobi

$$\frac{\partial S}{\partial t}(q_1, \dots, q_n, t) + H(q_1, \dots, q_n, \frac{\partial S}{\partial q_1}, \dots, \frac{\partial S}{\partial q_n}, t) = 0 \quad (9.11)$$

Il n'est pas nécessaire de trouver l'intégrale *générale* de cette équation mais seulement l'intégrale *complète* qui est en général beaucoup plus facilement calculable.

Définition Soit $u(x_1, \dots, x_n)$ une fonction de n variables x_i , $i = 1, \dots, n$. Une équation aux dérivées partielles (EDP) d'ordre 1 est une équation de la forme

$$F(x_1, \dots, x_n, u, \frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_n}) = 0 \quad (9.12)$$

L'intégrale *générale* d'une EDP dépend de fonctions arbitraires. Ces fonctions sont déterminées par les conditions aux limites. On appelle l'intégrale *complète* d'une EDP d'ordre 1 et de n variables *une* solution qui possède n constantes d'intégration.

L'équation de Hamilton-Jacobi (Eq. 9.11) dépend de $n + 1$ variables q_1, \dots, q_n, t . En principe l'intégrale complète dépend alors de $n + 1$ constantes d'intégration. Cependant S n'apparaît pas directement dans l'équation de Hamilton-Jacobi, il y a donc une constante additive A triviale

$$S' = S(q_1, \dots, q_n, t, \alpha_1, \dots, \alpha_n) + A \quad (9.13)$$

où $\alpha_1, \dots, \alpha_n$ sont des constantes arbitraires. La solution complète de Jacobi $S(q, t, \alpha)$ qui dépend de n constantes permet de trouver la solution du problème. Pour cela on introduit les quantités β_i définies par

$$\beta_i = \left(\frac{\partial S}{\partial \alpha_i} \right)_{q,t} \quad (9.14)$$

Dans cette approche on considère les coordonnées q_i et les constantes α_i comme variables indépendantes. On montre que les β_i sont des constantes du mouvement

$$\begin{aligned} \frac{d\beta_i}{dt} &= \frac{d}{dt} \left(\frac{\partial S}{\partial \alpha_i} \right)_{q,t} = \frac{\partial}{\partial t} \left(\frac{\partial S}{\partial \alpha_i} \right)_{q,t} + \sum_k \frac{\partial}{\partial q_k} \left(\frac{\partial S}{\partial \alpha_i} \right)_{q,t} \dot{q}_k \\ &= \frac{\partial}{\partial \alpha_i} \left(\frac{\partial S}{\partial t} \right)_{q,t} + \sum_k \left(\frac{\partial p_k}{\partial \alpha_i} \right)_{q,t} \frac{\partial H}{\partial p_k} \\ &= \frac{\partial}{\partial \alpha_i} \left(\frac{\partial S}{\partial t} + H \left(q, \frac{\partial S}{\partial q}, t \right) \right) = 0 \end{aligned} \quad (9.15)$$

On trouve la dernière égalité simplement en remarquant que

$$\left(\frac{\partial H}{\partial \alpha_i} \right)_{q,t} = \sum_k \frac{\partial H}{\partial p_k} \frac{\partial p_k}{\partial \alpha_i}. \quad (9.16)$$

La définition des β_i donne les n équations

$$\beta_i = \frac{\partial S}{\partial \alpha_i}(q_1, \dots, q_n, t, \alpha_1, \dots, \alpha_n) \quad (9.17)$$

que l'on peut inverser pour obtenir la solution générale du problème

$$q_i = q_i(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n, t) \quad (9.18)$$

qui dépend donc bien de $2n$ constantes. On peut ensuite dériver simplement les moments pour compléter la solution

$$p_i = \frac{\partial S}{\partial q_i} = p_i(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n, t) \quad (9.19)$$

III. Systèmes conservatifs et séparables

En général la recherche d'une solution complète à l'équation de Hamilton-Jacobi est très compliquée. Pour les systèmes conservatifs il est toutefois possible de simplifier un peu le problème. Dans le cas d'un système conservatif, l'Hamiltonien ne dépend pas du temps $H(q, p) = E = \text{cste}$ et la solution de l'équation de Hamilton-Jacobi doit s'écrire sous la forme

$$S(q_1, \dots, q_n, t) = W(q_1, \dots, q_n) - Et \quad (9.20)$$

La fonction $W(q_1, \dots, q_n)$ est la fonction caractéristique de Hamilton ¹ qui est alors déterminée par l'équation de Hamilton-Jacobi

$$H\left(q_1, \dots, q_n, \frac{\partial W}{\partial q_1}, \dots, \frac{\partial W}{\partial q_n}\right) = E \quad (9.21)$$

La solution complète de cette équation dépend de n constantes $W(q_1, \dots, q_n, \alpha_1, \dots, \alpha_n)$. D'une manière générale, l'énergie qui est une constante peut s'exprimer en fonction des n constantes α_i , $E = E(\alpha_1, \dots, \alpha_n)$. On remarquera qu'il est toujours possible de choisir librement les constantes et on peut prendre l'énergie E comme une des constantes par exemple $E = \alpha_n$.²

Même dans le cas d'un système conservatif si le nombre de degrés de liberté est supérieur à un, trouver une solution à l'équation de Hamilton-Jacobi n'est pas garanti. Il existe cependant un certain nombre de système pour lequel la solution s'écrit sous la forme séparable

$$S(q_1, \dots, q_n, \alpha_1, \dots, \alpha_n, t) = \sum_k W_k(q_k, \alpha_1, \dots, \alpha_n) - E(\alpha_1, \dots, \alpha_n)t, \quad (9.22)$$

Si un système possède une telle solution on parle alors de système séparable. Le problème de Kepler fait parti des systèmes séparables. D'une manière générale les systèmes intégrables sont séparables pour peu que l'on trouve le bon jeu de coordonnées (ce qui peut être une tache non triviale!)

Exemple : le problème du potentiel à force centrale Nous allons montrer ici que le problème du potentiel à force centrale constitue un problème séparable. Dans un repère sphérique la fonction de Lagrange s'écrit

$$\mathcal{L}(r, \theta, \dot{r}, \dot{\theta}, \dot{\varphi}) = \frac{m}{2} \left(\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\varphi}^2 \right) - V(r). \quad (9.23)$$

1. On remarquera que l'on peut voir la fonction caractéristique de Hamilton comme la transformée de Legendre de la fonction principale de Hamilton par rapport au temps.

2. En réalité la solution complète de l'EDP 9.21 ne dépend que de $n - 1$ constantes mais elle dépend aussi de l'énergie E qui intervient dans l'équation $W(q_1, \dots, q_n, \alpha_1, \dots, \alpha_{n-1}, E)$ On peut choisir $\alpha_n = E$ ou plus généralement $\alpha_n = \alpha_n(\alpha_1, \dots, \alpha_{n-1}, E)$.

La fonction de Hamilton s'écrit alors

$$\mathcal{H}(r, \theta, p_r, p_\theta, p_\varphi) = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\varphi^2}{2mr^2 \sin^2 \theta} + V(r), \quad (9.24)$$

où les moments conjugués sont définis par

$$p_r = m\dot{r}, \quad p_\theta = mr^2\dot{\theta}, \quad p_\varphi = mr^2 \sin^2 \theta \dot{\varphi}. \quad (9.25)$$

Ce problème est conservatif et nous allons rechercher une solution séparable, ainsi la fonction caractéristique de Hamilton s'écrit

$$W(r, \theta, \varphi) = W_r(r) + W_\theta(\theta) + W_\varphi(\varphi) \quad (9.26)$$

Nous savons que φ est une coordonnée cyclique, ainsi le moment p_φ est une constante que l'on notera α_φ . Ainsi on a

$$p_\varphi = \frac{\partial W}{\partial \varphi} = \frac{dW_\varphi}{d\varphi} = \alpha_\varphi, \quad (9.27)$$

Finalement la fonction caractéristique associée à la coordonnée φ s'écrit simplement

$$W_\varphi = \alpha_\varphi \varphi \quad (9.28)$$

Cette forme est en fait générale à toute coordonnée cyclique. En utilisant cette expression on obtient alors l'équation de Hamilton-Jacobi

$$\left(\frac{dW_r}{dr} \right)^2 + \frac{1}{r^2} \left[\left(\frac{dW_\theta}{d\theta} \right)^2 + \frac{\alpha_\varphi^2}{\sin^2 \theta} \right] = 2m(E - V(r)) \quad (9.29)$$

On notera que le terme entre crochets est une fonction de θ uniquement et que le reste de l'équation ne dépend que de r si bien que ce terme ne peut être qu'une constante que l'on notera α_θ^2 . Ainsi on a

$$\left(\frac{dW_\theta}{d\theta} \right)^2 + \frac{\alpha_\varphi^2}{\sin^2 \theta} = \alpha_\theta^2 \quad (9.30)$$

$$\left(\frac{dW_r}{dr} \right)^2 + \frac{\alpha_\theta^2}{r^2} = 2m(E - V(r)). \quad (9.31)$$

Nous n'allons pas résoudre ces deux équations mais ceci démontre que le problème de Kepler est bien un problème séparable car une telle solution existe.

IV. Transformation canonique de Jacobi

Nous allons voir que la solution complète de l'équation de Hamilton-Jacobi possède une interprétation en terme de la transformation canonique de Jacobi. Nous avons vu au chapitre précédent qu'une transformation canonique $(q, p) \rightarrow (Q, P)$ peut-être définie à partir d'une

fonction génératrice. Cette dernière nous permet de définir le nouvel Hamiltonien K . Il est peut-être judicieux de trouver une transformation canonique tel que le nouvel Hamiltonien soit nul $K = 0$ car dans ce cas la dynamique est particulièrement simple

$$Q_i = \text{cst}, \quad P_i = \text{cst}. \quad (9.32)$$

On recherche une telle transformation par le biais d'une fonction génératrice de type 2 $F_2(q, P, t)$. Les règles de transformation sont données par

$$p_i = \frac{\partial F_2}{\partial q_i}, \quad Q_i = \frac{\partial F_2}{\partial P_i}, \quad K = H + \frac{\partial F_2}{\partial t}. \quad (9.33)$$

En utilisant la première et dernière équation pour $K = 0$ on obtient

$$H\left(q, \frac{\partial F_2}{\partial q}, t\right) + \frac{\partial F_2}{\partial t} = 0 \quad (9.34)$$

qui n'est autre que l'équation de Hamilton-Jacobi et la solution n'est autre que la solution complète $F_2 = S$. Ainsi la solution complète de l'équation de Hamilton-Jacobi est la fonction génératrice de la transformation

$$(q_1, \dots, q_n, p_1, \dots, p_n) \rightarrow (\beta_1, \dots, \beta_n, \alpha_1, \dots, \alpha_n). \quad (9.35)$$

où les constantes α_i sont les nouveaux moments et les constantes β_i sont les nouvelles coordonnées. Avec ce nouveau jeu de coordonnées le problème est devenu complètement statique.

V. Variables angle-action

1. Principe

Au paragraphe précédent nous avons recherché la fonction génératrice de la transformation canonique permettant d'éliminer toute évolution temporelle dans le système. Cette fonction génératrice est la fonction principale de Hamilton. Cette approche radicale dans la résolution d'un problème est cependant peu pratique car il existe peu de système pour lesquels il est possible de calculer la fonction principale de Hamilton. Une approche plus judicieuse est de chercher une transformation canonique pour laquelle les moments sont constants et les coordonnées varient avec le temps. Cette approche est particulièrement puissante pour les systèmes quasi-périodiques comme le problème à N corps en mécanique céleste car il permet de faire apparaître de manière naturelle les différentes échelles de temps du système et de développer des schémas perturbatif très puissants. D'une manière générale, ce type de méthode est propice aux approximations pour des systèmes non-intégrables proches d'un système intégrable. Les moments ont alors une variation lente comparée aux variations rapides des coordonnées.

On cherche une transformation canonique indépendante du temps

$$(q_1, \dots, q_n, p_1, \dots, p_n) \rightarrow (\theta_1, \dots, \theta_n, J_1, \dots, J_n) \quad (9.36)$$

où θ_i et J_i sont les nouvelles coordonnées et moments également appelés angles et actions. Pour que les moments soient des constantes du mouvement $J_i = \text{cst}$, il est nécessaire que le nouvel Hamiltonien $K = H$ soit indépendant des angles.

$$H = H(J_1, \dots, J_n). \quad (9.37)$$

Dans ce cas les équations de Hamilton donnent

$$\dot{J}_i = 0, \quad \dot{\theta}_i = \frac{\partial H}{\partial J_i} = \omega_i(J_1, \dots, J_n). \quad (9.38)$$

où les ω_i sont les fréquences du système. Leur identification permet de mettre à jour les différentes échelles de temps dans la dynamique (dynamique rapide/lente, effets de résonance) particulièrement utile dans le cadre de la théorie des perturbations. Ces équations s'intègrent simplement pour donner

$$J_i = J_i^0, \quad \theta_i = \omega_i(J_1, \dots, J_n)t + \theta_i^0, \quad (9.39)$$

Les actions sont donc bien des constantes du mouvement et les angles varient linéairement avec le temps.

2. Systèmes multiplement périodiques

Les variables angle-action permettent de décrire les systèmes multiplement périodiques. C'est-à-dire des systèmes pour lesquels il existe un jeu de coordonnées individuellement périodiques mais avec différentes périodes. Le système dans son ensemble n'est alors périodique que si l'ensemble des périodes sont commensurables, on parle alors de dégénérescence. Il est évident que puisque toutes les coordonnées q_i sont périodiques il doit exister pour chacune de ces coordonnées une constante du mouvement, par exemple l'amplitude dans le cas d'une oscillation. L'ensemble de ces n constantes sont alors utilisées comme variable d'action J_i . Un tel choix n'est évidemment pas unique, mais celui qui est présenté ici est tel que les pulsations $\omega_i = \frac{\partial H}{\partial J_i}$ correspondent aux pulsations du système.

Nous devons tout d'abord définir la notion de mouvement périodique. Il existe en effet deux types de mouvement périodiques : la libration et la rotation (Fig. 9.1). Dans le cas de la libration comme par exemple l'oscillateur harmonique, le mouvement est caractérisé par des orbites fermées dans l'espace des phases (q, p) . Dans le cas de la rotation, l'état du système est une fonction périodique de la variable q (par exemple dans le cas d'un angle) si bien que dans l'espace des phases (q, p) le moment p devient une fonction périodique de la coordonnée q qui n'est pas bornée. On remarquera que l'oscillateur non-linéaire possède les deux types de

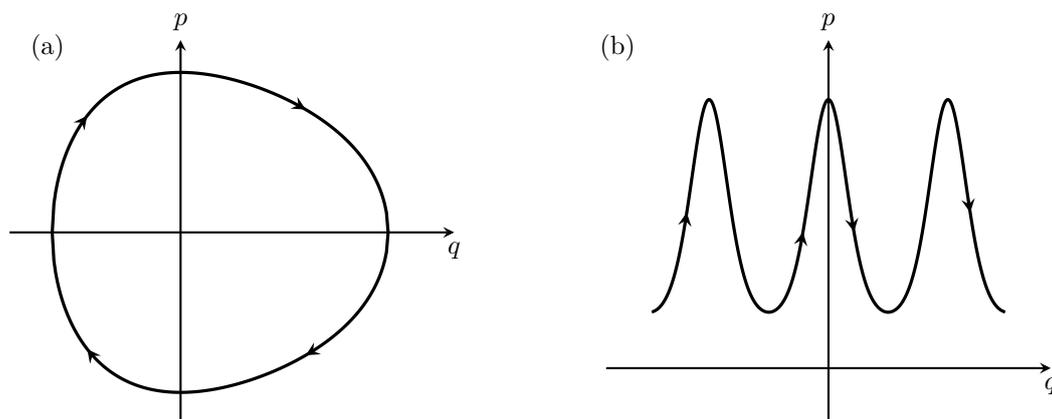


FIGURE 9.1. – (a) mouvement de libration, l'orbite dans l'espace est fermée. (b) la rotation, l'état du système dépend périodiquement de la coordonnée q .

mouvement. Nous allons définir les variables d'actions dans le contexte des librations mais il est possible de généraliser aux rotations.

Pour illustrer la notion de système multiplement périodique, nous allons tout d'abord considérer le cas de deux oscillateurs harmoniques non couplés de coordonnées q_i , de moments p_i de fréquences ω_i , $i = 1, 2$. L'Hamiltonien du système s'écrit

$$H(q_1, q_2, p_1, p_2) = \frac{1}{2}p_1^2 + \frac{1}{2}\omega_1^2 q_1^2 + \frac{1}{2}p_2^2 + \frac{1}{2}\omega_2^2 q_2^2 \quad (9.40)$$

On montre trivialement que la solution de ce problème s'écrit

$$q_i(t) = a_i \cos(\omega_i t + \varphi_i), \quad (9.41)$$

chaque coordonnée est donc périodique mais la trajectoire du système correspond aux courbes de Lissajous. Ces courbes ne sont fermées que si les fréquences ω_1 et ω_2 sont commensurables c'est à dire $\exists n, m \in \mathbb{N}$ tel que $n\omega_1 = m\omega_2$ (Fig. 9.2).

Nous allons maintenant considérer le cas plus complexe du problème d'une masse m dans un champ de force centrale exprimée dans un repère sphérique (r, θ, φ) . Dans ce cas l'Hamiltonien s'écrit

$$H = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\varphi^2}{2mr^2 \sin^2 \theta} + V(r) \quad (9.42)$$

Nous allons considérer un potentiel attractif $V(r)$ qui possède des orbites liées mais pas nécessairement fermées.³ Comme le montre la figure 9.3, les coordonnées $r(t)$ et $\theta(t)$ sont périodiques et la coordonnée $\varphi(t)$ est une fonction croissante du temps. En général, la trajectoire n'est pas fermée sauf pour le potentiel de Kepler qui est tel que la période d'oscillation de $\theta(t)$ est identique à la période d'oscillation de $r(t)$. Si l'on regarde maintenant le mouvement

3. Le théorème de Bertrand montre que seul le potentiel harmonique et le potentiel de Kepler ont toutes leurs orbites liées fermées.

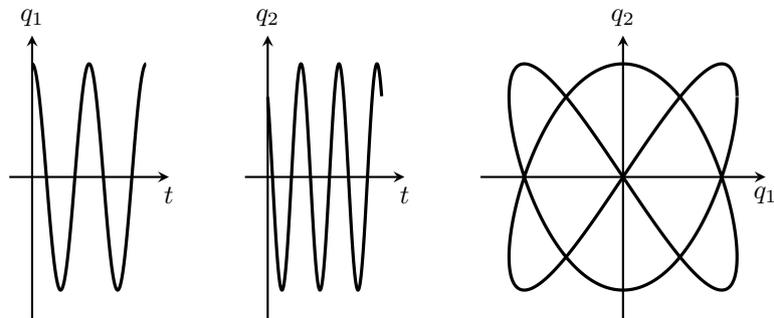


FIGURE 9.2. – Exemple de courbe de Lissajou générée par deux oscillateurs harmoniques dont les fréquences sont commensurables $2\omega_2 = 3\omega_1$. Dans le cas de deux fréquences non-commensurable, la trajectoire remplit l'ensemble du rectangle $|q_1| \leq a_1$ et $|q_2| \leq a_2$

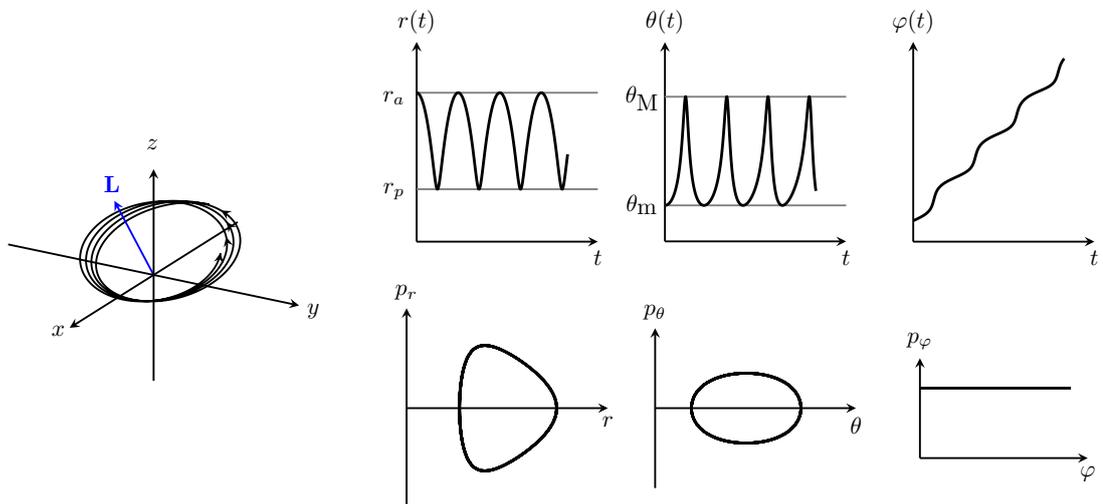


FIGURE 9.3. – Problème à force centrale, la séparabilité permet d'étudier chaque coordonnées dans l'espace des phases réduite de chaque coordonnée et moment.

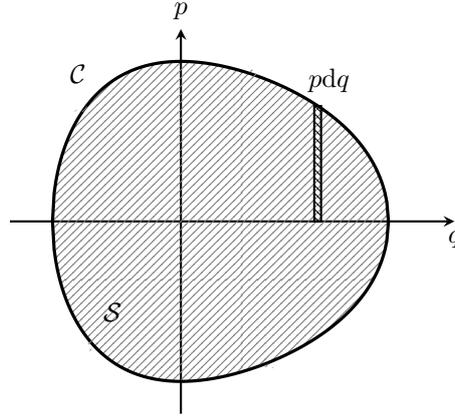


FIGURE 9.4. – Variable d'action définie comme l'intégrale curvilligne le long de l'orbite \mathcal{C} , elle correspond à la surface \mathcal{S} incluse dans la courbe \mathcal{C} .

dans l'espace des phases, on identifie parfaitement les coordonnées r et θ comme étant des librations et la coordonnée φ comme une rotation.

Nous allons donc ici considérer un système multiplement périodiques, conservatif et complètement séparable. Pour un tel système la fonction principale de Hamilton s'écrit

$$S(q_1, \dots, q_n, \alpha_1, \dots, \alpha_n, t) = \sum_k W_k(q_k, \alpha_1, \dots, \alpha_n) - E(\alpha_1, \dots, \alpha_n)t. \quad (9.43)$$

Ainsi pour un système complètement séparable le moment p_i s'écrit comme une fonction de la coordonnée q_i uniquement

$$p_i = \frac{\partial W}{\partial q_i} = \frac{\partial W_i}{\partial q_i}(q_i, \alpha_1, \dots, \alpha_n) \quad (9.44)$$

On définit les variables d'actions par

$$J_i = \frac{1}{2\pi} \oint_{\mathcal{C}_i} p_i dq_i = \frac{1}{2\pi} \iint_{\mathcal{S}_i} dq_i dp_i \quad (9.45)$$

où l'intégration s'effectue sur une période de la variable q_i le long de l'orbite \mathcal{C}_i dans l'espace des phases. Dans le cas des librations l'orbite dans l'espace des phases est une courbe fermée si bien que l'action correspond à la surface \mathcal{S}_i incluse dans la courbe \mathcal{C}_i (Fig. 9.4). Cette surface est bien une constante du mouvement. On peut aussi montrer que les actions sont constantes en remarquant que les actions sont fonctions uniquement des constantes α_i ,

$$J_i = \frac{1}{2\pi} \oint_{\mathcal{C}} \frac{\partial W_i}{\partial q_i}(q_i, \alpha_1, \dots, \alpha_n) dq_i = J_i(\alpha_1, \dots, \alpha_n), \quad (9.46)$$

ce qui montre bien que les actions sont des constantes du mouvement.

Pour finaliser la transformation angle-action et définir les angles, nous allons montrer que la fonction caractéristique de Hamilton exprimée en fonction des actions constitue la fonction

génératrice de la transformation angle-action. Ainsi on considère la fonction génératrice de deuxième espèce $F_2(q_1, \dots, q_n, J_1, \dots, J_n)$. Les règles de transformations donnent

$$p_i = \frac{\partial F_2}{\partial q_i}, \quad \theta_i = \frac{\partial F_2}{\partial J_i}, \quad K = H. \quad (9.47)$$

La première équation qui définit le moment conjugué est vérifiée facilement si la fonction génératrice est la fonction caractéristique de Hamilton $F_2 = W(q_1, \dots, q_n, J_1, \dots, J_n)$. Pour le confirmer il suffit de vérifier que la définition de θ_i à partir de W donnent bien l'équation de Hamilton $\dot{\theta}_i = \omega_i$.

$$\dot{\theta}_i = \frac{d}{dt} \left(\frac{\partial W}{\partial J_i} \right) = \sum_j \frac{\partial}{\partial q_j} \left(\frac{\partial W}{\partial J_i} \right) \dot{q}_j = \sum_j \frac{\partial}{\partial J_i} \left(\frac{\partial W}{\partial q_j} \right) \frac{\partial H}{\partial p_j} = \sum_j \frac{\partial p_j}{\partial J_i} \frac{\partial H}{\partial p_j} = \frac{\partial H}{\partial J_i} = \omega_i \quad (9.48)$$

Ce qui démontre que l'on retrouve l'équation de Hamilton. La fonction génératrice de la transformation est donc la fonction caractéristique de Hamilton, $W = \sum_j W_j$. Sachant que chaque W_j peut être définie à partir des moments on a

$$W(q_1, \dots, q_n, J_1, \dots, J_n) = \sum_i \int p_i(q_i, J_1, \dots, J_n) dq_i. \quad (9.49)$$

Ainsi la fonction caractéristique de Hamilton n'est autre que l'action réduite de Maupertuis.

3. Représentation des coordonnées en fonction des variables angle-action

Nous venons donc de montrer que la fonction caractéristique de Hamilton $W(q_1, \dots, q_n, J_1, \dots, J_n)$ permet de définir la transformation $(q, p) \rightarrow (\theta, J)$ dans le cas des systèmes complètement intégrables. Cette dernière permet d'obtenir les relations

$$\begin{aligned} q_i &= q_i(\theta, J) \\ p_i &= p_i(\theta, J) \end{aligned} \quad (9.50)$$

Pour obtenir ces relations il est nécessaire de connaître la forme précise de la fonction génératrice $W(q_1, \dots, q_n, J_1, \dots, J_n)$ et de résoudre les règles de génération des fonctions génératrices de seconde espèce. Ici nous allons extraire une forme générale d'équation reliant les coordonnées q_i aux variables angle-action. Pour simplifier nous allons considérer un système à un degré de liberté. La variable d'angle est définie par

$$\theta = \frac{\partial W}{\partial J}(q, J). \quad (9.51)$$

En effectuant une petite variation δq (à action constante), on obtient la variation de l'angle

$$\delta\theta = \frac{\partial^2 W}{\partial q \partial J} \delta q. \quad (9.52)$$

On intègre cette variation sur une période complète de q

$$\Delta\theta = \oint \frac{\partial^2 W}{\partial q \partial J} dq = \frac{\partial}{\partial J} \oint \frac{\partial W}{\partial q} dq_j = \frac{\partial}{\partial J} \oint p dq = 2\pi \frac{\partial J}{\partial J} = 2\pi. \quad (9.53)$$

Ainsi lors de la variation de la coordonnée q d'une période complète, l'angle θ varie de 2π . La coordonnée q est donc une fonction périodique de l'angle θ de période 2π . Il est donc possible de faire un développement en série de Fourier des coordonnées

$$q(\theta, J) = \sum_{m=-\infty}^{\infty} a_m(J) e^{im\theta} \quad (9.54)$$

où $a_m(J)$ sont les coefficients du développement. Puisque q est une fonction périodique de l'angle θ de période 2π et puisque θ varie linéairement en fonction du temps $\theta = \omega t + \theta_0$, avec la pulsation $\omega = \frac{\partial H}{\partial J}$, alors la période de la coordonnée q est donnée par $T = \frac{2\pi}{\omega}$. Ceci démontre qu'avec la définition de l'action (Eq. 9.45) la pulsation ω est bien la pulsation du système.

Pour un système à n degrés de liberté on généralise de manière triviale ce résultat et on écrit les coordonnées sous la forme du développement

$$q_i(\theta_1, \dots, \theta_n) = \sum_{m_1=-\infty}^{\infty} \dots \sum_{m_n=-\infty}^{\infty} a_{m_1, \dots, m_n}^{(i)} e^{im_1\theta_1 + \dots + im_n\theta_n} = \sum_{\mathbf{m}} a_{\mathbf{m}}^{(i)} e^{i\mathbf{m} \cdot \boldsymbol{\theta}}, \quad (9.55)$$

où les coefficients $a_{\mathbf{m}}^{(i)} = a_{\mathbf{m}}^{(i)}(J_1, \dots, J_n)$ sont des fonctions des actions. En résumé, les variables angle-action permettent de paramétrer un mouvement multiplement périodique sous la forme d'une représentation de la trajectoire dans l'espace le long d'un tore invariant T^n de dimension n .

C'est cette représentation qui a aboutit au théorème de Kolomogorov, Arnold et Moser (KAM) qui montre que la perturbation d'un système intégrable (qui peut donc être représenté sous la forme d'un tore invariant) n'aboutit pas nécessairement à la destruction complète du tore et que des tores invariants de dimension plus faible peuvent persister et sur lesquelles le mouvement reste multiplement périodique. Ceci remet notamment en cause l'hypothèse admise jusque là d'ergodicité de Boltzmann et mis en lumière par l'expérience numérique de Fermi-Pasta-Ulam en 1953.

VI. Théorie des perturbations adiabatiques

La représentation en variable angle-action est très puissante car elle permet d'identifier les différentes fréquences des systèmes mécaniques. Toutefois, elle ne peut s'appliquer qu'à une certaine classe de système dit intégrables. L'immense majorité des systèmes mécaniques est non-intégrable et il est impossible de définir des variables angle-action. Cependant dans bien

des cas le système non-intégrable est voisin d'un système intégrable, il est alors possible d'utiliser la théorie des perturbations. Un schéma perturbatif repose toujours sur une décomposition de l'Hamiltonien total H selon un Hamiltonien H_0 intégrable et une perturbation V . On suppose qu'il existe donc un jeu de coordonnées angle-action $(\boldsymbol{\theta}, \mathbf{J})$ associé à H_0 où $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ et $\mathbf{J} = (J_1, \dots, J_n)$. L'Hamiltonien total s'écrit donc

$$H(\boldsymbol{\theta}, J) = H_0(J) + V(\boldsymbol{\theta}, J). \quad (9.56)$$

La perturbation V introduit des couplages entre les différents degrés de libertés si bien que les actions ne sont plus des constantes et les angles ne sont plus des fonctions linéaires du temps. On peut séparer l'effet de la perturbation en deux catégories

- Régime résonant : Il existe $\mathbf{m} \in \mathbb{Z}^n$ et $\mathbf{J}_0 \in \mathbb{R}^n$ tel que $\mathbf{m} \cdot \boldsymbol{\omega}(\mathbf{J}_0) = 0$. Dans ce cas, pour un tore \mathbf{J}_0 , certaines fréquences sont commensurables. Cet effet de résonance implique des transferts important d'énergie d'un degré de liberté à un autre.
- Régime chaotique : Si il existe au moins deux résonances, le système peut entrer dans un régime dit chaotique pour lequel son comportement ne permet plus d'être prédit.
- Régime adiabatique : Dans ce régime il n'existe pas de résonance. Le système s'il n'est pas intégrable doit avoir un comportement proche d'un système intégrable si la perturbation V n'est pas trop forte.

C'est ce dernier régime que nous considérons ici dans le but d'introduire la théorie des perturbations adiabatiques à l'ordre 1. Pour construire le schéma perturbatif on introduit un paramètre λ qui caractérise l'intensité du couplage. Ce paramètre permet de suivre l'ordre de la perturbation. En fin de calcul on utilisera simplement $\lambda = 1$. L'Hamiltonien s'écrit donc

$$H(\boldsymbol{\theta}, \mathbf{J}) = H_0(\mathbf{J}) + \lambda V(\boldsymbol{\theta}, \mathbf{J}). \quad (9.57)$$

On recherche une transformation canonique perturbative permettant de trouver les coordonnées angle-action de l'Hamiltonien total $(\boldsymbol{\theta}, \mathbf{J}) \rightarrow (\boldsymbol{\varphi}, \mathbf{I})$. Le nouvel Hamiltonien ne doit pas dépendre des nouveaux angles et ne peut être une fonction que des nouvelles actions $K(\mathbf{I})$. Pour effectuer cette transformation on recherche une fonction génératrice de deuxième espèce $F(\boldsymbol{\theta}, \mathbf{I})$ pour laquelle les règles de transformation sont

$$\varphi_i = \frac{\partial F}{\partial I_i}, \quad J_i = \frac{\partial F}{\partial \theta_i}. \quad (9.58)$$

Si la perturbation est faible la transformation doit être proche de la transformation identité

$$F(\boldsymbol{\theta}, \mathbf{I}) = \boldsymbol{\theta} \cdot \mathbf{I} + \lambda F_1(\boldsymbol{\theta}, \mathbf{I}) + \mathcal{O}(\lambda^2). \quad (9.59)$$

La seconde équation de transformation donne

$$\mathbf{J} = \mathbf{I} + \lambda \frac{\partial F_1}{\partial \boldsymbol{\theta}} + \mathcal{O}(\lambda^2) \quad (9.60)$$

La transformation étant indépendante du temps, l'Hamiltonien K est simplement l'Hamiltonien H exprimé en fonction des nouvelles actions $K(\mathbf{I}) = H(\boldsymbol{\theta}, \mathbf{J})$. On peut donc écrire K sous la forme

$$K(\mathbf{I}) = H_0(\mathbf{I}) + \frac{\partial H_0}{\partial \mathbf{J}}(\mathbf{J} - \mathbf{I}) + \lambda V(\boldsymbol{\theta}, \mathbf{I}) + \mathcal{O}(\lambda^2) = K_0(\mathbf{I}) + \lambda K_1(\mathbf{I}) + \mathcal{O}(\lambda^2) \quad (9.61)$$

où $K_0(\mathbf{I})$ et $K_1(\mathbf{I})$ s'écrivent

$$\begin{aligned} K_0(\mathbf{I}) &= H_0(\mathbf{I}), \\ K_1(\mathbf{I}) &= V(\boldsymbol{\theta}, \mathbf{I}) + \boldsymbol{\omega}(\mathbf{I}) \cdot \frac{\partial F_1}{\partial \boldsymbol{\theta}}(\boldsymbol{\theta}, \mathbf{I}). \end{aligned} \quad (9.62)$$

où $\boldsymbol{\omega}(\mathbf{I}) = \frac{\partial H_0}{\partial \mathbf{J}}$ sont les fréquences du système. D'une manière générale le terme de couplage V doit être une fonction périodique des angles $\boldsymbol{\theta}$. On peut donc effectuer le développement en série de Fourier

$$V(\boldsymbol{\theta}, \mathbf{I}) = \sum_{\mathbf{m}} V_{\mathbf{m}}(\mathbf{I}) e^{i\mathbf{m} \cdot \boldsymbol{\theta}}, \quad (9.63)$$

où $V_{\mathbf{m}}(\mathbf{I})$ est donné par

$$V_{\mathbf{m}}(\mathbf{I}) = \frac{1}{2\pi^n} \int d\theta_1 \dots d\theta_n V(\boldsymbol{\theta}, \mathbf{I}) e^{-i\mathbf{m} \cdot \boldsymbol{\theta}} = \langle V(\boldsymbol{\theta}, \mathbf{I}) e^{-i\mathbf{m} \cdot \boldsymbol{\theta}} \rangle_{\boldsymbol{\theta}} \quad (9.64)$$

La perturbation contient un terme constant $V_0(\mathbf{I}) = \langle V(\boldsymbol{\theta}, \mathbf{I}) \rangle_{\boldsymbol{\theta}}$ et les autres termes sont oscillants

$$V(\boldsymbol{\theta}, \mathbf{I}) = V_0(\mathbf{I}) + \sum_{\mathbf{m} \neq \mathbf{0}} V_{\mathbf{m}}(\mathbf{I}) e^{i\mathbf{m} \cdot \boldsymbol{\theta}} \quad (9.65)$$

On peut effectuer un développement identique pour la fonction génératrice F_1

$$F_1(\boldsymbol{\theta}, \mathbf{I}) = \sum_{\mathbf{m}} F_{\mathbf{m}}(\mathbf{I}) e^{i\mathbf{m} \cdot \boldsymbol{\theta}} \quad (9.66)$$

Au final on obtient l'expression de la correction au premier ordre K_1 sous la forme

$$K_1(\mathbf{I}) = V_0(\mathbf{I}) + \sum_{\mathbf{m} \neq \mathbf{0}} [V_{\mathbf{m}}(\mathbf{I}) + i\mathbf{m} \cdot \boldsymbol{\omega}(\mathbf{I}) F_{\mathbf{m}}(\mathbf{I})] e^{i\mathbf{m} \cdot \boldsymbol{\theta}}. \quad (9.67)$$

Si il n'y a pas de résonances $\mathbf{m} \cdot \boldsymbol{\omega}(\mathbf{I}) \neq 0$ alors on peut toujours choisir $F_{\mathbf{m}}(\mathbf{I})$ tel que le second terme est nul. Au final on obtient le résultat très simple

$$K(\mathbf{I}) = H_0(\mathbf{I}) + \langle V(\boldsymbol{\theta}, \mathbf{I}) \rangle_{\boldsymbol{\theta}}. \quad (9.68)$$

Ainsi au premier ordre de la théorie des perturbations adiabatiques, l'effet d'une perturbation se résume simplement à la moyenne sur les angles de cette perturbation. Ce résultat est identique à la théorie des perturbations non-résonantes en mécanique quantique.