

Mécanique analytique

Version de 2007
révisée en 2024

Jean Hare

Avant-propos

Les formalismes lagrangien et hamiltonien de la mécanique ont été élaborés progressivement durant les années 1750 à 1850, même si certains des résultats figurant à la fin de cet exposé remontent seulement aux premières décennies du 20^e siècle. Ces deux approches de la mécanique ont le même « contenu physique » que les équations de Newton, mais constituent un autre point de vue, conduisant à des formalismes plus généraux et plus puissants, sur lesquels reposent une grande partie de la physique actuelle.

L'intérêt pratique de ces formalismes tient en particulier aux aspects suivants :

- la prise en compte des liaisons sans introduire les forces correspondantes, grâce à la notion de coordonnées généralisées, permettant une obtention plus aisée des équations du mouvement et souvent, une résolution plus rapide ;
- une description unifiée pour des systèmes mécaniques discrets et continus, ainsi que pour des systèmes que les principes de Newton ne décrivent pas (par exemple l'électromagnétisme de Maxwell) : la théorie classique des champs repose sur le même principe que la mécanique ! (cf. cours d'électromagnétisme) ;
- ce sont des outils puissants pour décrire des systèmes complexes (non-linéaires, chaotiques) ;
- ils s'appliquent aussi à d'autres systèmes que la mécanique (électronique par exemple) ;
- ils ont de nombreux prolongements en mécanique quantique, théorie quantique des champs, et en physique statistique.

L'objectif de ce document est de présenter, en près de 50 pages (parties **B** et **C**), l'ensemble des connaissances en mécanique classique correspondant au niveau de la troisième année de licence. En ce sens, il ne fait pas nombre avec les nombreux et excellents manuels disponibles. Ceux-ci restent une référence indispensable pour ceux qui voudraient approfondir, mais (hormis peut-être [1]) sont généralement trop volumineux et trop généraux.

Il ne s'agit pas non plus d'un document d'initiation à la mécanique classique, dans la mesure où il s'appuie souvent sur des connaissances de base de mécanique du point ou du solide qui sont enseignées durant les premières années de licence.

Il comporte aussi, essentiellement à partir de la partie **D**, de nombreux développements, parfois assez qualitatifs, qui dépassent le niveau du public visé, mais sont comme des approfondissements utiles et des exemples d'application, qui motivent en grande partie cet enseignement qui, sans cela, pourrait paraître un peu désuet. De façon plus générale, tous les paragraphes marqués par le symbole ★, ou figurant à l'intérieur d'une telle section, constituent des compléments et non des connaissances exigibles.

Au sujet des notations

Dans ce document, nous utilisons les notations françaises usuelles, où les quantités vectorielles sont surmontées d'une flèche tout en restant en italique mathématique, comme \vec{r} ou \vec{L} , et le produit vectoriel par le symbole \wedge . Comme il est de coutume en physique, la conjugaison complexe est représentée par une $*$ en exposant. En ce qui concerne les opérateurs, classiques comme quantiques, ils sont notés non pas avec des chapeaux comme \hat{x} , mais avec des lettres sans-sérif droites comme x et \vec{p} . La transposition de matrice est notée avec un \top en exposant, comme A^\top . Enfin, nous n'utiliserons pas les conventions d'Einstein que les indices répétés, aussi les signes de sommation seront tous explicites.



Table des matières

Avant-propos	3
Table des matières	5
Table des figures	7
A Introduction	7
1 Notion de principe variationnel	8
2 Exemples	8
2.1 Courbe d'équilibre d'une corde pesante	8
2.2 Forme d'équilibre d'un film de savon	10
B Formulation lagrangienne	13
3 Description du système	13
3.1 Configurations d'un système mécanique	13
3.2 Coordonnées généralisées	13
3.3 Quelques exemples	14
4 Principe de moindre action	15
4.1 Énoncé du principe	15
4.2 Équations d'Euler–Lagrange	15
4.3 Propriétés mathématiques du lagrangien	16
5 Expression du lagrangien	16
5.1 Cas d'une particule libre	17
5.2 Généralisation	18
5.3 Forces généralisées	19
5.4 Forces conservatives	19
5.5 Quelques exemples	21
5.6 Cas des coordonnées dépendant explicitement du temps	25
5.7 Potentiels généralisés	25
5.8 Forme générale du lagrangien	25
6 Symétries et lois de conservation	27
6.1 Notion d'intégrale première	27
6.2 Énergie mécanique	27
6.3 Théorème de Noether	28
6.4 Énergie	29
6.5 Impulsion	29
6.6 Moment cinétique	30

7 Variations sous contraintes	31
7.1 Multiplicateurs de Lagrange pour une fonction	31
7.1.1 Cas avec une contrainte	31
7.1.2 Cas avec plusieurs contraintes	32
7.2 Contraintes intégrales pour les fonctionnelles	33
7.3 Liaisons holonomes	34
7.4 Liaisons non-holonomes	35
8 Applications importantes	36
8.1 Modes normaux	36
8.2 Variables continues ★	40
9 L'action comme fonction des coordonnées	42
C Formulation hamiltonienne	43
10 Équations de Hamilton	44
10.1 Point de vue de Hamilton	44
10.2 Équations de Hamilton	45
10.3 Lien avec le point de vue de Lagrange	45
10.4 Lien avec l'énergie	46
10.5 Particule chargée dans un champ électromagnétique	47
10.6 Propriétés du hamiltonien	47
10.7 Crochets de Poisson	48
10.8 Trajectoires et portraits de phase	50
11 Structure symplectique et transformations canoniques	50
11.1 Définition	50
11.2 Transformations canoniques	52
11.2.1 Motivation et définition	52
11.2.2 Caractérisation	52
11.2.3 Exemples simples	53
11.2.4 Propriétés des transformations symplectiques	53
11.2.5 Transformations canoniques infinitésimales	54
11.3 Fonctions génératrices	55
11.3.1 Présentation générale	55
11.3.2 Un exemple en détail : le référentiel tournant	56
11.3.3 Identité des deux points de vue ★	58
11.4 Transformation canonique engendrée par l'action	59
11.5 Théorèmes de Liouville et de Poincaré	60
12 Systèmes intégrables ★	61
12.1 Notion de système intégrable	61
12.2 Variables d'action	62

D	Hamilton–Jacobi et variables actions-angles ★	63
13	Formalisme de Hamilton–Jacobi	64
13.1	Équation de Hamilton–Jacobi	64
13.2	Notion de solution complète	65
13.3	Exemple de l’oscillateur harmonique	65
13.4	Séparation des variables	66
13.4.1	Équation de H-J stationnaire	66
13.4.2	Cas « générique »	66
13.5	Application au potentiel central	67
14	Variables canoniques « action-angle »	67
14.1	Transformation canonique engendrée par l’action réduite	67
14.2	Interprétation géométrique	68
14.3	Variables d’angle	68
14.4	Évolution des variables d’action	69
15	Liens avec la mécanique quantique	70
15.1	Quantification canonique	71
15.2	Crochets de Poisson	71
15.3	Transformations canoniques	71
15.4	Équation de Hamilton–Jacobi	71
15.5	Variables action-angle	72
	Références	75

Table des figures

1	La chaînette	9
2	Paramétrisation de la bulle de savon	11
3	Résolution graphique pour le film de savon	12
4	Angles d’Euler	14
5	Les quatre coordonnées généralisées proposées pour l’exemple 5	14
6	Illustration des exemples (I/II)	21
7	Illustration des exemples (II/II)	23
8	Exemple de multiplicateurs de Lagrange	33
9	Matrices V et U de la chaîne d’oscillateurs avec CLP	39
10	Portraits de phase (de gauche à droite) de l’oscillateur harmonique, du problème de Kepler (partie radiale), du pendule pesant	51

A – Introduction

1 Notion de principe variationnel

Les principes variationnels jouent un grand rôle dans tous les domaines de la physique. Ils consistent à rechercher des fonctions (trajectoires par exemple, mais aussi surfaces, etc...) qui rendent *extrémale* une « fonctionnelle », c'est à dire une forme linéaire d'un espace de fonctions (généralement C^1) dans \mathbb{R} .

2 Exemples

Voici quelques exemples, plus ou moins bien connus, dont les deux premiers seront traités ci-dessous, les trois suivants seront vus en TD, et le dernier requiert un outillage mathématique plus élaboré, lequel sera présenté au §7.

Ex. 1 Courbe d'équilibre d'une corde pesante

Ex. 2 Courbes d'équilibre en capillarité

Ex. 3 Principe de Fermat en optique

Ex. 4 Dissipation minimale

Ex. 5 Courbes brachistochrones¹ dans un champ uniforme

Ex. 6 Lignes iso-périmétriques.

2.1 Courbe d'équilibre d'une corde pesante

Une corde de masse linéique μ et de longueur L est suspendue entre deux points A et B dont la distance est inférieure à L . La corde est supposée inextensible et sans raideur, et on note $s \in [0, L]$ son abscisse curviligne. Le problème est de déterminer sa forme d'équilibre $\vec{r}(s)$ sous l'effet de son propre poids, représenté par le potentiel gravitationnel $V(\vec{r}) = gz$.

On doit donc rechercher un minimum de son énergie potentielle de pesanteur :

$$U = \int_0^L \mu V(\vec{r}(s)) ds \quad \text{où } s \in [0, L] \text{ est l'abscisse curviligne,}$$

avec bien sûr $\vec{r}(0) = A$ et $\vec{r}(L) = B$.

Nous cherchons un minimum de l'énergie U qui se traduit par un accroissement nul (au premier ordre) si l'on s'écarte légèrement de la forme d'équilibre. Notons $\vec{r}(s)$ cette forme d'équilibre, $\vec{r}'(s)$ une autre forme très proche de celle-ci, et $\delta\vec{r}(s) = \vec{r}'(s) - \vec{r}(s)$ la petite variation correspondante², vérifiant les conditions aux limites $\delta\vec{r}(0) = \delta\vec{r}(L) = \vec{0}$.

On peut alors écrire :

$$dU = \int_0^L \mu (\delta V(\vec{r}(s)) ds + V(\vec{r}(s)) \delta ds) .$$

¹ C'est-à-dire en temps minimum.

² On utilise le symbole $\delta\cdot$ pour désigner les *variations*, afin de les distinguer immédiatement des incréments infinitésimaux $d\cdot$ du calcul différentiel ordinaire.

Le premier terme de l'intégrale s'écrit évidemment : $\delta V = \overrightarrow{\text{grad}} V \cdot \delta \vec{r}$. Mais le second est plus subtil puisqu'il fait apparaître la variation de l'abscisse curviligne due à $\delta \vec{r}$. On peut introduire le vecteur *unitaire* tangent $\vec{t}(s)$, tel que, par définition, $d\vec{r} = ds \vec{t}$. On a alors :

$$ds = \vec{t} \cdot d\vec{r} \quad \text{et donc} \quad \delta(ds) = \delta(\vec{t}) \cdot d\vec{r} + \vec{t} \cdot \delta(d\vec{r}) = \vec{t} \cdot \delta(d\vec{r}).$$

En effet, comme $\|\vec{t}\| = 1$, $\delta \vec{t}$ est nécessairement perpendiculaire à \vec{t} , donc à $d\vec{r}$, ce qui entraîne l'annulation du premier terme³. Il vient donc :

$$dU = \int_0^L \mu \left(\overrightarrow{\text{grad}} V(\vec{r}(s)) \cdot \delta \vec{r} ds + V(\vec{r}(s)) \vec{t} \cdot d(\delta \vec{r}) \right) = \int_0^L \mu \left(\overrightarrow{\text{grad}} V \cdot \delta \vec{r} + V \vec{t} \cdot \frac{d\delta \vec{r}}{ds} \right) ds.$$

Toutefois cette expression ne permet pas encore de conclure, puisque l'on a dans un terme $\delta \vec{r}$, et dans un autre sa dérivée $d(\delta \vec{r})$, mais une intégration par parties donne :

$$\begin{aligned} dU &= \int_0^L \mu \overrightarrow{\text{grad}} V(\vec{r}(s)) \cdot \delta \vec{r} ds + \int_0^L \mu V(\vec{r}(s)) \vec{t} \cdot \frac{d\delta \vec{r}}{ds} ds \\ &= \int_0^L \mu \overrightarrow{\text{grad}} V(\vec{r}(s)) \cdot \delta \vec{r} ds - \int_0^L \mu \frac{d}{ds} \left(V(\vec{r}(s)) \vec{t} \right) \cdot \delta \vec{r} ds \\ &= \int_0^L \mu \left(\overrightarrow{\text{grad}} V(\vec{r}(s)) - \frac{d}{ds} \left(V(\vec{r}(s)) \vec{t} \right) \right) \cdot \delta \vec{r} ds \end{aligned}$$

où la deuxième égalité résulte de l'intégration par partie, la partie toute intégrée étant nulle en raison des conditions aux limites. La variation $\delta \vec{r}$ étant quelconque, on peut en déduire :

$$\overrightarrow{\text{grad}} V(\vec{r}(s)) - \frac{d}{ds} \left(V(\vec{r}(s)) \vec{t}(s) \right) = 0.$$

Pour obtenir la solution du problème initial, on utilise la forme particulière $V = gz$, et on suppose que la corde peut être décrite par $z(x)$. Projetée sur l'axe Ox , l'équation précédente s'écrit simplement :

$$\frac{d}{ds} \left(\frac{z(x)}{\sqrt{1+z'^2}} \right) = 0 \iff \frac{z(x)}{\sqrt{1+z'^2}} = a \quad \text{où } a \text{ est une constante d'intégration.}$$

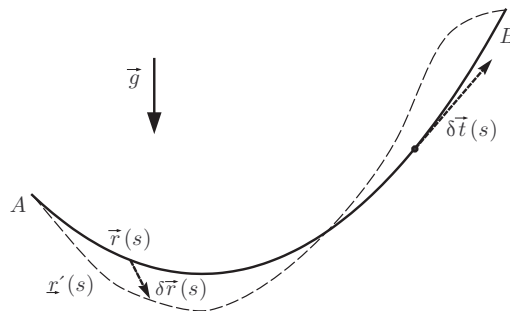


Fig. 1 – La chaînette : en trait plein la courbe d'équilibre $\vec{r}(s)$, et en tirets la courbe variée (fictive) $\vec{r}'(s)$. La variation $\delta \vec{r}$ est la différence entre les deux. Noter que le vecteur tangent $\vec{t}(s)$ et l'abscisse curviligne s sont ceux de $\vec{r}(s)$, mais peuvent différer (au premier ordre) de ceux de $\vec{r}'(s)$.

³ Si l'on est sceptique sur cette démonstration, pourtant exacte, on peut considérer que :

$$\delta(ds) = \|d(\vec{r} + \delta \vec{r})\| - \|d\vec{r}\| \quad \text{où} \quad \|d(\vec{r} + \delta \vec{r})\| = \left(d\vec{r}^2 + 2 d\vec{r} \cdot d\delta \vec{r} + o(d\delta \vec{r}) \right)^{1/2} = \|d\vec{r}\| + \vec{t} \cdot d\delta \vec{r} + o(d\delta \vec{r}),$$

donc $\delta(ds) = \vec{t} \cdot d(\delta \vec{r})$ comme annoncé.

Il vient :

$$z'(x) = \sqrt{(z/a)^2 - 1} \quad \text{et donc} \quad z(x) = a \cosh\left(\frac{x - x_0}{a}\right).$$

Dans cette expression a et x_0 sont deux constantes d'intégration, fixées *in fine* par $z_B - z_A$ et $x_B - x_A$. On retrouve la « chaînette », de façon un peu technique, mais en partant d'un point de vue global au lieu de la méthode locale habituelle.

Si l'on utilise au contraire la projection sur l'axe Oz , on a moins simplement :

$$\sqrt{1 + z'^2} = \frac{d}{dx} \left(\frac{z(x) z'(x)}{\sqrt{1 + z'^2}} \right) \quad \text{d'où} \quad z(x) z''(x) = 1 + z'^2(x)$$

qui peut être résolue (*cf.* exemple suivant § 2.2) et conduit naturellement au même résultat que la projection sur Ox .

2.2 Forme d'équilibre d'un film de savon

Paramétrisation On rappelle qu'un film de savon possède une énergie de tension de surface (ou « capillaire ») qui est proportionnelle à son aire. En conséquence, sa forme d'équilibre est celle qui réalise le minimum de surface pour des conditions aux limites données. Un exemple simple est celui d'un film qui joint deux cerceaux circulaires coaxiaux et de même rayon R . Si on note z la coordonnée de long de leur axe commun, on peut supposer que l'un est contenu dans le plan $z = -L/2$ et l'autre dans le plan $z = L/2$ (*cf.* figure 2).

En vertu de la symétrie du problème (le poids étant ici supposé négligeable) on peut admettre que la solution est une surface de révolution qui – en coordonnées cylindriques d'axe z – est décrite par la courbe $\rho = f(z)$; la fonction f vérifie alors les conditions aux limites $f(-L/2) = f(L/2) = R$. La surface s'écrit, en considérant l'élément de longueur ds le long du méridien qui vérifie $ds^2 = dz^2 + d\rho^2$:

$$S = \int_{-L/2}^{L/2} 2\pi\rho \, ds = 2\pi \int_{-L/2}^{L/2} f(z) \sqrt{1 + f'^2(z)} \, dz$$

Calcul de la variation δS De même qu'une fonction dérivable est stationnaire – c'est-à-dire possède une tangente horizontale – en tout extremum, le caractère minimal de la surface implique que son accroissement soit nul si l'on s'écarte légèrement de la forme d'équilibre. On doit donc évaluer la variation $\delta S = S(f + \delta f) - S(f)$ qui résulte d'une petite variation δf quelconque de la fonction f , et cette variation doit être nulle, au premier ordre, indépendamment du choix de δf ⁴.

La seule contrainte sur δf résulte des conditions aux limites sur f :

$$\delta f(-L/2) = \delta f(L/2) = 0. \quad (1)$$

Il vient alors :

$$\frac{\delta S}{2\pi} = \int_{-L/2}^{L/2} \delta(f(z) \sqrt{1 + f'^2(z)}) \, dz = \int_{-L/2}^{L/2} \left((\delta(f(z))) \sqrt{1 + f'^2(z)} + f(z) \delta(\sqrt{1 + f'^2(z)}) \right) dz,$$

où la variation de la racine carrée se calcule comme une différentielle :

$$\delta \sqrt{1 + f'^2(z)} = \frac{f' \delta f'}{\sqrt{1 + f'^2(z)}}.$$

⁴ On utilise le symbole $\delta \cdot$ pour désigner les *variations*, afin de les distinguer immédiatement des incréments infinitésimaux $d \cdot$ du calcul différentiel ordinaire.

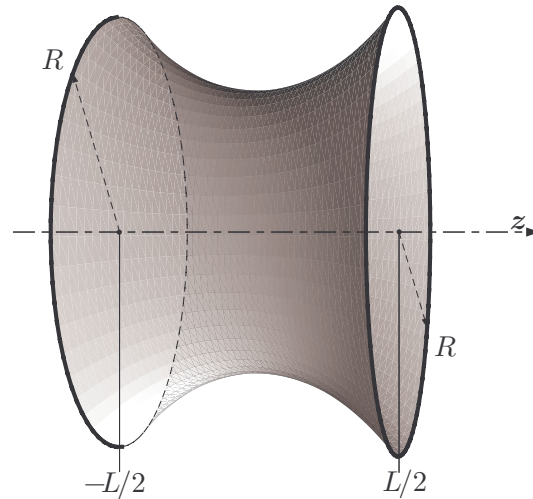


Fig. 2 – Paramétrisation du problème du film de savon

En réunissant les deux termes on obtient :

$$\frac{\delta S}{2\pi} = \int_{-L/2}^{L/2} \delta f \sqrt{1 + f'^2(z)} dz + \delta f' \frac{f f'}{\sqrt{1 + f'^2(z)}} dz .$$

À ce stade il n'est pas encore possible de conclure, puisqu'on a un terme en δf et un autre en $\delta f'$. On peut néanmoins faire une intégration par parties, en utilisant $\delta f' = (\delta f)'$ et en constatant que le terme tout intégré s'annule en raison de (1), la condition de stationnarité :

$$\frac{\delta S}{2\pi} = \int_{-L/2}^{L/2} \left\{ \sqrt{1 + f'^2(z)} - \frac{d}{dz} \left(\frac{f f'}{\sqrt{1 + f'^2(z)}} \right) \right\} \delta f dz = 0 .$$

La fonction δf étant quelconque (hormis la condition (1)), on en déduit que la fonction de z qui apparaît dans les $\{ \}$ doit être identiquement nulle pour assurer l'annulation de δS . En évaluant la dérivée et en simplifiant, on obtient⁵ avec une constante d'intégration A (qui est l'inverse d'une longueur), l'équation différentielle :

$$f'(z) = \sqrt{(f(z)/A)^2 - 1}$$

dont la solution est finalement de la forme :

$$f(z) = A \cosh(z/A).$$

⁵ On écrit successivement :

$$\begin{aligned} \sqrt{1 + f'^2(z)} &= \frac{d}{dz} \left(\frac{f f'}{(1 + f'^2(z))^{1/2}} \right) \equiv \frac{f f'' - f'^2(1 + f'^2)}{(1 + f'^2)^{3/2}} \\ \iff \frac{f(z) f''(z)}{1 + f'^2(z)} &= (f'(z))^2 \iff \frac{f'(z) f''(z)}{1 + f'^2(z)} = \frac{f'(z)}{f(z)} \iff \ln f(z) = \frac{1}{2} \ln(1 + f'(x)^2) + \ln A \end{aligned}$$

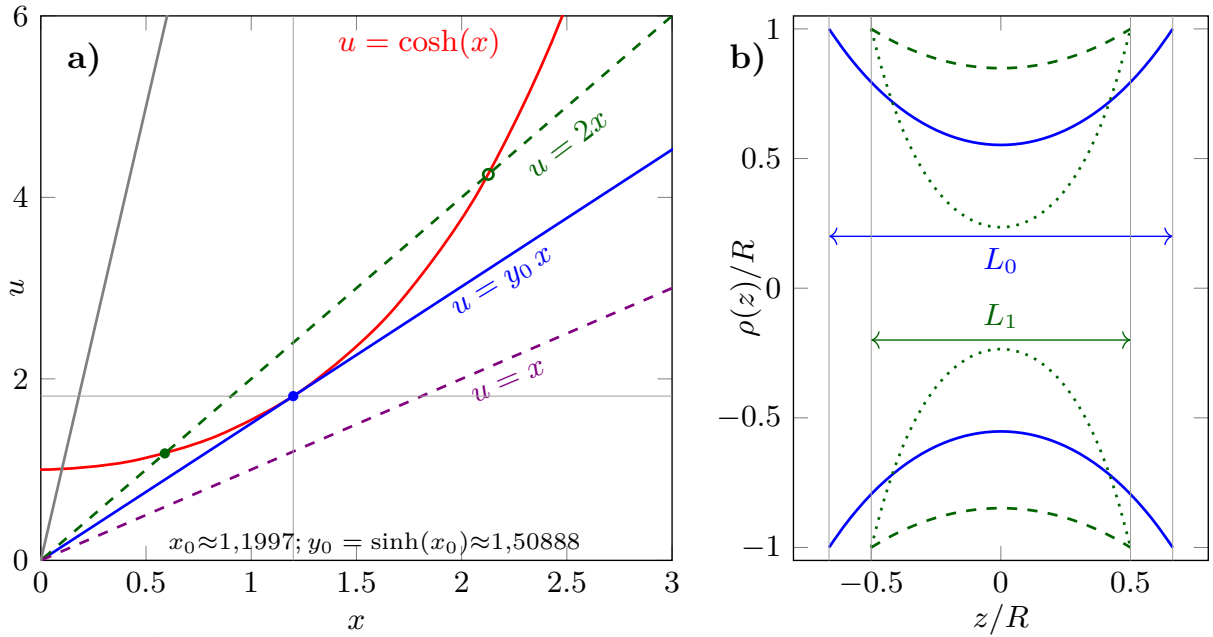


Fig. 3 – **a)** À gauche, résolution graphique de l'équation adimensionnée (2), illustrant le cas $y = 1 < y_0$ sans solution (en tirets violets), le cas limite $y = y_0$ (trait continu bleu) avec une solution (disque bleu), et le cas $y = 2 > y_0$ (en tirets verts) avec deux solutions dont la plus faible est stable (disque vert) et la plus grande instable (cercle vert).

b) À droite, tracé des solutions $\rho(z)$ pour $y = y_0$ (trait continu bleu) correspondant à la distance maximale L_0 , et des deux solutions de (2) pour $y = 2$, à la distance inférieure L_1 , dont la solution physique est en tirets verts, et la solution instable en pointillés verts.

Détermination de A et discussion physique Les conditions aux limites en $z = \pm L/2$ imposent alors la condition $R/A = \cosh(L/2A)$. En posant $x = L/2A$ et $y = 2R/L$, on obtient l'équation adimensionnée :

$$xy = \cosh(x). \quad (2)$$

Cette équation peut être résolue graphiquement en recherchant l'intersection du graphe de $x \mapsto u = \cosh(x)$ avec celui de la fonction linéaire de pente y , $x \mapsto u = yx$ (voir la figure 3-a).

On observe sur cette figure que pour y assez grand c'est-à-dire pour L assez faible, il y a deux intersections. Par exemple, pour $y = 2$, elles sont représentées par les points verts sur la figure 3-a). Elles correspondent à deux solutions de surfaces différentes. On pourrait montrer que la solution stationnaire associée au x le plus grand, soit le A le plus faible, et donc la surface la plus grande, est instable.

Si au contraire on a $y < y_0$, (par exemple $y = 1$) il n'y plus de intersection, ni de solution.

La valeur limite entre ces deux régimes est celui où la droite est tangente à la courbe, ce qui correspond à la valeur ⁶ de $x_0 \approx 1,1997$, soit $u_0 \equiv \cosh(x_0) \approx 1,8102$, $y_0 = \sinh(x_0) \approx 1,5089$, tracée en trait continu rouge sur la figure 3-a). Ainsi, pour $y > y_0$, c'est-à-dire $L/R < 1/2y_0 \approx 0,3314$, on a une solution stable. On peut noter en passant que pour y grand, soit $L/R \ll 1$, la solution stable correspond à un x tendant vers zéro, et alors $xy = R/A$ tend vers 1, ce qui signifie que la surface devient de plus en plus proche d'un cylindre.

La forme de ces solutions est illustrée sur la figure 3-b), pour la valeur limite $y = y_0$ en trait continu bleu.

⁶ C'est la solution de l'équation $x_0 \tanh(x_0) = 1$.

B – Formulation lagrangienne

3 Description du système

3.1 Configurations d'un système mécanique

Considérons pour commencer un système mécanique constitué de N points matériels en interaction mutuelle, et éventuellement soumis à un champ extérieur. On note $\{\vec{r}_\alpha\}$ ($\alpha \in \{1 \dots N\}$) leurs positions. L'état dynamique du système à un instant donné est décrit en toute généralité par ces positions, et les vitesses correspondantes $\{\vec{v}_\alpha\} \equiv \{\dot{\vec{r}}_\alpha\}$, qui moyennant les lois de forces, permettent – en principe – de prévoir son évolution. La « configuration » du système est alors déterminée par les $6N$ composantes des positions et des vitesses.

Toutefois, il arrive fréquemment que ces coordonnées, soient liées entre elles, comme par exemple deux masses reliées par une tige rigide, ce qui impose $(\vec{r}_1 - \vec{r}_2)^2 - a^2 = 0$, ou plus généralement doivent satisfaire des contraintes, comme une condition de roulement dans glissement, *etc.* La résolution des équations de Newton implique alors, en général, d'introduire les forces de liaisons, qui sont autant de nouvelles inconnues, et que l'on doit éliminer entre les diverses équations différentielles satisfaites par les $3N$ positions, lesquelles ne sont pas toutes indépendantes.

Une telle procédure n'est guère efficace, et il serait préférable de disposer d'un nombre réduit de variables dynamiques, dont on rechercherait l'évolution sans avoir à introduire les forces de liaison. En outre, il est fréquent qu'un choix particulier de coordonnées, comme des coordonnées polaires par exemple, simplifie grandement les équations de la dynamique, voire permette d'éliminer directement certaines forces inconnues. Il est donc intéressant de pouvoir les traiter avec un formalisme général, qui ne dépende pas du choix particulier que l'on a fait. C'est précisément l'objet des coordonnées généralisées que nous introduisons ci-dessous.

3.2 Coordonnées généralisées

Un tel programme est en partie réalisé en introduisant les « coordonnées généralisées », que nous allons définir. On suppose qu'il existe n relations indépendantes du type :

$$f_k(\{\vec{r}_\alpha\}, t) = 0 \quad k \in \{1, \dots, n\}, \quad (3)$$

qui définissent ce que l'on appelle des **liaisons holonomes** (les autres types de liaisons, par exemple une inégalité, seront dites « non-holonomes »). On dit alors que le système possède $D = 3N - n$ degrés de liberté, qui est le nombre de variables indépendantes parmi les $3N$ positions. Il suffit donc de $3N - n$ variables, et de leurs dérivées, pour caractériser la configuration du système.

On appellera alors *coordonnées généralisées* tout jeu de D variables indépendantes, notées $q_1 \dots q_D$, ou plus simplement $\{q_i\}$, dont la nature et la dimension est indifférente, mais telles que les N positions \vec{r}_α s'expriment comme des fonctions de celles-ci. On pourra alors écrire, à l'aide de fonctions *supposées connues* $\vec{R}_\alpha(\{q_i(t)\}, t)$:

$$\vec{r}_\alpha(t) = \vec{R}_\alpha(\{q_i(t)\}, t), \quad \text{et} \quad \vec{v}_\alpha(t) = \frac{d}{dt} \vec{R}_\alpha(\{q_i(t)\}, t) = \sum_i \frac{\partial \vec{R}_\alpha}{\partial q_i} \dot{q}_i + \frac{\partial \vec{R}_\alpha}{\partial t}. \quad (4)$$

3.3 Quelques exemples

1. En l'absence de liaisons, les coordonnées cartésiennes des positions des particules constituent un exemple trivial de coordonnées généralisées.
2. Pour un pendule pesant, on pourra choisir les angles sphériques θ et φ comme coordonnées généralisées, et si l'on suppose en outre que le mouvement est nécessairement plan (liaison holonome $y = 0$), il suffira de l'angle θ comme unique coordonnée généralisée.
3. Pour une perle mobile sur un cerceau tournant à la pulsation ω autour de son diamètre vertical, il n'y a, à nouveau, qu'un seul degré de liberté, et on pourra prendre comme coordonnée l'angle θ , de même que pour le pendule pesant. On écrira donc :

$$\vec{r}(\theta, t) = R \begin{pmatrix} \sin \theta \cos \omega t \\ \sin \theta \sin \omega t \\ \cos \theta \end{pmatrix} \text{ et } \vec{v}(\theta, \dot{\theta}, t) = R \dot{\theta} \begin{pmatrix} \cos \theta \cos \omega t \\ \cos \theta \sin \omega t \\ -\sin \theta \end{pmatrix} + R \omega \begin{pmatrix} -\sin \theta \sin \omega t \\ \sin \theta \cos \omega t \\ 0 \end{pmatrix} .$$

4. Pour un corps solide, on a un très grand nombre de points matériels reliés par presque autant de contraintes. Il suffit de préciser la position de trois de ses points pour caractériser le solide, mais ces positions sont contraintes par les trois distances relatives : il reste donc $D = 3 \times 3 - 3 = 6$ degrés de liberté. Usuellement, on utilise la position d'un point, et l'orientation d'un trièdre XYZ lié au solide par rapport au trièdre xyz de référence, repérée par les trois angles d'Euler (α, β, γ) .

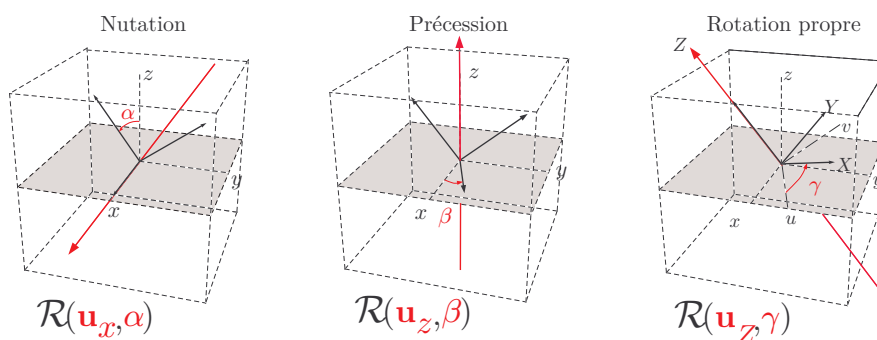


Fig. 4 – Les angles d'Euler : nutation α , précession β , rotation propre γ , dans la définition de Cardan.

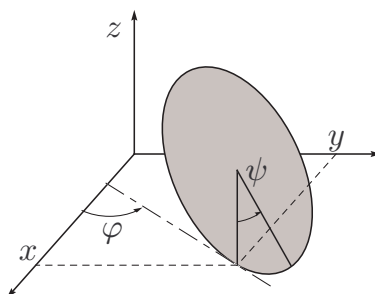


Fig. 5 – Les quatre coordonnées généralisées proposées pour l'exemple 5

5. Pour un disque vertical qui roule sans glisser sur un plan horizontal, on a *a priori* 3 positions et 3 angles pour repérer le disque, mais une liaison holonome sur la position, et une sur l'angle : il reste donc 4 degrés de liberté (la condition de roulement sans glissement, qui

relie les *vitesses*, est non-holonyme). On prendra comme coordonnées : les coordonnées x et y du centre du disque, l'angle φ repérant le plan du disque, et l'angle ψ mesurant sa rotation. Le roulement sans glissement s'écrira alors :

$$R \dot{\psi}(\cos \varphi, \sin \varphi) + (\dot{x}, \dot{y}) = 0 .$$

4 Principe de moindre action

4.1 Énoncé du principe

Le principe « de moindre action » ou « principe de Hamilton », est le fondement de tout ce qui va suivre. Il repose sur plusieurs définitions et hypothèses :

- On se donne une fonction, dite fonction de Lagrange, ou *lagrangien*, $\mathcal{L}(\{q_i\}, \{\dot{q}_i\}, t)$, qui est une fonction de la configuration et éventuellement du temps ;
- On définit l'*action* \mathcal{S} entre les deux positions $\{q_i\}(t_1)$ et $\{q_i\}(t_2)$ aux instants initial t_1 et t_2 , comme une fonction du chemin $\{q_i\}(t) \equiv \gamma(t)$ dans l'espace des configurations par ⁷ :

$$\mathcal{S}(\gamma) = \int_{t_1}^{t_2} \mathcal{L}(\{q_i\}(t), \{\dot{q}_i\}(t), t) dt ; \quad (5)$$

- On postule que que la trajectoire effectivement suivie par le système entre deux points correspond à un chemin pour lequel l'action est *minimale* (ou plus généralement, stationnaire).

4.2 Équations d'Euler–Lagrange

Le principe variationnel énoncé ci-dessus est une propriété *globale* de la trajectoire, mais ne pourra être satisfait que si *localement*, celle-ci se comporte d'une certaine façon, que nous allons maintenant déterminer.

Pour ce faire nous allons évaluer la variation $\delta\mathcal{S}$ de l'action pour une variation infinitésimale $\{\delta q_i\}(t)$ du chemin, autour de la trajectoire, puis imposer sa nullité pour toute variation $\{\delta q_i\}(t)$ avec $t \in [t_1, t_2]$. Nous supposerons toutefois que $\delta q_i(t_1) = \delta q_i(t_2) = 0$, car on cherche la trajectoire menant d'un point à un autre en un temps donné, et on doit donc varier le chemin à extrémités fixées.

Les coordonnées généralisés et les vitesses correspondantes étant – pour le lagrangien – des variables indépendantes, nous pouvons écrire, en développant au premier ordre :

$$\delta\mathcal{S} = \int_{t_1}^{t_2} \sum_i \frac{\partial \mathcal{L}}{\partial q_i} \delta q_i dt + \int_{t_1}^{t_2} \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta \dot{q}_i dt = 0 .$$

Néanmoins les δq_i et les $\delta \dot{q}_i \equiv d(\delta q_i)/dt$ ne sont pas indépendantes, puisque les secondes sont les dérivées des premières. On peut alors effectuer sur le second terme une intégration par parties, ce qui donne :

$$\delta\mathcal{S} = \int_{t_1}^{t_2} \sum_i \left[\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \right] \delta q_i dt + \sum_i \left[\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta q_i \right]_{t_1}^{t_2} . \quad (6)$$

Comme les variations δq_i s'annulent aux extrémités, le second terme est identiquement nul. L'intégrale devant s'annuler pour des $\{\delta q_i(t)\}$ arbitraires et les D coordonnées étant supposées

⁷ Il est d'usage de désigner comme « fonctionnelle » cette fonction dont la variable est elle-même une fonction.

indépendantes, chaque terme de la somme figurant sous l'intégrale doit s'annuler, et il en résulte les D équations différentielles :

$$\boxed{\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) = 0 \quad \text{pour } i \in \{1, \dots, D\}. \quad (7)}$$

Ces équations, dites équations d'Euler-Lagrange, ou de Lagrange pour simplifier, sont donc les équations du mouvement, qui seront généralement des équations différentielles couplées du second ordre pour les $\{q_i\}$.

4.3 Propriétés mathématiques du lagrangien

Avant de nous intéresser aux choix physiquement pertinents pour le lagrangien, nous pouvons relever un certain nombre de propriétés mathématiques :

1. Les équations du mouvement ne sont pas affectées par un changement d'échelle sur le lagrangien ; il est usuel, pour des raisons qui apparaîtront bientôt (*cf.* § 6.2), que celui-ci ait la dimension d'une énergie, cette indétermination correspondant alors au choix de l'unité.
2. Les équations du mouvement ne sont pas affectées par l'addition au lagrangien d'une constante (origine des énergies) ou plus généralement par l'ajout d'une fonction des $\{q_i\}$, des $\{\dot{q}_i\}$ et du temps qui soit la *dérivée totale par rapport au temps* d'une fonction des $F(\{q_i\}, t)$ indépendante des vitesses. En effet, l'action \mathcal{S} associée aux deux lagrangien \mathcal{L} et $\mathcal{L}' = \mathcal{L}$ sera alors modifiée par l'addition de la variation de F entre les deux extrémités du chemin, sans affecter son caractère stationnaire. Dans les équations de Lagrange, en utilisant :

$$f(\{q_i\}, \{\dot{q}_i\}, t) = \sum_j \frac{\partial F}{\partial q_j} \dot{q}_j + \frac{\partial F}{\partial t} ,$$

on constate que les deux termes sont modifiés ainsi :

$$\frac{\partial \mathcal{L}'}{\partial \dot{q}_i} = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} + \frac{\partial F}{\partial \dot{q}_i} \quad \text{et} \quad \frac{\partial \mathcal{L}'}{\partial q_i} = \frac{\partial \mathcal{L}}{\partial q_i} + \sum_j \frac{\partial^2 F}{\partial q_i \partial q_j^2} \dot{q}_j + \frac{\partial^2 F}{\partial t \partial q_i^2} ,$$

et que les termes additionnels vont se compenser pour redonner les mêmes équations d'Euler-Lagrange.

3. Si le système physique possède certaines symétries qui affectent les équations du mouvement, on doit les retrouver dans le lagrangien (nous reviendrons sur ce point, très important physiquement, un peu plus tard).
4. la fonction de Lagrange est additive : si deux systèmes sont correctement décrits, l'un par le lagrangien \mathcal{L}_1 impliquant ses d_1 coordonnées indépendantes $\{q_i^{(1)}\}$, et de même pour \mathcal{L}_2 et $\{q_i^{(2)}\}$, alors la fonction $\mathcal{L} = \mathcal{L}_1(\{q_i^{(1)}\}, \{\dot{q}_i^{(1)}\}, t) + \mathcal{L}_2(\{q_i^{(2)}\}, \{\dot{q}_i^{(2)}\}, t)$ décrira bien l'évolution des $d_1 + d_2$ degrés de liberté du système complet (si les deux systèmes n'interagissent pas).

5 Expression du lagrangien

Nous allons maintenant chercher l'expression appropriée du lagrangien, en commençant par les cas les plus simples, pour en inférer progressivement la structure générale.

5.1 Cas d'une particule libre

On se limite pour commencer à une unique particule (point matériel), libre de se mouvoir dans l'espace, et on se place dans un référentiel galiléen \mathcal{R} . On utilise en outre les *coordonnées cartésiennes* comme coordonnées généralisées. Nous allons montrer que le choix naturel est de prendre un lagrangien égal à *l'énergie cinétique* de la particule.

L'invariance par translation nous permet d'affirmer que le lagrangien ne dépend pas des q_i , et l'isotropie de l'espace assure qu'il ne peut dépendre que du *module* de la vitesse, et pas de sa direction ; de la même façon, l'uniformité du temps permet de supposer que \mathcal{L} ne dépend pas du temps. On peut en déduire que le lagrangien ne dépend que de la norme de la vitesse, et peut donc s'écrire $\mathcal{L}(\vec{r}, \vec{v}, t) = L(v^2)$, où l'on note pour simplifier $v \equiv \|\vec{v}\|$ cette norme.

Les équations de Lagrange peuvent alors s'écrire⁸ :

$$\frac{\partial L}{\partial \vec{r}} = \vec{0} = \frac{d}{dt} \left(\frac{\partial L}{\partial \vec{v}} \right) = \frac{d}{dt} \left(\frac{dL}{dv^2} \frac{\partial v^2}{\partial \vec{v}} \right) = 2 \frac{d}{dt} \left(\frac{dL}{dv^2} \vec{v} \right).$$

Cette égalité implique (tant que L n'est pas réduit à une constante) que $d\vec{v}/dt = \vec{0}$, et on retrouve donc le mouvement uniforme $\vec{v} = \overrightarrow{cste}$ attendu (principe d'inertie). Pour achever de déterminer la forme de L et donc de \mathcal{L} , il faut encore introduire le principe de relativité de Galilée, qui postule que les équations du mouvement sont les mêmes dans tout référentiel galiléen.

Nous considérons donc le mouvement dans deux référentiels, \mathcal{R} et \mathcal{R}' , où \mathcal{R}' est en translation uniforme par rapport à \mathcal{R} à la vitesse $-\vec{V}$. Le principe de relativité nous conduit à rechercher le lagrangien dans \mathcal{R}' sous la même forme que dans \mathcal{R} , soit $\mathcal{L}'(\vec{v}') = L(v'^2)$, avec la même fonction L que dans \mathcal{R} . Les équations de Lagrange conduiront alors nécessairement au résultat attendu $\vec{v}' = \overrightarrow{cste}'$. Observons alors que, en utilisant la relation « cinématique » $\vec{v}' = \vec{v} + \vec{V}$, cela nous permet de définir dans \mathcal{R} un nouveau Lagrangien $\mathcal{L}''(\vec{v}) = L((\vec{v} + \vec{V})^2)$. Il est alors naturel de chercher à relier ce nouveau lagrangien au précédent en utilisant l'invariance des équations de Lagrange lorsqu'on ajoute la dérivée temporelle totale d'une fonction de la position (cf. § 2 au paragraphe précédent). En supposant en outre – pour l'instant – que \vec{V} est infinitésimal, on peut donc écrire :

$$\mathcal{L}'' - \mathcal{L} = L((\vec{v} + \vec{V})^2) - L(\vec{v}^2) = 2 \frac{dL}{dv^2} \vec{v} \cdot \vec{V} + \mathcal{O}(\vec{V}^2) \simeq 2 \left(\frac{dL}{dv^2} \vec{V} \right) \cdot \vec{v}.$$

Or la dérivée totale par rapport au temps d'une fonction de la position est *nécessairement linéaire* en \vec{v} : il en résulte que dL/dv^2 doit être indépendant de v . On en déduit que $L(v^2)$ est une fonction linéaire de v^2 , et on est finalement conduit à écrire :

$$\mathcal{L}(\vec{r}, \vec{v}, t) = A v^2 = \frac{m}{2} \vec{v}^2,$$

où le choix de prendre la constante $A = m/2$ est arbitraire mais d'une part impose au lagrangien d'avoir la dimension d'une énergie ; il est en outre cohérent avec les équations usuelles, et notamment avec la définition de l'énergie mécanique totale qui apparaîtra plus tard (cf. § 10). On a donc établi que, pour une particule libre, le lagrangien est simplement son énergie cinétique et naturellement indépendante de \vec{r} et de t dans ce contexte.

⁸ Nous utiliserons ici, et occasionnellement par la suite, la notation $\partial f / \partial \vec{x}$, où f est une quantité scalaire et \vec{x} un vecteur, qui désigne simplement le vecteur gradient de f par rapport à \vec{x} . Avec ces notations, on a

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial \vec{x}} \cdot d\vec{x}.$$

Coordonnées polaires Le raisonnement précédent utilise les coordonnées cartésiennes, mais conduit à un résultat général indépendant de ces coordonnées. Il est instructif à cet égard de considérer la même situation physique (en se réduisant à un mouvement plan avec $D = 2$), et en prenant pour coordonnées généralisées des *coordonnées polaires* (r, θ) d'origine arbitraire. Ainsi :

$$\mathcal{L} = \frac{m}{2} (\dot{r}^2 + (r\dot{\theta})^2)$$

d'où les équations de Lagrange :

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right) \equiv \frac{d}{dt} (mr^2\dot{\theta}) = \frac{\partial \mathcal{L}}{\partial \theta} = 0 \quad \text{et} \quad \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{r}} \right) \equiv m\ddot{r} = \frac{\partial \mathcal{L}}{\partial r} = mr\dot{\theta}^2,$$

dont la première exprime la conservation du moment cinétique $\sigma = \partial \mathcal{L} / \partial \dot{\theta}$, et la seconde indique que l'accélération radiale se réduit à la « pseudo-force » centrifuge. Il est à noter que cette dernière s'introduit très naturellement par le terme en $\partial \mathcal{L} / \partial r$, alors que \mathcal{L} ne contient encore que l'énergie cinétique.

Résolution des équations du mouvement en (r, θ) L'indépendance de \mathcal{L} par rapport à la variable θ , qui résulte bien sûr de l'invariance par rotation, permet de mettre en évidence la « constante des aires » $C = r^2\dot{\theta} = \sigma/m$. Celle-ci permet l'élimination de $\dot{\theta}$ dans l'équation radiale, qui se met sous la forme $\ddot{r} = C^2/r^3$. En multipliant les deux membres de cette équation par \dot{r} , on obtient une intégrale première (constante du mouvement) $E = \dot{r}^2 + C^2/r^2$, qui n'est autre que l'énergie cinétique. On introduit alors la distance minimale d'approche r_0 , caractérisée par $\dot{r} = 0$ d'où $2E/m = C^2/r_0^2$. On obtient alors :

$$\frac{dr}{dt} = C \sqrt{\frac{1}{r_0^2} - \frac{1}{r^2}} \quad \text{et} \quad \frac{d\theta}{dt} = \frac{C}{r^2}.$$

En éliminant le temps entre ces deux équations et en prenant $\theta = \theta_0$ au point d'approche minimale, il vient :

$$\theta = \theta_0 + \int_{r_0}^r r' \sqrt{\frac{r'^2}{r_0^2} - 1} dr'.$$

Cette intégrale s'obtient aisément en faisant le changement de variable $r \rightarrow u = 1/r$, et conduit à $\theta = \arccos(r_0/r)$, soit $r(\theta) = r_0 \cos(\theta - \theta_0)$, qui est bien l'équation polaire d'une droite. L'énergie cinétique étant constante, cette trajectoire est parcourue à vitesse (linéaire) constante, et on retrouve le mouvement rectiligne uniforme attendu.

5.2 Généralisation

Il est alors assez logique de supposer que, dans le cas de plus d'une particule, le lagrangien fera intervenir l'énergie cinétique totale, traditionnellement notée $K \equiv E_{\text{cin}}$. Cela sera aussi vrai dans le cas où les fonctions $\vec{R}_\alpha(\{q_i\}, t)$ dépendent explicitement du temps. On pourra alors écrire, en dérivant l'équation (4) :

$$K = \sum_{\alpha} \frac{m_{\alpha}}{2} \vec{v}_{\alpha}^2 = \sum_{\alpha, i, j} \frac{m_{\alpha}}{2} \frac{\partial \vec{R}_{\alpha}}{\partial q_i} \cdot \frac{\partial \vec{R}_{\alpha}}{\partial q_j} \dot{q}_i \dot{q}_j + \sum_{\alpha, i} m_{\alpha} \frac{\partial \vec{R}_{\alpha}}{\partial q_i} \cdot \frac{\partial \vec{R}_{\alpha}}{\partial t} \dot{q}_i + \sum_{\alpha} \frac{m_{\alpha}}{2} \left(\frac{\partial \vec{R}_{\alpha}}{\partial t} \right)^2. \quad (8)$$

Un exemple simple mais non trivial est celui de la perle sur un cerceau est celui de l'exemple 3 du paragraphe § 3.3, où les liaisons permettent de se ramener à une seule coordonnée généralisée θ . Son énergie cinétique s'écrit $K = m/2 R^2(\dot{\theta}^2 + \omega^2 \sin^2 \theta)$, dont le deuxième terme, bien que

lui-même indépendant du temps, vient de la dépendance explicite en temps de la position (ici, le dernier terme de (8)).

Cependant, pour limiter la complexité des équations, nous supposons dans les paragraphes qui suivent que les positions \vec{R}_α ne dépendent pas explicitement du temps.

5.3 Forces généralisées

Il nous reste à étudier comment les forces peuvent être prises en compte dans le nouveau formalisme, de façon à ce qu'il soit physiquement équivalent à celui de Newton.

Écrivons donc, pour chaque particule, le principe fondamental de la dynamique à partir de l'équation (4) :

$$\vec{a}_\alpha = \dot{\vec{v}}_\alpha = \sum_i \frac{\partial \vec{R}_\alpha}{\partial q_i} \ddot{q}_i + \sum_{i,j} \frac{\partial^2 \vec{R}_\alpha}{\partial q_i \partial q_j^2} \dot{q}_i \dot{q}_j = \frac{\vec{F}_\alpha}{m_\alpha} \quad (9)$$

et formons, pour l'énergie cinétique K donné par (8) ci-dessus, les équations de la forme de celles d'Euler-Lagrange :

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial K}{\partial \dot{q}_k} \right) - \frac{\partial K}{\partial q_k} &= \frac{d}{dt} \left(\sum_{\alpha,i} m_\alpha \frac{\partial \vec{R}_\alpha}{\partial q_k} \cdot \frac{\partial \vec{R}_\alpha}{\partial q_i} \dot{q}_i \right) - \sum_{\alpha,i,j} m_\alpha \frac{\partial \vec{R}_\alpha}{\partial q_i} \cdot \frac{\partial^2 \vec{R}_\alpha}{\partial q_k \partial q_j^2} \dot{q}_i \dot{q}_j \\ &= \sum_{\alpha,i} m_\alpha \frac{\partial \vec{R}_\alpha}{\partial q_k} \cdot \frac{\partial \vec{R}_\alpha}{\partial q_i} \ddot{q}_i + \sum_{\alpha,i,j} m_\alpha \frac{\partial \vec{R}_\alpha}{\partial q_k} \cdot \frac{\partial^2 \vec{R}_\alpha}{\partial q_i \partial q_j^2} \dot{q}_i \dot{q}_j \end{aligned} \quad (10)$$

car la dérivation temporelle du terme en $\partial_k \vec{R}_\alpha \cdot \partial_i \vec{R}_\alpha$ fait apparaître deux termes, dont le premier se simplifie avec celui qui était affecté d'un signe $-$ sur la première ligne, et le second est celui qui reste sur la deuxième ligne⁹.

En factorisant, dans (10) pour chaque α , le terme en $m_\alpha \partial \vec{R}_\alpha / \partial q_k$, on fait apparaître l'expression de l'accélération \vec{a}_α écrite dans (9) :

$$\frac{d}{dt} \left(\frac{\partial K}{\partial \dot{q}_k} \right) - \frac{\partial K}{\partial q_k} = \sum_\alpha m_\alpha \frac{\partial \vec{R}_\alpha}{\partial q_k} \cdot \vec{a}_\alpha = \mathfrak{F}_k \quad \text{avec} \quad \mathfrak{F}_k = \sum_\alpha \frac{\partial \vec{R}_\alpha}{\partial q_k} \cdot \vec{F}_\alpha, \quad (11)$$

où les \mathfrak{F}_k sont appelées *forces généralisées* associées aux variables q_k (notons que la dimension de q_k étant arbitraire, celle de \mathfrak{F}_k n'est pas celle d'une force, mais telle que, pour chaque degré de liberté, $[\mathfrak{F}_k] \times [q_k]$ soit une énergie).

Évidemment, les forces \vec{F}_α dont nous sommes partis sont les forces totales subies par les particules et elles contiennent fatalement les forces associées aux liaisons holonomes, qui sont inconnues. C'est là que le formalisme des coordonnées généralisées démontre toute son efficacité : ces coordonnées décrivant exclusivement les mouvements compatibles avec les liaisons, les termes $\partial \vec{R}_\alpha / \partial q_k$ sont *tangents* aux « surfaces » décrites par les \vec{R}_α lorsque on fait varier les $\{q_i\}$, tandis que les forces de liaison y sont *normales*. Il en résulte que les forces généralisées \mathfrak{F}_k ne contiennent plus les forces de liaison !

5.4 Forces conservatives

Une classe très importante de problèmes est celle où les forces mises en jeu conservent l'énergie mécanique, et dérivent d'une énergie potentielle $U(\{\vec{R}_\alpha\}) \equiv V(\{q_i\})$. Dans cette écriture, nous

⁹ On remarque bien sûr que le terme en $\dot{q}_i \dot{q}_j$ n'existe que si $\partial_{ij}^2 \vec{R}_\alpha$ est non-nul, comme c'est le cas en général pour des coordonnées curvilignes, contrairement aux coordonnées cartésiennes.

incluons à la fois le cas d'un potentiel extérieur, comme celui du champ gravitationnel, ou électrostatique, et celui d'un potentiel décrivant l'interaction mutuelle des particules. On écrira alors, pour les forces totales « appliquées » :

$$\vec{F}_\alpha = -\frac{\partial U}{\partial \vec{R}_\alpha} \quad \text{et donc} \quad \mathcal{F}_k = -\sum_\alpha \frac{\partial \vec{R}_\alpha}{\partial q_k} \cdot \frac{\partial U}{\partial \vec{R}_\alpha} = -\frac{\partial V}{\partial q_k} . \quad (12)$$

En définitive, nous avons établi que, dans le cas étudié ici, les équations de Newton sont équivalentes à des équations de la forme :

$$\frac{d}{dt} \left(\frac{\partial K}{\partial \dot{q}_k} \right) - \frac{\partial K}{\partial q_k} = \mathcal{F}_k = -\frac{\partial V}{\partial q_k} , \quad (13)$$

et comme $\partial V / \partial \dot{q}_k = 0$, on retrouvera les équations de Lagrange en posant simplement :

$$\mathcal{L}(\{q_i\}, \{\dot{q}_i\}) = K(\{q_i\}, \{\dot{q}_i\}) - V(\{q_i\}) .$$

Inversement, on appellera « **lagrangien standard** » un lagrangien qui peut se mettre sous cette forme, et où l'énergie cinétique K est une **forme quadratique définie positive** des \dot{q}_i , dont les coefficients sont des fonctions des seuls q_j (et non des \dot{q}_i), et l'énergie potentielle V est elle-même indépendante des \dot{q}_i .

Vocabulaire On appelle **moment conjugué** de la variable généralisée q_i la quantité :

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} .$$

Pour les coordonnées cartésiennes d'une particule, et dans le cas d'un lagrangien standard, ces 3 moments conjugués sont les composantes de la quantité de mouvement $\vec{p} = \partial K / \partial \vec{r} = m \vec{v}$, aussi appelée « impulsion »¹⁰. En termes des moments, les équations d'Euler–Lagrange sont : $dp_i/dt = \partial \mathcal{L} / \partial q_i$.

Ces moments jouent un rôle clé dans la construction de la théorie hamiltonienne.

Remarque 1 Lorsque seulement une partie des forces généralisées dérive d'un potentiel, on peut utiliser la même procédure, tout en conservant dans le membre de droite de (11) la partie \mathcal{F}'_k des forces qui est non-conservative. Cela fournit un moyen de prendre en compte la dissipation (frottements, etc...) et nous servira aussi à prendre en compte des liaisons non-holonomes.

Remarque 2 Notons bien que le potentiel utilisé ici ne contient pas les interactions de liaison, ni les contributions que les liaisons rendent superflues. Par exemple, si deux particules sont astreintes à garder une distance constante, et qu'elles interagissent *en outre* par un potentiel $U(\|\vec{R}_1 - \vec{R}_2\|)$, le potentiel correspondant $V(\{q_i\})$ sera constant, et ne contribuera donc pas aux forces généralisées. Mathématiquement, la force généralisée comportera en effet un terme :

$$\frac{\partial \vec{R}_1}{\partial q_k} \cdot \frac{\partial U}{\partial \vec{R}_1} + \frac{\partial \vec{R}_2}{\partial q_k} \cdot \frac{\partial U}{\partial \vec{R}_2} = \frac{\partial U}{\partial \vec{R}_1} \cdot \left(\frac{\partial \vec{R}_1}{\partial q_k} - \frac{\partial \vec{R}_2}{\partial q_k} \right) = \frac{\partial U}{\partial \vec{R}_1} \cdot \frac{\partial \vec{R}_1 - \vec{R}_2}{\partial q_k} = 0 ,$$

qui est bien nul car le premier terme est radial, et le second orthoradial à cause de la liaison.

¹⁰ L'impulsion est une notion plus générale, qui coïncide ici avec la quantité de mouvement. Elle joue un rôle particulier la théorie électrodynamique (classique), éventuellement relativiste où ne n'est plus forcément le cas.

5.5 Quelques exemples

a) La machine d'Atwood

Elle est constituée d'une poulie sur laquelle s'appuie une corde (inextensible, et dont on néglige la masse) lestée à ses deux extrémités et reposant sur une poulie parfaite (cf. figure 6a ci-dessous). On note x la position verticale de la masse m_1 (orientée vers le haut). En utilisant les liaisons (dont notamment $\theta = (x - x_0)/R$), on peut alors écrire :

$$\mathcal{L} = \frac{m_1 \dot{x}^2}{2} + \frac{m_2 \dot{x}^2}{2} + \frac{J \dot{x}^2}{2R^2} - (m_1 - m_2)gx$$

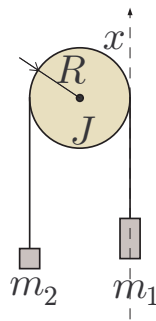
d'où :

$$p_x \equiv \frac{\partial \mathcal{L}}{\partial \dot{x}} = (m_1 + m_2 + J/R^2)\dot{x} \quad \text{et} \quad \mathcal{F}_x \equiv \frac{\partial \mathcal{L}}{\partial x} = -(m_1 - m_2)g$$

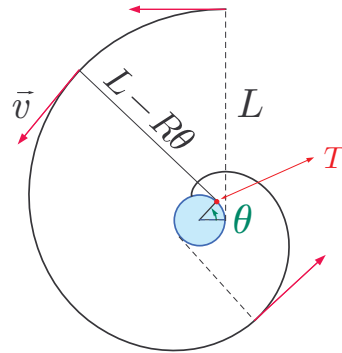
d'où l'équation du mouvement :

$$(m_1 + m_2 + J/R^2)\ddot{x} = -(m_1 - m_2)g ,$$

qui (pour $m_1 > m_2$) décrit une *chute libre ralentie* avec une accélération $\ddot{x} = -g'$ avec $g' = (m_1 - m_2)/(m_1 + m_2 + J/R^2) g$, qui est toujours inférieur à g .



(a) Machine d'Atwood



(b) Speedball

Fig. 6 – Illustration des exemples a) et b) dans cette section 5.5.

b) Corde qui s'enroule autour d'un mât

Ce problème physique n'est pas sans rapport avec un jeu traditionnel redécouvert il y a quelques années sous le nom de « speedball ». Nous nous limitons au problème à deux dimensions : une balle est fixée au bout d'une corde (inextensible, de raideur et de masse négligeables), laquelle s'enroule autour d'un mât (cf. figure 6b ci-dessus).

En adoptant la coordonnée généralisée θ correspondant à la position du point $T = R\vec{u}(\theta)$ de tangence de la corde sur le cercle¹¹, et en désignant par $\vec{u}(\theta)$ le vecteur unitaire correspondant, la cinématique donne successivement :

$$\begin{cases} \vec{r}(\theta) = R\vec{u}(\theta) + (L - R\theta)\vec{u}(\theta + \pi/2) \\ \vec{v}(\theta, \dot{\theta}) = R\dot{\theta}\vec{u}(\theta + \pi/2) + (L - R\theta)(-\dot{\theta}\vec{u}(\theta)) - R\dot{\theta}\vec{u}(\theta + \pi/2) \\ \qquad \qquad \qquad = -(L - R\theta)\dot{\theta}\vec{u}(\theta + \pi/2) . \end{cases}$$

¹¹ Le point de contact T est, pour tout t , le centre de rotation : la balle parcourt une « développante » de cercle.

On en déduit le lagrangien \mathcal{L} (pas d'énergie potentielle) et le moment p_θ :

$$\mathcal{L}(\theta, \dot{\theta}) = \frac{1}{2} m(L - R\theta)^2 \dot{\theta}^2 \quad \Rightarrow \quad p_\theta \equiv \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = m(L - R\theta)^2 \dot{\theta} .$$

Malgré l'absence d'énergie potentielle, \mathcal{L} dépend explicitement de θ et l'équation de Lagrange s'écrit :

$$\frac{dp_\theta}{dt} = \frac{\partial \mathcal{L}}{\partial \theta} = -R(L - R\theta) \dot{\theta}^2 \quad \Rightarrow \quad (L - R\theta) \ddot{\theta} - R\dot{\theta}^2 = 0 .$$

Cette équation s'intègre aisément en exprimant $\ddot{\theta}/\dot{\theta}$, et on obtient l'intégrale première :

$$(L - R\theta) \dot{\theta} = cste ,$$

laquelle n'est autre que la vitesse¹², que nous noterons v_0 . Cette équation s'intègre elle-même aisément, et conduit à la loi horaire :

$$\theta(t) = \frac{L}{R} (1 - \sqrt{1 - t/\tau}), \quad \text{où} \quad \tau = \frac{L^2}{2Rv_0}$$

est le temps au bout duquel $L - R\theta$ s'annule, et où la balle heurte le mât¹³.

Notons bien, à la lumière des exemples précédents, que le formalisme lagrangien n'a que peu d'intérêt pour un système à un degré de liberté, pour lequel la conservation de l'énergie mécanique fournit immédiatement une description minimale du mouvement. Nous allons donc nous intéresser maintenant à deux problèmes à deux degrés de liberté.

c) Le pendule d'axe mobile

Un pendule pesant de masse m et de longueur L est accroché sur un « chariot », de masse M , libre de se mouvoir sur un axe horizontal. On choisit comme coordonnées généralisées l'écart du pendule à la verticale Oz , et l'abscisse x d'un point du chariot. On écrit simplement leurs vitesses respectives :

$$\vec{v}_{\text{ch}} = \dot{x} \vec{u}_x \quad \text{et} \quad \vec{v}_{\text{pend}} = (\dot{x} + L \cos \theta) \dot{\theta} \vec{u}_x + L \sin \theta \dot{\theta} \vec{u}_z .$$

Le lagrangien s'écrit donc, en rassemblant les termes :

$$\mathcal{L} = \frac{M + m}{2} \dot{x}^2 + \frac{m}{2} (L^2 \dot{\theta}^2 + 2L \cos \theta \dot{\theta} \dot{x}) - mgL(1 - \cos \theta) .$$

Le moments conjugués sont :

$$\begin{cases} p_x = \frac{\partial \mathcal{L}}{\partial \dot{x}} = (M + m)\dot{x} + mL \cos \theta \dot{\theta} \\ p_\theta = \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = mL^2 \dot{\theta} + mL \cos \theta \dot{x} \end{cases}$$

Comme la variable x ne figure pas dans \mathcal{L} (c'est une variable cyclique), la quantité $\partial \mathcal{L} / \partial x$ est nulle, et le moment conjugué p_x est conservé. Il est évident sur son expression qu'il s'agit de la

¹² La force de liaison, portée par la corde, étant en chaque instant perpendiculaire à la vitesse, ne travaille pas.

¹³ On constate ainsi que la vitesse angulaire croît rapidement avec le temps, ce qui est souvent interprété à tort comme une conséquence de la conservation du moment cinétique. En effet la force n'est pas centrale et moment cinétique s'écrit $\sigma = (L - R\theta)^2 \dot{\theta}$, qui n'est visiblement pas une constante du mouvement.

quantité de mouvement totale, dans la direction horizontale, laquelle est conservée en raison de l'invariance par translation du système. On peut alors éliminer la variable x , en supposant en outre que p_x est initialement nul¹⁴, et remplacer partout \dot{x} par $-m/(M+m)L\cos\theta\dot{\theta}$. Si l'on fait ce remplacement dans le lagrangien plutôt que dans les équations du mouvement (c'est-à-dire on traite l'intégrale première comme une liaison holonome) on obtient alors :

$$\mathcal{L} = m \frac{m \sin^2 \theta + M}{M + m} L^2 \frac{\dot{\theta}^2}{2} + mgL \cos \theta \quad \text{et} \quad p_\theta = m \frac{m \sin^2 \theta + M}{M + m} L^2 \dot{\theta} .$$

On calcule alors $\partial\mathcal{L}/\partial\theta = m^2 L^2 / (M + m) \sin\theta \cos\theta \dot{\theta}^2 - mgL \sin\theta$, et après simplification, l'équation différentielle satisfaite par θ est la suivante :

$$\frac{m \sin^2 \theta + m}{M + m} L \ddot{\theta} + g \sin \theta = 0 .$$

Dans la limite des « petites oscillations », on peut *linéariser* cette équation, c'est-à-dire supposer que θ est un infiniment petit, ainsi que ses dérivées successives (elles sont de l'ordre de $\omega^p \theta$ si ω est la fréquence caractéristique). En ne conservant que les termes d'ordre 1, il vient :

$$\ddot{\theta} + \omega^2 \theta = 0 , \quad \text{avec} \quad \omega = \sqrt{\frac{g}{L}} \sqrt{\frac{M + m}{M}} .$$

Selon le rapport des deux masses, deux cas limites sont plus simples :

- Dans la limite où $M/m \rightarrow \infty$, ω tend vers la valeur $\omega_0 = \sqrt{g/L}$ du pendule simple, puisque la masse du pendule devient trop faible pour entraîner significativement le chariot (cf. figure 7a).
- Dans la limite $M/m \rightarrow 0$, c'est-à-dire lorsque le chariot est très léger devant le pendule, on obtient un mouvement rapide du chariot, sous l'effet d'oscillations de très faible amplitude du pendule. La fréquence est élevée car c'est le poids de la grande masse m qui met en mouvement la masse plus faible M (cf. figure 7b).

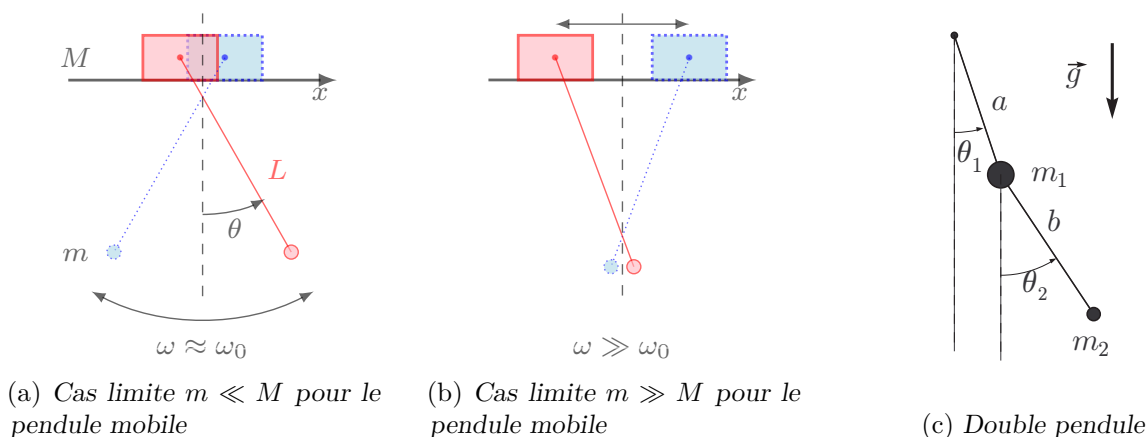


Fig. 7 – Illustration des exemples c) et d) dans cette section 5.5.

¹⁴ Dans le cas contraire, on pourra considérer que l'on a résolu le problème dans le référentiel galiléen qui se meut à la vitesse $v = p_x / (m + m')$ et en déduire simplement le mouvement dans le référentiel « fixe » initial.

d) **Le double pendule ★**

Le mouvement du « double pendule » dans un plan vertical (*cf.* figure 7c ci-dessus). En choisissant les écarts à la verticale θ_1 et θ_2 comme coordonnées généralisées, on a :

$$\begin{cases} \vec{r}_1 = -a \cos \theta_1 \vec{u}_z + a \sin \theta_1 \vec{u}_x \\ \vec{r}_2 = \vec{r}_1 - b \cos \theta_2 \vec{u}_z + b \sin \theta_2 \vec{u}_x \\ \vec{v}_1 = \dot{\theta}_1 (a \sin \theta_1 \vec{u}_z + a \cos \theta_1 \vec{u}_x) \\ \vec{v}_2 = \vec{v}_1 + \dot{\theta}_2 (b \sin \theta_2 \vec{u}_z + b \cos \theta_2 \vec{u}_x) \end{cases}$$

et donc :

$$\begin{cases} \vec{v}_1^2 = a^2 \dot{\theta}_1^2 \\ \vec{v}_2^2 = a^2 \dot{\theta}_1^2 + b^2 \dot{\theta}_2^2 + 2ab \cos(\theta_1 - \theta_2) \dot{\theta}_1 \dot{\theta}_2 . \end{cases}$$

On en déduit :

$$K = \frac{m_1 + m_2}{2} a^2 \dot{\theta}_1^2 + \frac{m_2}{2} b^2 \dot{\theta}_2^2 + m_2 ab \cos(\theta_1 - \theta_2) \dot{\theta}_1 \dot{\theta}_2$$

$$V(\theta_1, \theta_2) = -m_1 g a \cos \theta_1 - m_2 g (a \cos \theta_1 + b \cos \theta_2)$$

puis

$$p_1 \equiv \frac{\partial \mathcal{L}}{\partial \dot{\theta}_1} = (m_1 + m_2) a^2 \dot{\theta}_1 + m_2 ab \dot{\theta}_2 \cos(\theta_1 - \theta_2)$$

$$p_2 \equiv \frac{\partial \mathcal{L}}{\partial \dot{\theta}_2} = m_2 b^2 \dot{\theta}_2 + m_2 ab \dot{\theta}_1 \cos(\theta_1 - \theta_2)$$

et les équations de Lagrange :

$$\begin{aligned} \frac{dp_1}{dt} = \frac{\partial \mathcal{L}}{\partial \theta_1} &= (m_1 + m_2) a^2 \ddot{\theta}_1 + m_2 ab \ddot{\theta}_2 \cos(\theta_1 - \theta_2) - m_2 ab \dot{\theta}_2 (\dot{\theta}_1 - \dot{\theta}_2) \sin(\theta_1 - \theta_2) \\ &= -m_2 ab \dot{\theta}_1 \dot{\theta}_2 \sin(\theta_1 - \theta_2) + (m_1 + m_2) a g \sin \theta_1 , \end{aligned}$$

$$\begin{aligned} \frac{dp_2}{dt} = \frac{\partial \mathcal{L}}{\partial \theta_2} &= m_2 b^2 \ddot{\theta}_2 + m_2 ab \ddot{\theta}_1 \cos(\theta_1 - \theta_2) - m_2 ab \dot{\theta}_1 (\dot{\theta}_1 - \dot{\theta}_2) \sin(\theta_1 - \theta_2) \\ &= m_2 ab \dot{\theta}_1 \dot{\theta}_2 \sin(\theta_1 - \theta_2) + m_2 b g \sin \theta_2 , \end{aligned}$$

et après simplification :

$$\begin{aligned} (m_1 + m_2) \ddot{\theta}_1 + m_2 \frac{b}{a} (\cos(\theta_1 - \theta_2) \ddot{\theta}_2 + \sin(\theta_1 - \theta_2) \dot{\theta}_1 \dot{\theta}_2) + m_2 \frac{b}{a} \sin(\theta_1 - \theta_2) \dot{\theta}_2^2 &= (m_1 + m_2) \frac{g}{a} \sin \theta_1 \\ \ddot{\theta}_2 + \frac{a}{b} \cos(\theta_1 - \theta_2) \ddot{\theta}_1 - \frac{a}{b} \sin(\theta_1 - \theta_2) \dot{\theta}_1^2 &= \frac{g}{b} \sin \theta_2 . \end{aligned}$$

Ces équations différentielles couplées et *non-linéaires* du second ordre ne peuvent être résolues de façon analytique, mais si l'on se limite aux petites oscillations, il est possible de les « linéariser »¹⁵.

Cela conduit aux équations notablement plus simples :

$$\begin{aligned} (m_1 + m_2) \ddot{\theta}_1 + m_2 \frac{b}{a} \ddot{\theta}_2 &= (m_1 + m_2) \frac{g}{a} \theta_1 \\ \ddot{\theta}_2 + \frac{a}{b} \ddot{\theta}_1 &= \frac{g}{b} \theta_2 , \end{aligned}$$

dont nous laissons aux lecteurs la recherche des « modes propres », et la discussions physique pour les cas limites $m_1 \ll m_2$ ou $m_2 \ll m_1$.

¹⁵ C'est à dire négliger les termes d'ordre 2 et plus.

5.6 Cas des coordonnées dépendant explicitement du temps

Dans le cas plus général où la relation $\{q_i\} \mapsto \{\vec{R}_\alpha\}$ dépend explicitement du temps, les équations (9) et (10), ne sont évidemment plus valables, mais on peut montrer, et nous admettrons dans ce cours, sans poser les calculs correspondants, qui sont élémentaires mais affreux (il y a treize termes!), que les équations (7), avec les forces généralisées de (11), restent valables. On notera que en général, les forces généralisées \mathcal{F}_k dépendent alors du temps.

En outre, une éventuelle dépendance en temps, explicite, ou via \vec{R}_α , du potentiel V n'affectera en rien la validité des équations (12) et (13). On en déduit le résultat plus général :

$$\mathcal{L}(\{q_i\}, \{\dot{q}_i\}, t) = K(\{q_i\}, \{\dot{q}_i\}, t) - V(\{q_i\}, t) .$$

5.7 Potentiels généralisés

L'expression que nous avons obtenue pour la fonction de Lagrange nécessite d'être encore étendue aux cas, rares mais physiquement importants où les forces *non-dissipatives* dépendent de la vitesse : c'est le cas de la force de Lorentz, et celui de la force de Coriolis¹⁶.

Si nous voulons rendre compte de cette dépendance, nous pouvons chercher un « potentiel généralisé » $V(\{q_i\}, \{\dot{q}_i\}, t)$ dépendant des vitesses, et éventuellement du temps. Il faut remarquer que la structure des équations de Lagrange sera automatiquement conservée si, dans l'équation (13), les forces \mathcal{F}_k s'écrivent de façon plus générale :

$$\mathcal{F}_k = \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{q}_k} \right) - \frac{\partial V}{\partial q_k} ,$$

Si nous voulons obtenir, pour les cas physique évoquées ci-dessus, une force linéaire en vitesse, nous voyons que le potentiel V doit lui-même être linéaire¹⁷ en $\{\dot{q}_i\}$, soit :

$$V(\{q_i\}, \{\dot{q}_i\}, t) = \sum_i a_i(\{q_i\}, t) \dot{q}_i \quad \text{et donc} \quad \mathcal{F}_k = \frac{da_k}{dt} - \sum_i \frac{\partial a_i}{\partial q_k} \dot{q}_i = \sum_i \left(\frac{\partial a_k}{\partial q_i} - \frac{\partial a_i}{\partial q_k} \right) \dot{q}_i + \frac{\partial a_k}{\partial t} .$$

On observe que la force ainsi obtenue repose nécessairement sur une relation $\dot{q} \mapsto \mathcal{F}$ linéaire *antisymétrique*, généralisant le produit vectoriel qu'on obtient en coordonnées cartésiennes.

5.8 Forme générale du lagrangien

En récapitulant tout ce qui précède, on a en définitive une expression générale du lagrangien :

Forme générale du lagrangien

Si l'évolution du système est régie par des forces intérieures ou extérieures, éventuellement linéaires en vitesse, non dissipatives, on peut introduire un potentiel généralisé $V(\{q_i\}, \{\dot{q}_i\}, t)$, et écrire son lagrangien :

$$\mathcal{L}(\{q_i\}, \{\dot{q}_i\}, t) = K(\{q_i\}, \{\dot{q}_i\}, t) - V(\{q_i\}, \{\dot{q}_i\}, t) .$$

et toujours appliquées les équations d'Euler-Lagrange (7), où cette fois les dérivées totales comme partielles doivent tenir compte des dépendances explicites vis-à-vis des vitesses et du temps.

¹⁶ Notons que toute force dépendant de la vitesse qui ne serait pas normale à celle-ci serait forcément dissipative!

¹⁷ Un potentiel non-linéaire en $\{\dot{q}_i\}$ ferait apparaître l'accélération \ddot{q}_i dans l'expression des forces généralisées.

Exemple important : particule chargée dans un champ électromagnétique

On sait depuis les classes élémentaires qu'une particule chargée (dont nous noterons Q la charge électrique pour éviter les confusions avec les coordonnées généralisées), est soumise à la « force de Lorentz » :

$$\vec{F}_L = Q(\vec{E} + \vec{v} \times \vec{B}) = Q\left(\left(-\frac{\partial \vec{A}}{\partial t} - \overrightarrow{\text{grad}} \Phi\right) + \vec{v} \times \overrightarrow{\text{rot}} \vec{A}\right), \quad (14)$$

où l'on a écrit les champs en fonction des potentiels scalaire Φ et vecteur \vec{A} associés.

On notera que le gradient s'écrit simplement $\partial \cdot / \partial \vec{r}$, mais le produit vectoriel est plus subtil, et peut poser des problèmes techniques en coordonnées généralisées. Aussi nous contenterons nous ici des coordonnées cartésiennes.

Cette force peut être retrouvée en formalisme lagrangien en utilisant le potentiel généralisé :

$$V(\vec{r}, \vec{v}, t) = Q\left(\Phi(\vec{r}, t) - \vec{v} \cdot \vec{A}(\vec{r}, t)\right). \quad (15)$$

En effet, on a calculé alors¹⁸ :

$$\vec{p} = \frac{\partial \mathcal{L}}{\partial \vec{v}} = m\vec{v} + Q \vec{A}(\vec{r}, t) \quad \text{et} \quad \frac{\partial \mathcal{L}}{\partial \vec{r}} = -Q \overrightarrow{\text{grad}}\left(\Phi(\vec{r}, t) - \vec{v} \cdot \vec{A}(\vec{r}, t)\right)$$

et l'équation de Lagrange s'écrit :

$$m \frac{d\vec{v}}{dt} + Q \frac{d\vec{A}}{dt} = -Q \overrightarrow{\text{grad}} \Phi + Q \overrightarrow{\text{grad}}(\vec{v} \cdot \vec{A})$$

En utilisant alors les formules bien connues :

$$\begin{aligned} \frac{d\vec{A}}{dt} &= \frac{\partial \vec{A}}{\partial t} + (\vec{v} \cdot \overrightarrow{\text{grad}}) \vec{A}, \\ \overrightarrow{\text{grad}}(\vec{a} \cdot \vec{b}) &= (\vec{a} \cdot \overrightarrow{\text{grad}}) \vec{b} + (\vec{b} \cdot \overrightarrow{\text{grad}}) \vec{a} + \vec{a} \times (\overrightarrow{\text{rot}} \vec{b}) + \vec{b} \times (\overrightarrow{\text{rot}} \vec{a}), \end{aligned}$$

il vient :

$$\begin{aligned} m \frac{d\vec{v}}{dt} + Q \left(\frac{\partial \vec{A}}{\partial t} + \underbrace{(\vec{v} \cdot \overrightarrow{\text{grad}}) \vec{A}} \right) &= -Q \overrightarrow{\text{grad}} \Phi + Q \overrightarrow{\text{grad}}(\vec{v} \cdot \vec{A}) \\ &= -Q \overrightarrow{\text{grad}} \Phi + Q \left(\underbrace{(\vec{v} \cdot \overrightarrow{\text{grad}}) \vec{A}} + \vec{v} \wedge \overrightarrow{\text{rot}} \vec{A} \right), \end{aligned}$$

et après simplification et réagencement, on retrouve exactement la force de Lorentz (14).

Invariance de jauge On pourrait s'étonner de ce que le Lagrangien fasse intervenir les potentiels alors que ceux-ci sont connus pour n'être pas totalement univoques...

Il est bien connu en particulier que la « transformation de jauge » :

$$\vec{A} \mapsto \vec{A}' = \vec{A} + \overrightarrow{\text{grad}} \chi(\vec{r}, t) \quad \text{et} \quad \Phi \mapsto \Phi' = \Phi - \frac{\partial \chi}{\partial t}(\vec{r}, t)$$

¹⁸ Noter que l'impulsion \vec{p} ne coïncide plus avec la quantité de mouvement $m\vec{v}$.

a la propriété de changer les potentiels sans modifier les champs électromagnétiques qu'ils décrivent. On constate aisément que s'il l'on forme le lagrangien \mathcal{L}' à l'aide des nouveaux potentiels, on obtient le lagrangien :

$$\mathcal{L}' = \mathcal{L} + Q \left(\frac{\partial \chi}{\partial t}(\vec{r}, t) + \vec{v} \cdot \overrightarrow{\text{grad}} \chi(\vec{r}, t) \right) = \mathcal{L} + \frac{d}{dt} \left(Q \chi(\vec{r}, t) \right).$$

Il ne diffère du lagrangien initial que par l'ajout d'une dérivée totale par rapport au temps, dont nous avons vu au § 4.3 qu'il ne modifiait pas les équations d'Euler–Lagrange.

Le mouvement dans un référentiel en rotation uniforme, dans l'approche lagrangienne, sera étudié en TD. Pour l'approche hamiltonienne, voir § 11.3.2 ci-dessous.

6 Symétries et lois de conservation

6.1 Notion d'intégrale première

Nous avons déjà utilisé au § 5.1 certaines lois d'invariance fondamentales pour en inférer des informations sur la structure du lagrangien, dans le cas particulier d'une particule libre. Cette démarche est en fait beaucoup plus générale, et se traduit par la mise en évidence « d'intégrales premières » ou « constantes du mouvement », qui sont des variables dynamiques, c'est-à-dire des quantités de la forme $f(\{q_i\}, \{\dot{q}_i\}, t)$, qui ne varient pas au cours du mouvement.

Variables cycliques Un cas particulier, à la fois trivial et très intéressant du point de vue du formalisme hamiltonien (cf. § 14) est celui d'une variable q_i telle que $\partial \mathcal{L} / \partial q_i = 0$, qui est alors dite **cyclique**. L'équation de Lagrange correspondante donne immédiatement la conservation de « l'impulsion généralisée » correspondante $p_i = \partial L / \partial \dot{q}_i = \text{cste}$.

Cette situation est de façon générale celle de la variable d'angle φ dans un mouvement à force centrale, et l'impulsion généralisée correspondante n'est autre que le moment cinétique orbital, sur lequel nous reviendrons au § 6.6.

6.2 Énergie mécanique

Parmi les éventuelles constantes du mouvement, l'énergie joue un rôle particulier. Elle s'introduit naturellement en formant la dérivée temporelle (totale) du lagrangien :

$$\begin{aligned} \frac{d\mathcal{L}}{dt} &= \frac{\partial \mathcal{L}}{\partial t} + \sum_i \frac{\partial \mathcal{L}}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{d\dot{q}_i}{dt} = \frac{\partial \mathcal{L}}{\partial t} + \sum_i \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \dot{q}_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{d}{dt} (\dot{q}_i) \\ &= \frac{\partial \mathcal{L}}{\partial t} + \sum_i \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i \right) = \frac{\partial \mathcal{L}}{\partial t} + \frac{d}{dt} \left(\sum_i p_i \dot{q}_i \right) \end{aligned}$$

On peut en déduire que la quantité :

$$\mathcal{E} = \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i - \mathcal{L} = \sum_i p_i \dot{q}_i - \mathcal{L} \quad \text{vérifie} \quad \frac{d\mathcal{E}}{dt} = - \frac{\partial \mathcal{L}}{\partial t}. \quad (16)$$

et notamment dans le cas où le lagrangien ne dépend pas explicitement du temps ($\partial \mathcal{L} / \partial t = 0$), elle est indépendante du temps. Ainsi, l'invariance par rapport au temps se traduit-elle par la conservation d'une quantité, qui n'est autre que de l'énergie mécanique¹⁹.

¹⁹ C'est pour cette raison que l'on prend un lagrangien homogène à une énergie.

Notamment, dans le cas où le lagrangien est de forme standard (c'est-à-dire K forme quadratique des \dot{q}_i et V indépendant des vitesses), on montre à l'aide du théorème d'Euler sur les fonctions homogènes (de degré 2 ici) que :

$$\sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i = \sum_i \frac{\partial K}{\partial \dot{q}_i} \dot{q}_i = 2K,$$

et on peut alors écrire :

$$\mathcal{E} = 2K - (K - V) = K + V,$$

qui est bien sûr l'expression habituelle de l'énergie mécanique.

Cas des potentiels généralisés Dans le cas plus général, où l'énergie cinétique contient aussi des termes *linéaires* vis-à-vis des \dot{q}_i , et/ou lorsqu'on a un potentiel généralisé lui-même linéaire en \dot{q}_i , soit $\mathcal{L} = K_2 - V_0 + \mathcal{L}_1$, où les indices traduisent ici le degré de la dépendance en \dot{q}_i , on a $\sum_i \partial \mathcal{L} / \partial \dot{q}_i \dot{q}_i = 2K_2 + \mathcal{L}_1$, et $\mathcal{E} = K_2 + V_0$, dont l'interprétation dépend de la situation considérée.

Par exemple, pour une particule chargée dans des champs électrique et magnétique statiques, le terme en $\vec{v} \cdot \vec{A}$ du potentiel (15) s'élimine, et on retrouve bien l'énergie mécanique $\mathcal{E} = m\vec{v}^2/2 + \Phi$.

6.3 Théorème de Noether

Les propriétés d'invariance vis-à-vis des symétries sont systématiquement reliées à des quantités conservées par un théorème général, appelé « théorème de Noether ». Si celui-ci ne révèle toute sa richesse que dans le cadre de la théorie des champs, il n'est pas inutile d'en donner un aperçu dans le cadre le plus élémentaire qui nous intéresse ici.

Nous considérons donc un système physique dont la configuration est définie par un jeu de coordonnées $\{q_i\}$. Nous savons que ce jeu de coordonnées n'est pas unique : il existe au contraire une infinité de transformations *bijectives* $\{q_i\} \mapsto \{\tilde{q}_i\}$ permettant, au moins formellement, d'écrire un lagrangien \mathcal{L}' et d'en déduire des équations de Lagrange décrivant le même mouvement²⁰. Pour cela, on admet que le lagrangien est une fonction « intrinsèque » de l'espace des configurations, indépendant de la paramétrisation choisie. On parle alors d'*invariance ponctuelle*, qui se traduit par l'identité générale :

$$\mathcal{L}'(\{\tilde{q}_i\}, \{\dot{\tilde{q}}_i\}, t) \equiv \mathcal{L}(\{q_i\}, \{\dot{q}_i\}, t)$$

où les $\{q_i\}$ et $\{\dot{q}_i\}$ ne sont pas ici des variables indépendantes mais des fonctions des $\{\tilde{q}_i\}$ (et des $\{\dot{\tilde{q}}_i\}$ en ce qui concerne $\{\dot{q}_i\}$).

Nous nous restreignons ici à une classe très particulière de telles transformations qui possèdent la propriété d'*invariance formelle* : elle préservent la forme du Lagrangien, c'est-à-dire que l'expression de \mathcal{L}' en fonctions des nouvelles coordonnées est *mathématiquement* identique à celle de \mathcal{L} en fonction des anciennes : une telle possibilité, lorsqu'elle existe, est la manifestation explicite d'une certaine invariance dans le système.

Nous supposons en outre que nous avons affaire à un *ensemble continu et dérivable* de telles transformations, c'est-à-dire que l'on peut définir une fonction continue d'un paramètre réel s qui à chaque valeur associe une nouvelle transformation $\{q_i\} \mapsto \{\tilde{q}_i\}(s)$ des coordonnées, et telle que $\{\tilde{q}_i\}(0) = \{q_i\}$ et idem pour les vitesses. Ceci exclut bien sûr des transformations discrètes par

²⁰ Comme indiqué précédemment, cette indépendance formelle vis-à-vis des jeux de coordonnées est l'un des intérêts du formalisme lagrangien, car il permet de choisir un système adapté, qui simplifie les équations à résoudre ; un nouvel exemple particulièrement probant en sera donné dans le § 8.1 ci-dessous.

essence, comme par exemple des réflexions qui changeraient une ou plusieurs des coordonnées en leur opposé.

Dans ce cadre, le théorème de Noether affirme qu'il existe une constante du mouvement associée à cette invariance, qui s'écrit :

$$I = \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{\partial \tilde{q}_i}{\partial s} \Big|_{s=0} \quad (17)$$

et qui sera appelée « générateur infinitésimal » de la transformation étudiée²¹.

Après la démonstration, qui suit mais peut être omise en première lecture, nous présentons les 3 exemples essentiels (pour la mécanique) que sont l'énergie, l'impulsion, et le moment cinétique.

Ce paragraphe aurait dû être le premier exemple s'il n'était pas techniquement plus délicat.

Démonstration ★

Pour l'établir, il suffit d'utiliser les deux hypothèses essentielles : le lagrangien est à la fois ponctuellement invariant, et formellement invariant. Considérons alors la dérivée de $\mathcal{L}'(\{\tilde{q}_i\}, \{\dot{\tilde{q}}_i\}, t)$ par rapport à s : celle-ci est identiquement nulle en raison de l'invariance ponctuelle. Elle vérifie aussi :

$$0 = \frac{d\mathcal{L}'}{ds} = \sum_i \frac{\partial \mathcal{L}'}{\partial \tilde{q}_i} \frac{\partial \tilde{q}_i}{\partial s} + \frac{\partial \mathcal{L}'}{\partial \dot{\tilde{q}}_i} \frac{\partial \dot{\tilde{q}}_i}{\partial s} = \sum_i \frac{d}{dt} \left(\frac{\partial \mathcal{L}'}{\partial \dot{\tilde{q}}_i} \right) \frac{\partial \tilde{q}_i}{\partial s} + \frac{\partial \mathcal{L}'}{\partial \dot{\tilde{q}}_i} \frac{d}{dt} \left(\frac{\partial \tilde{q}_i}{\partial s} \right) = \frac{d}{dt} \left(\sum_i \frac{\partial \mathcal{L}'}{\partial \dot{\tilde{q}}_i} \frac{\partial \tilde{q}_i}{\partial s} \right),$$

où la condition d'évaluation en $s = 0$ est implicite, où l'on a simplement utilisé les (nouvelles) équations de Lagrange, et permuté les dérivations relatives à s et à t . En utilisant alors l'invariance formelle, qui assure que $\partial \mathcal{L}' / \partial \dot{\tilde{q}}_i = \partial \mathcal{L} / \partial \dot{q}_i$, on a bien obtenu la nullité de la dérivée temporelle de I .

NB Les hypothèses que nous avons faites sont en réalité plus fortes que nécessaire. L'invariance formelle globale n'est pas indispensable, puisqu'elle en se manifeste que par son comportement différentiel en $s = 0$.

6.4 Énergie

Nous avons déjà montré au § 6.2 qu l'invariance par rapport à des translations dans le temps, matérialisée par l'indépendance en t du lagrangien entraînait la conservation de l'énergie $\mathcal{E} = \sum_i p_i \dot{q}_i - \mathcal{L}$

La mise évidence de cette conservation comme exemple du théorème de Noether est techniquement assez délicate, car le changement de coordonnées continu correspondant doit utiliser un paramètre s qui est le temps lui-même, soit $q_i \mapsto \tilde{q}_i = q_i + \dot{q}_i s$. Cela complique la démonstration car il faut en principe distinguer la dérivation partielle par rapport à s de l'équation (17) de la dérivation par rapport à t . Pour ne pas alourdir inutilement, nous ne développerons pas cette démonstration .

6.5 Impulsion

Un autre cas fréquent de symétrie continue est l'invariance par translation dans l'espace, qui est vérifiée pour un ensemble de particules en interaction mutuelle, et traduit l'homogénéité de l'espace pour un système isolé. En considérant un lagrangien standard et en prenant comme coordonnées généralisées les positions cartésiennes $\{\vec{r}_\alpha\}$ des particules, on a :

$$\mathcal{L}(\{\vec{r}_\alpha\}, \{\dot{\vec{r}}_\alpha\}, t) = \sum_\alpha \frac{m_\alpha}{2} \dot{\vec{r}}_\alpha^2 - V(\{\vec{r}_\alpha\}),$$

²¹ Pour l'origine de ce terme, on se reportera au § 11.2, à la page 54.

où l'élément clef est bien sûr le fait que V ne dépend que des positions *relatives* des particules, donc est invariant par *translation d'ensemble* $\vec{r}_\alpha \mapsto \vec{r}_\alpha + \vec{a}$. Une transformation continue possible est alors une translation de la forme $\vec{r}_\alpha \mapsto \vec{Q}_\alpha = \vec{r}_\alpha + s\vec{u}$, où \vec{u} est un vecteur unitaire (uniforme et constant) et indépendant de α . On a alors $d\vec{Q}_\alpha/ds = \vec{u}$, et la quantité conservée s'écrit donc :

$$I = \sum_{\alpha} \frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_{\alpha}} \cdot \vec{u} = \sum_{\alpha} (\vec{p}_{\alpha} \cdot \vec{u}) = \left(\sum_{\alpha} \vec{p}_{\alpha} \right) \cdot \vec{u} \quad \text{où} \quad \vec{p}_{\alpha} = m_{\alpha} \dot{\vec{r}}_{\alpha}.$$

Comme \vec{u} est quelconque, l'invariant de Noether est *l'impulsion d'ensemble* :

$$\vec{P} = \sum_{\alpha} m_{\alpha} \dot{\vec{r}}_{\alpha} = \sum_{\alpha} \vec{p}_{\alpha},$$

qui, dans le cas considéré ici, est la *quantité de mouvement* totale.

Rappelons que l'on appelle « **moment conjugué** » ou « **impulsion généralisée** » associé(e) à la variable q_i la quantité $p_i = \partial \mathcal{L} / \partial \dot{q}_i$, qui coïncide avec la quantité de mouvement correspondante dans le cas particulier, considéré ici, du lagrangien standard et des coordonnées cartésiennes.

En revanche, on perd cette propriété si l'on utilise des coordonnées curvilignes, ou si le lagrangien implique un potentiel généralisé comme celui de l'électromagnétisme. Dans ce dernier cas, très important, on a pour chaque particule : $\vec{p}_{\alpha} = m_{\alpha} \vec{v}_{\alpha} + Q_{\alpha} \vec{A}(\vec{r}_{\alpha}, t)$, et c'est la somme de ces impulsions qui est potentiellement conservée. Toutefois, pour des champs \vec{A} et $\vec{\Phi}$ non triviaux, le système des particules seules n'est pas invariant par translation et la conservation de l'impulsion globale nécessite d'introduire celle du champ électromagnétique (pas étudiée ici).

Remarques

- Dans un champ électrostatique $\vec{E} = E \vec{u}_z$ (ou de pesanteur) uniforme selon x - y et dirigé selon la direction de z , le potentiel Φ dépend uniquement de z , et $\vec{A} = \vec{0}$ on a donc invariance seulement dans les directions perpendiculaires, et $\vec{P}_{\perp} = \text{cste}$.
- Dans un champ magnétostatique uniforme $\vec{B} = B \vec{u}_z$, on peut prendre $\vec{A} = Bx \vec{u}_y$, et alors $P_x = \sum_{\alpha} m_{\alpha} \dot{x}_{\alpha}$, et même chose pour P_z , tandis que $P_y = \sum_{\alpha} m_{\alpha} \dot{y}_{\alpha} + Q_{\alpha} Bx_{\alpha}$, et on constate que P_z (bien sûr !) et P_y sont des constantes du mouvement. Un autre choix pour \vec{A} aurait cependant conduit à d'autres constantes (*cf. infra*).

6.6 Moment cinétique

Un autre cas important est celui de l'invariance par rotation, observée tant pour une particule en mouvement dans un potentiel central, que pour un système de particules en interaction mutuelle. Dans ce dernier cas, elle traduit l'isotropie de l'espace pour un système isolé.

En décrivant à nouveau la configuration à l'aide les coordonnées cartésiennes, on peut écrire une rotation *infinitésimale* d'axe \vec{u} et d'angle s : $\vec{r}_{\alpha} \mapsto \vec{Q}_{\alpha} = \vec{r}_{\alpha} + s\vec{u} \wedge \vec{r}_{\alpha}$. Dans ce cas, on a $d\vec{Q}_{\alpha}/ds = d(s\vec{u} \wedge \vec{r}_{\alpha})/ds$, et la quantité conservée s'écrit :

$$I = \sum_{\alpha} \vec{p}_{\alpha} \cdot (\vec{u} \wedge \vec{r}_{\alpha}) = \sum_{\alpha} (\vec{r}_{\alpha} \wedge \vec{p}_{\alpha}) \cdot \vec{u},$$

et comme \vec{u} est quelconque, on obtient la conservation du « moment cinétique » total :

$$\vec{L} = \sum_{\alpha} (\vec{r}_{\alpha} \wedge \vec{p}_{\alpha}).$$

Dans le cas d'un lagrangien standard, ce moment cinétique est simplement le moment $\vec{\sigma}$ de la quantité de mouvement, appelé « moment angulaire »²²; toutefois, il est construit à partir des impulsions, et pourra donc différer de $\vec{\sigma}$ lorsque les impulsions diffèrent de la quantité de mouvement.

Notons que, de même que pour l'impulsion, l'invariance de rotation peut n'être vérifiée qu'autour de certains axes, donc certaines valeurs de \vec{u} : on n'aura alors conservation que des composantes du moment cinétique total noté $\vec{\sigma}$ correspondantes. C'est le cas en particulier d'une particule en mouvement dans un champ magnétique uniforme et constant de direction Oz , qui est visiblement invariant par rotation autour de Oz . Si l'on choisit un lagrangien vérifiant cette symétrie, avec le potentiel vecteur $\vec{A} = \vec{B} \times \vec{r}/2$, on obtiendra la conservation de L_z , qui s'écrit :

$$L_z = \vec{r} \times \vec{p} \cdot \vec{u}_z = x p_y - y p_x = x(m\dot{y} + Q \frac{Bx}{2}) - y(m\dot{x} - Q \frac{By}{2}) = \sigma_z + QB \frac{x^2 + y^2}{2} = cste ,$$

où σ_z est le moment de la quantité de mouvement, mais *n'est pas* une constante du mouvement !

7 Variations sous contraintes

Il arrive souvent que l'on ait à résoudre un problème de minimisation²³ sous *contrainte(s)*, comme par exemple pour trouver le format de la boîte de conserve le plus économe en fer pour un volume donné... Plus près de notre sujet, c'est notamment le cas si le mouvement d'un système doit satisfaire certaines conditions, comme des liaisons non-holonomes que l'on a pas pu, ou des liaisons holonomes que l'on a pas voulu, intégrer dans la définition des coordonnées généralisées. La méthode générale pour résoudre ce genre de problème est la méthode dite « des multiplicateurs de Lagrange » (en anglais ULM pour *Undetermined Lagrange Multipliers*) que nous introduisons ici.

7.1 Multiplicateurs de Lagrange pour une fonction

À des fins purement pédagogiques que nous allons dans paragraphe introduire les multiplicateurs de Lagrange dans un cas plus simple, celui de la minimisation d'une fonction ordinaire et non d'une fonctionnelle.

7.1.1 Cas avec une contrainte

Soient donc deux fonctions $F, C : \mathbb{R}^n \rightarrow \mathbb{R}$, et on cherche un point \vec{u}_0 qui réalise un minimum de $F(\vec{u})$ en même temps que la condition $C(\vec{u}) = cste$, où la constante peut être choisie nulle sans restriction. Les fonctions F et C sont supposées de classe \mathcal{C}^1 sur le domaine utile. Le problème considéré est donc de trouver un extremum de $F(\vec{u})$ sur une iso-surface donnée S de C , en supposant qu'il admet au moins une solution.

De façon assez naturelle, une solution²⁴ est un point \vec{u}_0 où le gradient de F est *normal* à l'iso-surface S : de la sorte, toute variation infinitésimale $d\vec{u}$ autour de \vec{u}_0 respectant la contrainte sera tangente à la surface et vérifiera donc $dF = \overrightarrow{\text{grad}} F \cdot d\vec{u} = 0$. Or la normale à S est elle-même dirigée par le gradient de la fonction C . On en déduit alors que les deux gradients sont colinéaires, et qu'il existe donc une constante λ telle que :

$$\overrightarrow{\text{grad}} F = \lambda \overrightarrow{\text{grad}} C , \quad (18)$$

²² Ou en physique atomique moment cinétique « orbital » pour le distinguer de celui dû au spin des particules.

²³ Nous emploierons abusivement *minimum* et *minimiser* là où on devrait dire *extremum* ou *stationnaire*.

²⁴ On exclut le cas où l'extremum serait atteint sur les bord du domaine.

et la constante λ , qui sera déterminée en même temps que \vec{u}_0 , est appelée « multiplicateur de Lagrange ». Cela conduit à chercher des points (et de valeurs de λ) où le gradient de $F - \lambda C$ s'annule. Cela ramène le problème initial à celui de la minimisation de la fonction :

$$F' = F - \lambda C ,$$

On peut s'assurer que le problème est maintenant un problème à $n + 1$ inconnues, et que l'on a bien $n + 1$ équations, à savoir les n relations de (18), plus la relation de contrainte $C(\vec{u}_0) = 0$.

7.1.2 Cas avec plusieurs contraintes

Au lieu d'une simple contrainte décrite par C , on a maintenant $m < n$ contraintes, soit :

$$C_i(\vec{u}) = 0 \quad \text{pour } i = 1 \dots m , \quad (19)$$

les fonctions $C_i : \mathbb{R}^n \rightarrow \mathbb{R}$ étant toujours supposées \mathcal{C}^1 , et indépendantes. Cette fois, le gradient de F ne peut évidemment pas être colinéaire avec le gradient de chacune des contraintes, mais l'annulation de $dF = \overrightarrow{\text{grad}} F \cdot d\vec{u}$ pour tout $d\vec{u}$ vérifiant l'ensemble des contraintes doit donner un résultat similaire.

A cet effet, considérons le problème mathématique général suivant : soit $(\vec{v}_1, \dots, \vec{v}_m)$ une famille de vecteurs d'un \mathbb{R} -espace vectoriel de dimension n . Considérons le supplémentaire orthogonal \mathcal{E}_p de chacun des vecteurs \vec{v}_i , et l'ensemble :

$$\mathcal{E} = \bigcap_{i=1}^m \mathcal{E}_p$$

des vecteurs orthogonaux à tous les vecteurs \vec{v}_i . Cet ensemble est un espace vectoriel qui n'est autre que le supplémentaire orthogonal de l'espace $\mathcal{F}(\vec{v}_1, \dots, \vec{v}_m)$ engendré par la famille de vecteurs considérés. En effet, on a bien sûr, pour tout vecteur \vec{V} de \mathbb{R}^n :

$$\forall i \in \{1, \dots, m\} \quad \vec{v}_i \cdot \vec{V} = 0 \quad \iff \quad \forall (\lambda_1 \dots \lambda_m) \in \mathbb{R}^m \quad \sum_{i=1}^m \lambda_i \vec{v}_i \cdot \vec{V} = 0 .$$

Si l'on applique ce lemme très général aux vecteurs $\vec{v}_p \equiv \overrightarrow{\text{grad}} C_p$ et aux vecteurs $\vec{V} \equiv \vec{u}$, on constate que le vecteur $\overrightarrow{\text{grad}} F$ doit appartenir à l'espace \mathcal{F} engendré par les gradients des contraintes. On en déduit alors le « théorème des multiplicateurs de Lagrange » selon lequel les extrema de F sous les contraintes $C_i(\vec{u}) = 0$ sont les point \vec{u} tels qu'il existe $(\lambda_1, \dots, \lambda_m)$ vérifiant :

$$\overrightarrow{\text{grad}} F(\vec{u}) = \sum_{i=1}^m \lambda_i \overrightarrow{\text{grad}} C_i(\vec{u}) . \quad (20)$$

Par conséquent, on généralise le résultat à une dimension, en identifiant le problème initial à celui de la recherche des extrema de la fonction :

$$F' = F - \sum_{i=1}^m \lambda_i C_i .$$

Les multiplicateurs de Lagrange $(\lambda_1 \dots \lambda_m)$ sont alors des inconnues supplémentaires qu'il reste à déterminer. On a ainsi $n + m$ inconnues, qui doivent vérifier à la fois les n équations (20) et les m équations de contrainte (19), et sont donc bien déterminées.

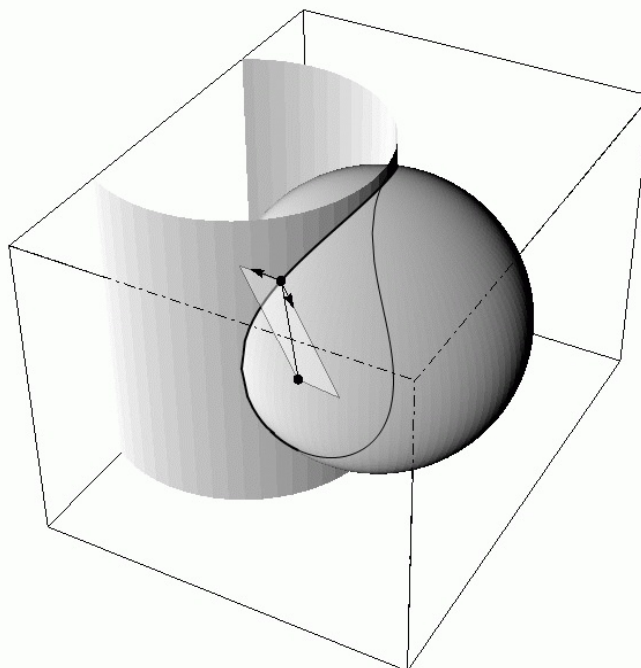


Fig. 8 – Exemple : la fonction F à minimiser est la distance au point situé au premier plan. Les contraintes sont matérialisées par la sphère et le cylindre ($n = 3, m = 2$). L'intersection des deux surfaces est une ligne gauche, dont la tangente est perpendiculaire aux deux vecteurs représentant la normale aux surfaces de contrainte. La ligne joignant les deux points matérialise la direction $\overrightarrow{\text{grad}} F$. Cette ligne étant dans le plan engendré par les normales, elle est orthogonale à la courbe, et on a donc bien trouvé le minimum !

Exemples simples Un problème bien connu est celui de la « boîte de conserve » cylindrique, dont on doit minimiser la surface à volume fixé. Avec des notations évidentes, on a un volume $\mathcal{V} = \pi a^2 h$, et une aire $\mathcal{A} = 2\pi a^2 + 2\pi a h$. Au lieu d'exprimer h en fonctions de a pour un volume \mathcal{V} donné et de reporter dans \mathcal{A} , il est intéressant de chercher l'extremum de $\mathcal{A}' = \mathcal{A} - \lambda \mathcal{V}$, défini par les équations (en supposant a fixé) :

$$\begin{cases} \frac{\partial \mathcal{A}'}{\partial a} \equiv (4\pi a + 2\pi h) - \lambda 2\pi a h = 0 \\ \frac{\partial \mathcal{A}'}{\partial h} \equiv (2\pi a) - \lambda \pi a^2 = 0 \end{cases} \implies \begin{cases} h = 2a \\ \lambda = 2/a \end{cases}$$

qui est le résultat bien connu, « hauteur=diamètre ».

Un autre exemple simple est la maximisation de l'aire \mathcal{A} d'un triangle à périmètre p fixé. On utilise pour cela la relation : $\mathcal{A} = \sqrt{p(p-2a)(p-2b)(p-2c)}/16$ où a, b , et c sont les longueurs des trois cotés. Nous laissons au lecteur le soin de calculer le gradient de $\mathcal{A} - \lambda p$ par rapport aux trois variables indépendantes a, b, c , et d'en déduire, après élimination de λ la condition attendue $a = b = c$.

7.2 Contraintes intégrales pour les fonctionnelles

Lorsqu'on cherche un extremum d'une fonctionnelle $\Phi : \vec{\gamma}(t) \mapsto \Phi(\vec{\gamma}) \in \mathbb{R}$, où $\vec{\gamma}(t)$ est un chemin dans l'espace des configurations, c'est une variation de la forme :

$$\delta\Phi = \int \phi(\vec{\gamma}(t)) \cdot \delta\vec{\gamma}(t) dt$$

qu'il faut considérer. Plus précisément, si on observe que la relation :

$$(f, g) \mapsto \langle f, g \rangle = \int f(\vec{\gamma}(t)) \cdot g(\vec{\gamma}(t)) dt$$

définit un *produit scalaire* sur l'espace de fonctions, l'équivalent du « gradient » de Φ évalué au « point » $\vec{\gamma}$, que l'on pourrait écrire abusivement $\delta\Phi/\delta\vec{\gamma}$, est exactement la fonction $\phi(\vec{\gamma}(t))$.

Dans un certain nombre de cas, les contraintes s'expriment elles-mêmes à l'aide de *fonctionnelles* sous la forme $G_i(\vec{\gamma}) = cste$, où les m fonctionnelles G_i , sont des intégrales du type :

$$G_i : \vec{\gamma}(t) \mapsto G_i(\vec{\gamma}) \in \mathbb{R} \quad \text{avec} \quad G_i(\vec{\gamma}) = \int \mathcal{G}_i(\vec{\gamma}(t), \dot{\vec{\gamma}}(t), t) dt. \quad (21)$$

On rencontrera notamment ce type de contraintes en étudiant des problèmes de mécanique plus élaborés (corde inextensible, courbe brachistochrone), ou des formes d'équilibre de films de savon, lorsqu'on cherche à minimiser l'énergie à périmètre donné²⁵.

Dans le problème d'extrémisation de l'action qui nous intéresse, le problème n'est pas exactement de cette forme. On peut toutefois, dans une version discrétisée, assimiler les valeurs prises par $\vec{\gamma}(t)$ aux instants t successifs aux composantes des vecteurs \vec{u} du paragraphe § 7.1 ci-dessus ; si l'on préfère, on dira que la contrainte porte sur l'objet $\vec{\gamma}$ dans sa globalité. On peut donc appliquer les résultats précédents, et introduire les multiplicateurs de Lagrange λ_i , qui permettent d'écrire (avec les notations abusives introduites alors) :

$$\frac{\delta\mathcal{S}}{\delta\vec{\gamma}} = \sum_{i=1}^m \lambda_i \frac{\delta G_i}{\delta\vec{\gamma}} \quad (22)$$

En termes des équations de Euler-Lagrange, cela donne :

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) - \frac{\partial \mathcal{L}}{\partial q_k} = \sum_i \lambda_i \left[\frac{d}{dt} \left(\frac{\partial \mathcal{G}_i}{\partial \dot{q}_k} \right) - \frac{\partial \mathcal{G}_i}{\partial q_k} \right],$$

et le problème se réduit donc à minimiser la fonctionnelle :

$$\mathcal{S}'(\vec{\gamma}) = \mathcal{S}(\vec{\gamma}) - \sum_i \lambda_i G_i(\vec{\gamma}).$$

7.3 Liaisons holonomes

Des contraintes de type holonome apparaissent naturellement dans le cas où on *ajoute* une (ou plusieurs) liaison(s) supplémentaire(s) après avoir choisi un jeu de coordonnées généralisées, ou bien lorsqu'il est techniquement *plus simple* de conserver des coordonnées non-indépendantes et de prendre en compte la ou les liaison(s) « après coup ». La forme choisie de ces contraintes reste alors de la forme (3) :

$$f_i(\{\vec{r}_\alpha\}, t) = 0 \quad i \in \{1, \dots, m\}, \quad (3)$$

et ne peut pas être mis sous la forme intégrale (21). Il est clair en particulier que ceci est une *contrainte beaucoup plus forte* que la contrainte globale étudiée au paragraphe précédent.

²⁵ On parle de ce fait du « problème iso-périmétrique » ; la contrainte porte alors sur la longueur du chemin, et on aura donc dans ce cas $G(\sigma, \dot{\sigma}, t) = ds/dt = \sqrt{\dot{\sigma}}$.

En reprenant la démonstration des équations de Lagrange (cf. § 4.2), on retrouve naturellement sous l'intégrale la quantité :

$$\sum_{k=1}^n \left(\frac{\partial \mathcal{L}}{\partial q_k} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) \right) \delta q_k , \quad (23)$$

mais il faut ici tenir compte de la dépendance des q_k , donc des δq_k , qui résulte des contraintes. D'après (3), la liaison peut s'écrire :

$$\sum_{k=1}^n \frac{\partial}{\partial q_k} \left(f_i(\{q_j\}, t) \right) \delta q_k = 0 \quad \text{pour } i \in \{1, \dots, m\} ,$$

ce qui traduit que l'action n'est plus extrémale par rapport à *toutes* les variations δq_k possibles, mais seulement par rapport au sous-ensemble de celles qui satisfont les contraintes. Il s'en suit que l'équation (23) n'impose plus aux termes dans les grandes parenthèses d'être nul, mais seulement d'être *perpendiculaires à chaque instant* aux δq_k autorisés. Comme dans le cas précédent, le supplémentaire orthogonal de l'espace des δq_k licites est engendré par les gradients des fonctions de contraintes. On aura donc l'équation similaire à (22) :

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) - \frac{\partial \mathcal{L}}{\partial q_k} = \sum_{i=1}^m \lambda_i(t) \frac{\partial f_i}{\partial q_k} , \quad (24)$$

à la différence essentielle près que les multiplicateurs de Lagrange sont ici des *fonctions du temps*, et non plus des constantes !

On pourrait certes avoir des doutes sur la pertinence d'introduire m fonctions du temps inconnues (et donc de passer de n à $n + m$ fonctions inconnues) pour rendre compte de ce qu'il a seulement $n - m$ « vrais » degrés de liberté. Un intérêt de cette approche est cependant mis en évidence par l'écriture de l'équation ci-dessus sous la forme (cf. équation (13)) :

$$\frac{d}{dt} \left(\frac{\partial K}{\partial \dot{q}_k} \right) - \frac{\partial K}{\partial q_k} = \mathfrak{F}_k + \sum_{i=1}^m \lambda_i(t) \frac{\partial f_i}{\partial q_k} .$$

Nous avons souligné au § 5.3 que les forces généralisées \mathfrak{F}_k ne dépendaient pas des forces de liaison ; il en va tout autrement ici puisque les termes additionnels sont précisément normaux aux surfaces de contraintes. Nous constatons donc que les multiplicateurs de Lagrange, loin d'être de simples artifices mathématiques, sont en fait – à $\|\partial f_i / \partial q_k\|$ près – les *forces de liaison* associées aux contraintes additionnelles. Dans nombre de cas concrets, il pourra être très utile de connaître cette force (par exemple pour réaliser un système suffisamment rigide), obtenue comme un « sous-produit » de la méthode que nous venons d'exposer.

7.4 Liaisons non-holonomes

Nous avons laissé de côté, en commençant cet exposé (§ 3.3), le traitement des liaisons non-holonomes. Il se trouve que leur prise en compte peut maintenant être réalisée à l'aide des mêmes concepts que ceux que nous venons de développer pour les liaisons holonomes. En effet, ces dernières ne sont intervenues ci-dessus que par leur forme différentielle :

$$df_i = \sum_{k=1}^n \frac{\partial f_i}{\partial q_k} dq_k = 0 , \quad (25)$$

mais le fait qu'il s'agisse d'une forme différentielle *exacte* n'a pas réellement joué de rôle.

Roulement sans glissement Les liaisons non-holonomes s'expriment souvent comme une liaisons linéaires entre les vitesses, et c'est notamment le cas pour le roulement sans glissement, que nous décrivons sommairement ici.

Un ensemble de m liaisons linéaires entre les vitesses peut de façon générale s'écrire sous la forme :

$$\sum_{k=1}^n a_{ik}(\{q_j\}, t) \dot{q}_k = 0 \quad \text{soit} \quad d\omega_i = \sum_{k=1}^n a_{ik}(\{q_j\}, t) dq_k = 0 \quad \text{pour } i \in \{1, \dots, m\} .$$

Cette équation est très analogue à (25), et pour y satisfaire, on se limite à nouveau à des variations δq_k qui satisfont à chaque instant :

$$\sum_{k=1}^n a_{ik}(\{q_j\}, t) \delta q_k = 0 \quad \text{pour } i \in \{1, \dots, m\} ,$$

et on a alors à nouveau m multiplicateurs de Lagrange $\lambda_i(t)$ vérifiant une relation identique à (24), au remplacement près des $\partial f_i / \partial q_k$ par les a_{ik} :

$$\frac{d}{dt} \left(\frac{\partial K}{\partial \dot{q}_k} \right) - \frac{\partial K}{\partial q_k} = \mathcal{F}_k + \sum_{i=1}^m \lambda_i(t) a_{ik} .$$

De cette façon les termes $\sum_{i=1}^m \lambda_i(t) a_{ik}(\{q_j\}, t)$ apparaissent dans les nouvelles équations de Lagrange comme des forces additionnelles, découlant de la liaison, holonome ou pas. Le cas qui nous intéresse ici étant celui où la forme différentielle $d\omega$ de (7.4) n'est pas fermée, contrairement au cas holonome. On peut montrer que cela se traduit par le fait que les forces de liaison ne sont généralement plus normales aux déplacements autorisés.

L'évaluation des multiplicateurs de Lagrange permet alors de (i) tester le signe des force de réaction normale pour vérifier que le contact est maintenu, (ii) d'appliquer les lois de Coulomb du frottement pour vérifier que le roulement sans glissement est possible.

8 Applications importantes

8.1 Modes normaux

Nous l'avons dit, l'un des intérêts de l'approche lagrangienne est de présenter un formalisme unique indépendamment du choix des coordonnées, bien que ce choix puisse évidemment influer de façon déterminante sur la facilité de résolution des équations de Lagrange, une fois celles-ci explicitées. Nous allons mettre ces possibilités à profit pour traiter de façon extrêmement générale le problème important des « petites oscillations ».

Position du problème Soit un système physique décrit par ses coordonnées généralisées $\{q_i\}$, et par le lagrangien standard $\mathcal{L} = K(\{q_i\}, \{\dot{q}_i\}) - V(\{q_i\})$, indépendant du temps. Nous supposons que le potentiel V possède un minimum en un point $\{q_i^{(0)}\}$, qui définit donc une position d'équilibre stable, et que nous prenons pour simplifier comme origine des coordonnées. Nous nous intéressons aux mouvements de faible amplitude de ce système autour du point d'équilibre.

L'énergie cinétique K est par hypothèse une forme quadratique *définie positive* des vitesses, dont les coefficients sont évalués au point d'équilibre $\{q_i^{(0)}\}$. De même, on peut faire un développement limité du potentiel autour de $\{q_i^{(0)}\}$, dont le terme d'ordre 1 sera nul par hypothèse. Dans le

cas général on pourra donc représenter V par le terme d'ordre 2, et le lagrangien s'écrira donc :

$$\mathcal{L}(\{q_i\}, \{\dot{q}_i\}) = \frac{1}{2} \sum_{ij} M_{ij} \dot{q}_i \dot{q}_j - \frac{1}{2} \sum_{ij} V_{i,j} q_i q_j ,$$

où la seconde forme quadratique donnée par $V_{i,j} = \partial^2 V / \partial q_i \partial q_j (\{q_i^{(0)}\})$ est positive (mais pas nécessairement définie s'il existe un équilibre indifférent pour certaines coordonnées). Dans ce qui suit, nous adoptons une notation vectorielle (abusives) qui simplifie les écritures, les $\{q_i(t)\}$ étant représentés par le vecteur colonne $\mathbf{q}(t)$, et les formes quadratiques par les matrices symétriques \mathbf{M} et \mathbf{V} .

Cette forme du lagrangien conduit nécessairement à des équations d'Euler-Lagrange qui sont D équations différentielles linéaires couplées du second ordre, que l'on peut écrire symboliquement :

$$\mathbf{M} \ddot{\mathbf{q}}(t) + \mathbf{V} \mathbf{q}(t) = 0 .$$

les solutions de ce système Peprésentent des oscillations de différentes fréquences, caractéristiques du système. De façon générale, on appelle « **modes propres** » du système les solutions oscillantes de la forme :

$$\mathbf{q}_n(t) = \Re e(\mathbf{q}_n(0) \exp[-i\omega_n t]) ,$$

ne comportant qu'une seule fréquence. En reportant cette expression dans la précédente, on obtient l'équation :

$$-\omega_n^2 \mathbf{M} \mathbf{q}_n + \mathbf{V} \mathbf{q}_n = 0 \quad \Rightarrow \quad \det(\mathbf{V} - \omega_n^2 \mathbf{M}) = 0 ,$$

puisque'un système linéaire *homogène* n'a de solution non triviale que si son déterminant est nul. L'équation ainsi obtenue s'appelle « équation séculaire », en raison d'une équation analogue sur l'évolution lente des orbites planétaires en astronomie. Cette équation s'apparente à une recherche de valeurs propres/vecteurs propres de la matrice \mathbf{V} , si ce n'est qu'on aurait dans ce cas la matrice identité $\mathbf{1}$ à la place de \mathbf{M} .

Solution générale La solution la plus générale à ce problème consiste à opérer un changement de coordonnées généralisées $\{q_i\} \mapsto \{Q_i\}$ de telle sorte que le lagrangien puisse s'écrire sous la *forme canonique* :

$$\mathcal{L}(\{Q_i\}, \{\dot{Q}_i\}) = \frac{1}{2} \sum_n \dot{Q}_n^2 - \omega_n^2 Q_n^2 ,$$

dans laquelle chaque coordonnée correspond à un mode propre. Cette transformation correspond à ce qu'on appelle en mathématiques une « diagonalisation simultanée » des deux formes quadratiques, dont un théorème assure qu'elle est possible si l'une au moins est définie positive. Par chance la démonstration est constructive et nous allons l'exposer.

Une première étape consiste à diagonaliser la « matrice des masses » \mathbf{M} , en faisant ensuite un changement d'échelle pour la ramener à l'identité.

Si \mathbf{P} désigne la matrice de passage orthogonale formée des D vecteurs propres de \mathbf{M} , on a $\mathbf{P}^\top \mathbf{M} \mathbf{P} = \mathbf{D}$, où \mathbf{D} est une matrice diagonale dont les éléments sont les valeurs propres de \mathbf{M} , positives par hypothèse. On peut alors écrire $\mathbf{M} = \mathbf{P} \Delta^2 \mathbf{P}^\top$ où $\Delta = \mathbf{D}^{1/2}$, et il vient donc :

$$K = \frac{1}{2} \dot{\mathbf{q}}^\top \mathbf{M} \dot{\mathbf{q}} = \frac{1}{2} \dot{\mathbf{q}}^\top (\mathbf{P} \Delta^2 \mathbf{P}^\top) \dot{\mathbf{q}} = \frac{1}{2} (\dot{\mathbf{q}}^\top \mathbf{P} \Delta) (\Delta \mathbf{P}^\top \dot{\mathbf{q}}) = \frac{1}{2} \dot{\mathbf{Q}}^\top \dot{\mathbf{Q}} .$$

On a ici introduit des coordonnées généralisées intermédiaires $\mathbf{Q} \equiv \Delta \mathbf{P}^\top \mathbf{q}$. En termes de ces coordonnées, l'énergie potentielle s'écrit :

$$V = \frac{1}{2} \mathbf{q}^\top \mathbf{V} \mathbf{q} = \frac{1}{2} (\mathbf{P} \Delta^{-1} \mathbf{R})^\top \mathbf{V} (\mathbf{P} \Delta^{-1} \mathbf{R}) = \frac{1}{2} \mathbf{Q}^\top (\Delta^{-1} \mathbf{P}^\top \mathbf{V} \mathbf{P} \Delta^{-1}) \mathbf{Q} = \frac{1}{2} \mathbf{Q}^\top \mathbf{W} \mathbf{Q} ,$$

où la matrice $\mathbf{W} \equiv \Delta^{-1} \mathbf{P}^\top \mathbf{V} \mathbf{P} \Delta^{-1}$ est toujours une matrice symétrique et positive. On peut finalement la diagonaliser sous la forme : $\mathbf{W} = \mathbf{U} \cdot \omega^2 \mathbf{U}^\top$, où \mathbf{U} est une nouvelle matrice de passage orthogonale, formée des vecteurs propres de \mathbf{W} . En passant enfin aux coordonnées généralisées $\mathbf{R} = \mathbf{U}^\top \mathbf{Q} = \mathbf{U}^\top \Delta \mathbf{P}^\top \mathbf{q}$, on obtient effectivement le résultat annoncé :

$$\mathcal{L} = \frac{1}{2} \dot{\mathbf{R}}^\top \mathbf{U}^\top \mathbf{U} \dot{\mathbf{R}} - \frac{1}{2} \mathbf{R}^\top \mathbf{U}^\top (\mathbf{U} \cdot \omega^2 \mathbf{U}^\top) \mathbf{U} \mathbf{R} = \frac{1}{2} \dot{\mathbf{R}}^\top \dot{\mathbf{R}} - \frac{1}{2} \mathbf{R}^\top \omega^2 \mathbf{R} .$$

Cette démonstration, bien qu'un peu lourde, a le mérite d'être totalement constructive. Dans bien des cas, la « matrice des masses » \mathbf{M} sera déjà diagonale, avec sur la diagonale la masse des particules, et le problème se réduit alors à la diagonalisation de $\mathbf{M}^{-1/2} \mathbf{V} \mathbf{M}^{-1/2}$.

Les résultats physique sont bien sûr les mêmes que ceux obtenus via le déterminant séculaire, puisque les fréquences obtenues sont les valeurs propres de $\mathbf{M}^{-1} \mathbf{V}$ qui sont bien sûr les mêmes que celles de $\mathbf{M}^{-1/2} \mathbf{V} \mathbf{M}^{-1/2}$ et donc de \mathbf{W} , et les vecteurs propres sont bien donnés par la décomposition des $\mathbf{Q}^{(n)}$ sur les \mathbf{q} .

Chaîne linéaire Un exemple physiquement important est celui d'une chaîne linéaire de N « atomes » de masse m , séparés à l'équilibre de la distance a . le nombre e degrés e liberté est lors $D = N$. Nous utilisons comme coordonnées généralisées les *écarts* aux positions d'équilibre, notés x_i , avec $i = 1 \cdots N$, et supposons que la force de rappel vers l'équilibre dépend uniquement de la distance relative aux deux plus proches voisins. Elle est modélisée – dans la limite des petites oscillations – par un ressort de raideur $k = m\Omega^2$ entre chaque masse.

Moyennant ces hypothèses, le lagrangien s'écrit :

$$\mathcal{L} = \sum_{p=1}^N \frac{m}{2} \dot{x}_p^2 - \sum_{p=1}^{N-1} \frac{k}{2} (x_p - x_{p+1})^2 .$$

Dans ce cas, la matrice $\mathbf{M} = m\mathbf{1}$ est déjà diagonale et scalaire, c'est simplement la matrice \mathbf{V} qu'il faut diagonaliser. Cette matrice s'écrit comme indiqué sur la figure 9, où les termes extrêmes de la diagonale sont différents des autres (*i.e.* de -2) en raison des conditions aux limites qui font que les masses 1 et N ne sont soumises qu'au rappel d'un unique ressort.

Conditions au limites périodiques Avant de revenir à ce problème, nous pouvons traiter celui, plus simple, où l'on impose artificiellement des « conditions au limites périodiques » (CLP), en introduisant un atome fictif de position $x_0 \equiv x_N$ et un ressort additionnel entre les masses 0 et 1. Ceci permet de rendre tous les sites atomiques identiques, et de restaurer artificiellement l'invariance par translation. La matrice \mathbf{V}' ainsi obtenue ne comporte plus que des -2 sur sa diagonale, et des termes 1 à la place des 0 dans les coins supérieur droit et inférieur gauche. C'est alors une matrice « circulante », c'est à dire dont les lignes sont obtenues par des permutations circulaires successives. Cette symétrie permet de rechercher²⁶ ses éléments propres en termes de la racine N -ième de l'unité $\eta = \exp(i2\pi/N)$, et la matrice de passage unitaire \mathbf{U} telle $U_{i,j} = \eta^{i \times j}$ (*cf.* figure 9), permet de la diagonaliser.

²⁶ Voir par exemple http://fr.wikipedia.org/wiki/Matrice_circulante

$$\mathbf{V} = \begin{bmatrix} \boxed{-1} & 1 & 0 & \cdots & 0 & \boxed{0} \\ 1 & -2 & 1 & \ddots & & 0 \\ 0 & 1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 & 0 \\ 0 & & \ddots & 1 & -2 & 1 \\ \boxed{0} & 0 & \cdots & 0 & 1 & \boxed{-1} \end{bmatrix} \quad \mathbf{U} = \begin{bmatrix} \eta & \eta^2 & \cdots & \eta^{(N-1)} & 1 \\ \eta^2 & \eta^4 & \cdots & \eta^{2(N-1)} & 1 \\ \eta^3 & \eta^6 & \cdots & \eta^{3(N-1)} & 1 \\ \vdots & \vdots & & \vdots & 1 \\ \eta^{N-1} & \eta^{2(N-1)} & \cdots & \eta^{(N-1)^2} & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix}$$

Fig. 9 – Matrice \mathbf{V} représentant le potentiel, et matrice de passage \mathbf{U} vers la base qui diagonalise la matrice \mathbf{V}' correspondant aux CLP. Les éléments surlignés dans \mathbf{V} sont ceux affectés par les CLP.

On obtient ainsi N modes propres de la forme $\mathbf{q}_i^{(n)} = a^{(n)} \exp(in\phi)$, pour $n = 1 \cdots N$, avec $\phi = 2\pi/N$. Les fréquences propres correspondantes s'obtiennent en reportant cette structure dans l'équation, ce qui conduit à :

$$\omega_n = 2\Omega |\sin(n\phi/2)| \quad \text{pour } n = 1 \cdots N .$$

On doit noter qu'il y a bien exactement N modes propres, puisque des valeurs de n plus grandes donneraient la même dépendance spatiale et la même fréquence que $n \bmod (N)$.

La distance a ne joue ici aucun rôle, mais elle intervient de façon implicite si l'on interprète physiquement les modes propres en termes d'ondes progressives caractérisés par les vecteurs d'ondes $k^{(n)} = n\phi/a$, dont les fréquences sont données par la « relation de dispersion » $\omega(k) = 2\Omega |\sin(ka/2)|$, et la périodicité modulo N restreignant l'intervalle pertinent à $0 \leq k < 2\pi/a$, ou, de façon équivalente à la « première zone de Brillouin » $-\pi/a < k < \pi/a$.

Conditions aux limites physiques Si l'on revient maintenant aux conditions au problème initial, on constate que l'on peut toujours chercher des solutions sous forme d'ondes progressives de la forme $x_p = A \exp(ip\beta) \exp(-i\omega t)$, et que l'équation d'Euler-Lagrange relative à \ddot{x}_p pour $p \in \{2, \cdots, N-1\}$ donne :

$$-\omega_n^2 x_p = -\Omega^2 (2x_p - x_{p-1} - x_{p+1}) \quad \Rightarrow \quad \omega(\beta) = 2\Omega |\sin(\beta/2)| \quad \text{pour } \beta \in]-\pi, \pi]$$

On obtient naturellement la même relation de dispersion que précédemment, mais la valeur de la phase β reste à déterminer, et ce sont précisément les conditions aux limites qui vont en fixer les valeurs possibles. Si nous considérons les équations satisfaites par x_1 et x_N :

$$\ddot{x}_1 = -\Omega^2 (x_1 - x_2) \quad \text{et} \quad \ddot{x}_N = -\Omega^2 (x_N - x_{N-1}) ,$$

il est visible que les ondes progressives précédentes ne sauraient les satisfaire. On peut toutefois chercher des solutions mêlant des ondes progressives dans un sens et dans l'autre, qui possèdent la même fréquence. Physiquement cela signifie que les conditions aux bords induisent des réflexions des ondes progressives « incidentes ». On cherche alors légitimement des solutions du type :

$$x_p = (A \exp(ip\beta) + B \exp(-ip\beta)) \exp(-i\omega t) .$$

En utilisant les propriétés de symétrie qui imposent $B = A^*$, en reportant dans l'équation sur \ddot{x}_1 , et en utilisant la relation de dispersion $\omega(\beta)$ on obtient aisément le résultat $\arg(A) = -\arg(B) = -\beta/2$. On a donc simplement une onde stationnaire qui s'écrit :

$$x_p = 2|A| \cos\left(\left(p - \frac{1}{2}\right)\beta\right) \cos(\omega t) , \quad (26)$$

et ne possède *a priori* ni un nœud ni un ventre en $p = 1$. La prise en compte de la condition aux limites pour $p = N$ conduit maintenant, après un peu de trigonométrie, à la condition $\sin(N\beta) = 0$, ce qui fixe la valeur de β pour les N modes propres :

$$\beta^{(n)} = \frac{n\pi}{N} \quad \text{avec} \quad n = 0, 1, \dots, N-1,$$

et les fréquences propres correspondantes s'en déduisent à l'aide de la relation de dispersion. Il convient de noter que le mode $n = 0$ correspond à une fréquence d'appel ω_0 nulle, ce qui décrit une translation uniforme de l'ensemble des atomes. En outre, pour $N \gg n$ la phase aux extrémités est très faible, ce qui signifie que l'on a quasiment un ventre (et donc une réflexion sans changement de signe), alors que pour $N \gg N - n$, la phase est de l'ordre de $\pi/2$, ce qui fait que l'on a quasiment un nœud (réflexion avec changement de signe), comme si les extrémités avaient été fixées.

Les « variables normales » R_n sont alors définies par des combinaisons linéaires des x_p avec des poids qui sont distribués comme dans l'onde stationnaire de l'équation (26).

8.2 Variables continues ★

Passage à la limite continue Dans le cas précédent, si le nombre N devient infiniment grand et la distance caractéristique a infiniment petite, avec une longueur $L = Na$ fixée, il peut être intéressant de remplacer l'indice discret p par un indice continu x , conçu comme limite de la fonction $x(p) = p \times a$. Dans cette approche, les coordonnées généralisées $\{x_p(t)\}$ cèdent la place à une fonction $\varphi(x, t)$, et les vitesses généralisées aux dérivées temporelles évaluées à position fixée, soit $\dot{x}_p \rightarrow \partial\varphi(x, t)/\partial t$. De même, la somme discrète sur p devient une intégrale sur x , en utilisant $dx = a$. De façon moins évidente, les différences de positions mises en jeu dans l'énergie potentielle élastique peuvent être transformées par $(x_p - x_{p-1}) \rightarrow a \times \partial V/\partial x$. Ceci conduit à réécrire le lagrangien :

$$\mathcal{L} = \frac{1}{2} \int_0^L m \left(\frac{\partial\varphi}{\partial t} \right)^2 \frac{dx}{a} - \frac{1}{2} \int_0^L k \left(a \frac{\partial\varphi}{\partial x} \right)^2 \frac{dx}{a} \equiv \frac{1}{2} \int_0^L \left[\mu \left(\frac{\partial\varphi}{\partial t} \right)^2 - Y \left(\frac{\partial\varphi}{\partial x} \right)^2 \right] dx$$

en introduisant la masse linéique $\mu = m/a$, et le module de compression $Y = \lim_{N \rightarrow \infty} ka$, lequel a un sens physique bien défini dans la mesure où la raideur d'un ressort est inversement proportionnelle à sa longueur.

Description générale La « théorie classique des champs » repose sur une généralisation de cette expression, dont nous allons esquisser les grands traits pour un champ scalaire $\varphi(\vec{r}, t)$ sur l'espace usuel à trois dimensions :

- on définit une densité volumique de lagrangien $L(\varphi, \partial_t\varphi, \partial_{\vec{r}}\varphi, t)$ dépendant de façon générale du champ, de ses dérivées spatiales $\partial_{\vec{r}}\varphi \equiv \partial\varphi/\partial\vec{r} \equiv \nabla\varphi$ et temporelle $\partial_t\varphi \equiv \partial\varphi/\partial t$ et éventuellement du temps ;
- l'action est définie par l'intégrale *spatio-temporelle* :

$$\mathcal{S} = \int_{t_1}^{t_2} \left[\int_{\mathcal{V}} L(\varphi, \partial_t\varphi, \partial_{\vec{r}}\varphi, t) d^3\vec{r} \right] dt,$$

avec des états initial $\varphi(\vec{r}, t_1)$ et final $\varphi(\vec{r}, t_2)$ fixés dans tout le volume d'intégration \mathcal{V} ainsi que des conditions aux limites spatiales fixées sur les bords $\partial\mathcal{V}$ du volume, ou pour $\|\vec{r}\| \rightarrow \infty$ si \mathcal{V} est infini, et ce à chaque instant ;

– on applique le principe de Hamilton en cherchant les conditions qui rendent \mathcal{S} stationnaire vis à vis des variations du champ $\delta\varphi$.

On peut alors écrire la variation de l'action :

$$\delta\mathcal{S} = \int_{t_1}^{t_2} \int_{\mathcal{V}} \left\{ \frac{\partial L}{\partial\varphi} \delta\phi + \frac{\partial L}{\partial t\varphi} \delta(\partial_t\varphi) + \frac{\partial L}{\partial(\nabla\varphi)} \cdot \delta(\nabla\varphi) \right\} d^3\vec{r} dt .$$

La désormais traditionnelle intégration par parties temporelle donne :

$$\delta\mathcal{S} = \int_{t_1}^{t_2} \int_{\mathcal{V}} \left\{ \frac{\partial L}{\partial\varphi} \delta\phi - \frac{\partial}{\partial t} \left(\frac{\partial L}{\partial t\varphi} \right) \delta\varphi + \frac{\partial L}{\partial(\nabla\varphi)} \delta(\nabla\varphi) \right\} d^3\vec{r} dt + \left[\int_{\mathcal{V}} \frac{\partial L}{\partial t\varphi} \delta\varphi d^3\vec{r} \right]_{t_1}^{t_2}$$

dont le terme intégré est nul puisque par hypothèse $\delta\varphi(\vec{r}, t_1) = \delta\varphi(\vec{r}, t_2) = 0$ dans \mathcal{V} entier.

La dernière étape consiste dans une intégration par parties spatiale fondée sur le théorème de Green-Ostrogradski et sur l'identité :

$$\operatorname{div}(f \vec{u}) = \nabla f \cdot \vec{u} + f \operatorname{div} \vec{u}$$

où f est identifié à $\delta\varphi$ et \vec{u} à $\partial L/\partial(\nabla\varphi)$. Cela donne :

$$\int_{\mathcal{V}} \left\{ \frac{\partial L}{\partial(\nabla\varphi)} \cdot \nabla\delta\varphi + \operatorname{div}\left(\frac{\partial L}{\partial(\nabla\varphi)}\right) \delta\varphi \right\} d^3\vec{r} = \int_{\mathcal{V}} \operatorname{div}\left(\frac{\partial L}{\partial(\nabla\varphi)} \delta\varphi\right) d^3\vec{r} = \int_{\partial\mathcal{V}} \frac{\partial L}{\partial(\nabla\varphi)} \delta\varphi \cdot d\vec{S} .$$

Cette quantité est le terme « tout intégré », nul à chaque instant par hypothèse. En reportant dans $\delta\mathcal{S}$ il vient :

$$\delta\mathcal{S} = \int_{t_1}^{t_2} \int_{\mathcal{V}} \left\{ \frac{\partial L}{\partial\varphi} - \frac{\partial}{\partial t} \left(\frac{\partial L}{\partial t\varphi} \right) - \operatorname{div}\left(\frac{\partial L}{\partial(\nabla\varphi)}\right) \right\} \delta\varphi d^3\vec{r}$$

qui ne peut s'annuler pour $\delta\varphi$ quelconque que si le terme entre $\{\cdot\}$ est nul en tout point et à chaque instant, d'où l'équation au dérivées partielles de Euler-Lagrange :

$$\frac{\partial}{\partial t} \left(\frac{\partial L}{\partial t\varphi} \right) + \operatorname{div}\left(\frac{\partial L}{\partial(\nabla\varphi)}\right) - \frac{\partial L}{\partial\varphi} = 0 ,$$

Exemples Une forme simple de densité de lagrangien, possédant un maximum de symétries, est donnée par :

$$L = \frac{1}{2} \left[\frac{1}{v^2} \left(\frac{\partial\varphi}{\partial t} \right)^2 - \left(\frac{\partial\varphi}{\partial\vec{r}} \right)^2 - \frac{\varphi^2}{\Lambda^2} \right] ,$$

où c est une vitesse, et Λ une longueur. On calcule aisément $\partial L/\partial t\varphi = (\partial_t\varphi)/v^2$, $\partial L/\partial(\nabla\varphi) = -\nabla\varphi$, et $\frac{\partial L}{\partial\varphi} = \varphi/\Lambda^2$, d'où l'on déduit l'équation d'onde :

$$\frac{1}{v^2} \frac{\partial^2\varphi}{\partial t^2} - \Delta\varphi + \frac{1}{\Lambda^2} \varphi = 0 .$$

Dans le cas où $\Lambda^{-2} = 0$, on reconnaît une équation de d'Alembert usuelle, décrivant la propagation d'ondes analogues à celles de l'électromagnétisme, si ce n'est que l'on n'a considéré pour simplifier qu'un champ scalaire.

Notons que la version à une dimension que nous obtenions à la limite continue de la chaîne linéaire donne une onde de compression (onde sonore) de vitesse $v = \sqrt{K/\mu} = \Omega/a$, qui est bien la vitesse de phase des modes propres de chaîne d'oscillateurs à la limite où $k \ll 1/a$.

Dans le cas où $\Lambda^{-2} \neq 0$ et où c est la vitesse de la lumière, on obtient l'équation dite « de Klein-Gordon », dont les solutions harmoniques vérifient la relation de dispersion :

$$\frac{\omega^2}{c^2} = k^2 + \frac{1}{\Lambda^2} .$$

Les solutions *stationnaires* à symétrie *sphérique* ont alors la forme d'un potentiel de Yukawa :

$$\varphi(\vec{r}, t) = A \frac{e^{-r/\Lambda}}{r} .$$

Celui-ci est utilisé pour décrire de façon phénoménologique l'interaction forte dans les noyaux atomiques. La portée Λ du potentiel est reliée à la masse de la particule « vecteur » de l'interaction (pour l'interaction entre nucléons, il s'agit du pion π de masse environ $140 \text{ Mev}/c^2 \simeq 0.15 \text{ uma}$) par la relation $\Lambda = \lambda_c = \hbar/mc$ (longueur d'onde de Compton), ici de l'ordre de 10^{-15} m . Dans ce sens, l'électromagnétisme correspond une interaction dont le vecteur (le photon) est de masse nulle et de portée infinie ($\Lambda = \infty$).

Tous ces concepts seront approfondis, dans un cadre plus général, dans le cours de Relativité et Électromagnétisme.

9 L'action comme fonction des coordonnées

Au paragraphe § 4, nous avons défini l'action comme une *fonctionnelle*, qui attribuait une valeur réelle à un chemin (d'extrémités fixées) et nous avons utilisé le principe de Hamilton pour définir le « chemin physique » ou trajectoire.

Une fois cette trajectoire déterminée, on peut considérer l'action comme une *fonction* des positions initiale et finale, l'intégrale étant prise sur la trajectoire physique. Cette fonction action est appelée « fonction principale de Hamilton », et joue un rôle important dans le formalisme de Hamilton et de Hamilton–Jacobi. Nous la noterons S (et non plus \mathcal{S}) pour distinguer les deux versions.

Nous nous limiterons ici à la dépendance vis-à-vis de la position finale et de l'instant final. A cet effet, nous considérons deux trajectoires *physiques* arbitrairement proches $\{q_i^{(1)}(t)\}$ et $\{q_i^{(2)}(t)\}$, et la variation $\delta q_i(t) = q_i^{(2)}(t) - q_i^{(1)}(t)$. Nous supposons en outre que les positions initiales à l'instant t_0 sont identiques, c'est-à-dire $\delta q_i(t_0) = 0$.

Dépendance spatiale Supposons dans un premier temps que les instants finaux sont identiques, soit $t_1 = t_2$, et $q^{(1)}(t_2) = q^{(2)}(t_2)$ soit $\delta q(t_0) = 0$. La variation de l'action qui en résulte est comme à l'équation (6) :

$$\delta \mathcal{S} = \int_{t_0}^{t_1} \sum_i \left[\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \right] \delta q_i dt + \sum_i \left[\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta q_i \right]_{t_0}^{t_1} .$$

Au contraire de la démonstration de (6), le terme intégral est identiquement nul, puisque par hypothèse $q^{(1)}(t)$ et $q^{(2)}(t)$ sont des trajectoires physiques, vérifiant les équations d'Euler–Lagrange. Il reste donc le terme intégral, qui lui même ne contient que la valeur en t_1 puisque $\delta q(t_0) = 0$. On peut donc en déduire $\delta \mathcal{S} = \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \delta q_i$.

Comme le second membre ne dépend que de la variation de la position entre les deux points finaux $dq_i = \delta q_i(t_1)$, ceci constitue une expression simple de l'accroissement de la fonction S

entre ces deux points, on en déduit :

$$\frac{\partial S}{\partial q_i}(\{q_j\}, t) = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = p_i(\{q_j\}, t) .$$

Il est à noter que cette dérivée partielle est implicitement prise à t constant. En outre, comme on a considéré des trajectoires physiques, les vitesses \dot{q}_i et les moments p_i sont bien seulement des fonctions des q_i et éventuellement du temps, et non des variables indépendantes.

Dépendance temporelle Pour obtenir la dépendance temporelle de S , on pourrait faire un calcul de variation analogue, en relâchant la contrainte $t_1 = t_2$, et supposant les positions finales identiques, c'est-à-dire $q_i^{(1)}(t_1) = q_i^{(2)}(t_2)$.

Il est toutefois plus direct d'utiliser la forme différentielle de (5) : $dS/dt = \mathcal{L}$ et l'expression (16) de l'énergie, pour obtenir :

$$\frac{\partial S}{\partial t} = \frac{dS}{dt} - \sum_i \frac{\partial S}{\partial q_i} \dot{q}_i = \mathcal{L} - \sum_i p_i \dot{q}_i = -\mathcal{E} .$$

Différentielle de S En définitive, nous obtenons la différentielle de l'action S sous la forme :

$$dS = \sum_i p_i dq_i - \mathcal{E} dt . \quad (27)$$

qui nous sera utile pour la suite²⁷ aux paragraphes § 11 et suivants.

²⁷ Une question, cruciale, de savoir si cette différentielle définit bien une fonction (c'est à dire si elle est fermée) est techniquement très difficile dans le formalisme lagrangien, mais deviendra évidente en formalisme hamiltonien

C – Formulation hamiltonienne

10 Équations de Hamilton

Nous avons vu dans la section 7 que certaines contraintes peuvent être prises en compte directement en modifiant la fonctionnelle à minimiser, au prix d'une augmentation du nombre de variables indépendantes, à hauteur du nombre de contraintes. Or, dans tout ce que nous avons fait jusqu'ici, nous avons dû prendre en compte des contraintes subtiles liées à ce que les vitesses $\{\dot{q}_i(t)\}$ sont des variables indépendantes pour le lagrangien, mais doivent tout de même découler, par dérivation, de la loi horaire $\{q_i(t)\}$, que l'on ait ou non affaire à la trajectoire physique. Le premier objectif de la formulation hamiltonienne est précisément de résoudre ce problème, et nous verrons ensuite ce que cela ouvre de nombreuses perspectives, qui rendent ce point de vue bien plus fécond.

10.1 Point de vue de Hamilton

Dans cette nouvelle approche, pour décrire un système à D degrés de liberté, on introduit, à coté des coordonnées généralisées $\{q_i\}$, et à la place des vitesses généralisées $\{\dot{q}_i\}$, D variables indépendantes supplémentaires, notées $\{p_i\}$ et appelées « impulsions généralisées » ou « moments (canoniquement) conjugués ». La description d'un système physique n'est plus alors faite dans l'espace des configurations, mais dans un nouvel espace à $2D$ dimensions, appelé « **espace des phases** ». On se dote ensuite une fonction de Hamilton, ou hamiltonien, $\mathcal{H}(\{q_i\}, \{p_i\}, t)$ qui doit décrire la dynamique à l'aide du même principe variationnel. Mais comment construire ce hamiltonien de façon à retrouver la même dynamique que celle décrite par les équations de Lagrange ?

Lien entre les points de vue de Lagrange et de Hamilton

On peut former le hamiltonien en effectuant une transformation de Legendre du lagrangien, en posant (le signe $-$ global arbitraire est introduit par commodité) :

$$\mathcal{H}(\{q_i\}, \{p_i\}, t) = -\left(\mathcal{L}(\{q_i\}, \{\dot{q}_i\}, t) - \sum_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j} \dot{q}_j\right) = \sum_j p_j \dot{q}_j - \mathcal{L}(\{q_i\}, \{\dot{q}_i\}, t) \quad (28)$$

où les $\{\dot{q}_i\}$ de la deuxième expression ne doivent plus être vus comme les dérivées des $\{q_i\}$, mais comme des fonction des variables indépendantes $\{q_i\}$ et $\{p_i\}$ (et éventuellement du temps) résultant de l'inversion de l'identité $p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$.

En notant utilisant les équations de Lagrange, puis (28), on a :

$$\begin{aligned} d\mathcal{L} &= \sum_i \dot{p}_i dq_i + \sum_i p_i d\dot{q}_i + \frac{\partial \mathcal{L}}{\partial t} dt \\ \text{et donc } d\mathcal{H} &= \sum_i \dot{q}_i dp_i - \sum_i \dot{p}_i dq_i + \frac{\partial \mathcal{L}}{\partial t} dt, \end{aligned} \quad (29)$$

ce qui nous servira ultérieurement.

Si on exprime maintenant l'action en fonction des nouvelles variables, le principe de Hamilton du § 4.1 doit nous conduire à des équations du mouvement qui soient équivalentes aux équations de Lagrange. On pourra alors « oublier » tout le formalisme lagrangien, et utiliser le formalisme hamiltonien de façon totalement autonome.

10.2 Équations de Hamilton

En inversant l'expression (28) pour obtenir \mathcal{L} en fonction de \mathcal{H} , on peut alors écrire fonctionnelle action, en termes des nouvelles variables :

$$\mathcal{S}(\gamma) = \int_{t_1}^{t_2} \sum_i p_i d\dot{q}_i - \mathcal{H}(\{q_i\}, \{p_i\}, t) dt = \int_{t_1}^{t_2} \left[\sum_i p_i \dot{q}_i - \mathcal{H}(\{q_i\}, \{p_i\}, t) \right] dt .$$

où $\gamma(t)$ est un chemin dans l'espace des phases correspondant à un choix arbitraire de fonctions $q_i(t)$ et $p_i(t)$, et on cherche les conditions qui rendent l'action stationnaire, vis à vis d'une variation $\delta\gamma$ du chemin, avec des extrémités entièrement fixées dans l'espace des phases, c'est-à-dire $\delta q_i(t_1) = \delta q_i(t_2) = 0$ et $\delta p_i(t_1) = \delta p_i(t_2) = 0$.

On écrit alors la variation :

$$\delta\mathcal{S} = \sum_i \int_{t_1}^{t_2} \left[p_i \delta\dot{q}_i + \dot{q}_i \delta p_i - \frac{\partial\mathcal{H}}{\partial q_i} \delta q_i - \frac{\partial\mathcal{H}}{\partial p_i} \delta p_i \right] dt$$

dans laquelle on voit apparaître $\delta\dot{q}_i$, que l'on élimine comme à l'accoutumée par une intégration par parties, d'où, en regroupant les termes :

$$\delta\mathcal{S} = \left[\sum_i p_i \delta q_i \right]_{t_1}^{t_2} + \sum_i \int_{t_1}^{t_2} \left[\left(-\frac{dp_i}{dt} - \frac{\partial\mathcal{H}}{\partial q_i} \right) \delta q_i + \left(\frac{dq_i}{dt} - \frac{\partial\mathcal{H}}{\partial p_i} \right) \delta p_i \right] dt .$$

Comme le premier terme est nul par hypothèse, et que les q_i et p_i sont indépendants, la condition $\delta\mathcal{S} = 0$ entraîne l'annulation des $2D$ termes entre parenthèses, d'où les $2D$ « équations de Hamilton » :

$$\boxed{\begin{cases} \frac{dq_i}{dt} = \frac{\partial\mathcal{H}}{\partial p_i} \\ \frac{dp_i}{dt} = -\frac{\partial\mathcal{H}}{\partial q_i} \end{cases}} \quad (30a)$$

Comme on peut le constater, les D équations de Lagrange, du second ordre, ont laissé place à $2D$ équations du premier ordre, ce qui est mathématiquement équivalent, mais souvent plus pratique, qu'il s'agisse d'en faire une résolution formelle ou numérique.

10.3 Lien avec le point de vue de Lagrange

Il est essentiel de faire les observations suivantes :

NB1 Le formalisme hamiltonien est totalement indépendant du formalisme lagrangien : en particulier la première équation de Hamilton (30a) n'est pas une ré-écriture de l'impulsion « à la Lagrange », mais bel et bien l'une des équations du mouvement. En particulier, lorsqu'on varie le chemin γ , on envisage aussi les chemins où cette équation n'est pas vérifiée.

NB2 Les équations de Hamilton *ne sont pas* non plus une conséquence triviale de l'équation (29), mais résultent uniquement du principe de moindre action (« de Hamilton »), écrit sans référence à la formulation de Lagrange. En revanche cette équation prouve que *si* la fonction de Hamilton est déduite d'une fonction de Lagrange par la transformation de Legendre (28), *alors* les équations du mouvement obtenues dans l'un et l'autre point de vue seront équivalentes.

NB3 Nous appliquons le principe de moindre action en fixant les impulsions aux extrémités, mais le processus conduisant aux équations de Hamilton n'utilise en fait que la nullité des δq_i aux deux extrémités, et non celle des δp_i : malgré les apparences, nous n'avons donc pas utilisé des hypothèses plus fortes.

Cela dit, le lien entre les deux formulations est très utile pour les deux raisons suivantes :

- il permet de comprendre comment et pourquoi « ça marche » ;
- il fournit une méthode pour construire de façon systématique le hamiltonien physiquement convenable, par transformation de Legendre, ce qui implique impérativement l'élimination des \dot{q}_i au profit des p_i

10.4 Lien avec l'énergie

Dans le cas où le système est « autonome », c'est à dire qu'il ne subit pas d'interaction avec un environnement dépendant du temps, son lagrangien n'a pas de dépendance explicite en temps, et nous avons vu que l'énergie mécanique est conservée (cf. § 6.2). Il est important de noter que la valeur numérique du hamiltonien est alors constante, égale à l'énergie mécanique \mathcal{E} du système. Toutefois, alors que l'énergie peut être écrite indifféremment de bien des façons, la fonction de Hamilton, elle, doit impérativement être écrite en termes des coordonnées q_i et de leurs moments conjugués p_i . Une expression du hamiltonien contenant explicitement des \dot{q}_i est incorrecte et peut conduire à de résultats incohérents

Dans le cas, où \mathcal{L} dépend du temps, comme par exemple en présence de champs extérieurs variables, on sait que l'énergie n'est plus conservée, mais il reste toujours possible de définir un hamiltonien par la même procédure, celle-ci ayant du coup une dépendance explicite en temps. Cette dépendance s'exprime simplement :

$$\frac{\partial \mathcal{H}}{\partial t} = - \frac{\partial \mathcal{L}}{\partial t} , \quad (31)$$

mais il ne faut pas perdre de vue que la dérivation de \mathcal{H} et de \mathcal{L} sont faites en gardant constants des jeux de variables différents.

De façon complémentaire, dans l'optique de l'indépendance conceptuelle de l'approche de Hamilton vis-à-vis de celle de Lagrange, il convient de s'affranchir de (31), au profit d'une démonstration fondée sur les seules équations de Hamilton :

$$\frac{d\mathcal{H}}{dt} = \sum_i \left(\frac{\partial \mathcal{H}}{\partial q_i} \dot{q}_i + \frac{\partial \mathcal{H}}{\partial p_i} \dot{p}_i \right) + \frac{\partial \mathcal{H}}{\partial t} = \frac{\partial \mathcal{H}}{\partial t} = \sum_i \underbrace{\left(\frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right)}_0 + \frac{\partial \mathcal{H}}{\partial t} .$$

En conséquence un système décrit par un hamiltonien \mathcal{H} admet l'énergie comme intégrale première si et seulement si \mathcal{H} ne dépend pas explicitement du temps.

10.5 Particule chargée dans un champ électromagnétique

Pour illustrer le propos des trois paragraphes précédents, il est utile de reprendre le problème du § 5.8, pour voir comment il se présente dans l'approche de Hamilton. Nous nous placerons pour ce faire en coordonnées cartésiennes.

Construction du hamiltonien Il faut partir du lagrangien impliquant le potentiel généralisé (15). On peut alors écrire :

$$\vec{p} = \frac{\partial \mathcal{L}}{\partial \vec{v}} = m \vec{v} + Q \vec{A}(\vec{r}, t) \quad \text{d'où} \quad \vec{p} \cdot \vec{v} = m v^2 + Q \vec{A}(\vec{r}, t) \cdot \vec{v} ,$$

si bien que le terme de couplage en $\vec{A} \cdot \vec{v}$ s'élimine quand on forme le hamiltonien. Il se manifeste néanmoins par le biais du terme d'énergie cinétique, et on obtient²⁸ :

$$\mathcal{H}(\vec{r}, \vec{p}, t) = \frac{(\vec{p} - Q \vec{A}(\vec{r}, t))^2}{2m} + Q \Phi(\vec{r}, t) .$$

Expression et vérification des équations de Hamilton Les équations de Hamilton s'écrivent alors :

$$\begin{aligned} \frac{d\vec{r}}{dt} &= \frac{\partial \mathcal{H}}{\partial \vec{p}} = \frac{\vec{p} - Q \vec{A}(\vec{r}, t)}{m} \\ \frac{d\vec{p}}{dt} &= - \frac{\partial \mathcal{H}}{\partial \vec{r}} = - \overrightarrow{\text{grad}} \frac{(\vec{p} - Q \vec{A}(\vec{r}, t))^2}{2m} - Q \overrightarrow{\text{grad}} \Phi(\vec{r}, t) . \end{aligned}$$

On peut développer le gradient en utilisant la formule d'analyse vectorielle :

$$\overrightarrow{\text{grad}}(\vec{a} \cdot \vec{b}) = (\vec{a} \cdot \overrightarrow{\text{grad}}) \vec{b} + \vec{a} \times \text{rot} \vec{b} + (\vec{b} \cdot \overrightarrow{\text{grad}}) \vec{a} + \vec{b} \times \text{rot} \vec{a} ,$$

ce qui, en remplaçant \vec{v} par $(\vec{p} - Q \vec{A})/m$, donne :

$$\frac{d\vec{p}}{dt} = Q \left(\frac{\vec{p} - Q \vec{A}(\vec{r}, t)}{m} \cdot \overrightarrow{\text{grad}} \right) \vec{A}(\vec{r}, t) + Q \frac{(\vec{p} - Q \vec{A}(\vec{r}, t))}{m} \times \vec{B} - Q \overrightarrow{\text{grad}} \Phi(\vec{r}, t)$$

et en utilisant la première équation de Hamilton :

$$\frac{d}{dt} (m \vec{v} + Q \vec{A}(\vec{r}, t)) = Q (\vec{v} \cdot \overrightarrow{\text{grad}}) \vec{A}(\vec{r}, t) + Q \vec{v} \times \vec{B}(\vec{r}, t) - Q \overrightarrow{\text{grad}} \Phi(\vec{r}, t) .$$

Si l'on note que $d\vec{A}/dt = \partial \vec{A}/\partial t + (\vec{v} \cdot \overrightarrow{\text{grad}}) \vec{A}$, on constate que le terme en $(\vec{v} \cdot \overrightarrow{\text{grad}})$ s'élimine, et on obtient :

$$\frac{d}{dt} m \vec{v} = Q \vec{v} \times \vec{B}(\vec{r}, t) - Q \overrightarrow{\text{grad}} \Phi(\vec{r}, t) - Q \frac{\partial \vec{A}}{\partial t}(\vec{r}, t) ,$$

qui est bien le résultat attendu.

10.6 Propriétés du hamiltonien

Évolution temporelle Un calcul élémentaire donne :

$$\frac{d\mathcal{H}}{dt} = \sum_i \frac{\partial \mathcal{H}}{\partial q_i} \dot{q}_i + \frac{\partial \mathcal{H}}{\partial p_i} \dot{p}_i + \frac{\partial \mathcal{H}}{\partial t} ,$$

En utilisant les équation de Hamilton, ainsi que la définition de \mathcal{H} , ainsi que (31) il vient :

$$\frac{d\mathcal{H}}{dt} = \sum_i \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} - \frac{\partial \mathcal{L}}{\partial t} = - \frac{\partial \mathcal{L}}{\partial t} ,$$

et on en déduit que pour tout système autonome, le hamiltonien est une intégrale première du mouvement. En d'autres termes, la seule variation possible vient de la dépendance explicite de \mathcal{H} avec le temps.

²⁸ Notons bien que si l'énergie mécanique (non conservée en général) ne dépend pas du champ magnétique, le hamiltonien, lui le fait encore intervenir !

Propriétés mathématiques Le hamiltonien possède les mêmes propriétés mathématiques que le lagrangien dont il est dérivé : invariance d'échelle et d'origine (unité d'énergie), propriétés de symétrie éventuelles, additivité.

En ce qui concerne la non-unicité, on constate aisément que l'on peut, sans modifier les équations du mouvement, ajouter librement à \mathcal{H} la dérivée temporelle (totale) d'une quelconque fonction du temps, des positions *et des impulsions*²⁹. Cela se déduit de façon élémentaire de l'expression de l'action et de la dérivation des équations de Hamilton. En revanche les équations de Hamilton, elles, seront modifiées d'une façon non-triviale. Cette propriété sera très utile dans ce qui suit.

Cette même propriété conduit aussi au résultat très important que certains problèmes peuvent très bien être décrits par un hamiltonien alors qu'ils ne peuvent pas se mettre sous forme lagrangienne, ou encore que certains systèmes peuvent avoir un hamiltonien dont la valeur *numérique* est identiquement nulle pour toute trajectoire !

10.7 Crochets de Poisson

Les équations de Hamilton, grâce au traitement presque symétrique des positions et impulsions, permettent d'écrire de façon très élégante la variation temporelle d'une grandeur physique quelconque. Soit $f(p, q, t)$ une telle grandeur, on a :

$$\frac{df}{dt} = \sum_i \frac{\partial f}{\partial q_i} \dot{q}_i + \frac{\partial f}{\partial p_i} \dot{p}_i + \frac{\partial f}{\partial t} = \sum_i \frac{\partial f}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} + \frac{\partial f}{\partial t} .$$

Il est usuel de réécrire cette équation sous une forme plus compacte en définissant les *crochets de Poisson* de deux grandeurs physiques f et g quelconques par l'identité :

$$\{f, g\} = \sum_k \frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q_k} ,$$

ce qui donne la loi générale :

$$\frac{df}{dt} = \{f, \mathcal{H}\} + \frac{\partial f}{\partial t} . \quad (32)$$

Bien sûr que cela est aussi vrai pour les variables dynamiques fondamentales q_i et p_i , pour lesquelles les équations de Hamilton (30) s'écrivent alors sous la forme encore plus symétrique :

$$\frac{dq_i}{dt} = \{q_i, \mathcal{H}\} \quad \text{et} \quad \frac{dp_i}{dt} = \{p_i, \mathcal{H}\} ,$$

puisque, dans chaque cas, seul le terme pour lequel $k = i$ sera non nul et que par hypothèse $\partial p_i / \partial q_k = \partial q_i / \partial p_k = 0$.

Propriétés élémentaires Les crochets de Poisson jouent un rôle essentiel dans tout ce qui va suivre, il est donc nécessaire d'en dégager les propriétés essentielles :

- Il s'agit tout d'abord d'une forme bilinéaire, de même que le produit scalaire usuel.
- Au contraire du produit scalaire, il est antisymétrique : $\{f, g\} = -\{g, f\}$ (et donc $\{f, f\} = 0$).

²⁹ Contrairement au Lagrangien (cf. point 2 du § 4.3), pour lequel la fonction F pouvait dépendre des positions, mais pas dépendre des vitesses.

– Comme tout produit antisymétrique, il vérifie l'égalité de Jacobi :

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0 ,$$

qui peut elle-même être vue comme un caractère « associatif » du « produit » de Poisson.

– On a bien sûr, en raison des règles de dérivation du produit de deux fonctions, l'identité :

$$\{f, gh\} = \{f, g\}h + \{f, h\}g$$

Illustration : moment cinétique dans un champ magnétique On peut montrer qu'un objet muni un moment cinétique \vec{J} possède un moment magnétique $\vec{\mu} = \gamma \vec{J}$, où la quantité γ est appelée « facteur gyromagnétique »³⁰.

Avec l'expression bien connue de l'énergie d'interaction entre le dipôle et le champ magnétique $U = -\vec{\mu} \cdot \vec{B}$, on obtient le hamiltonien :

$$\mathcal{H} = \vec{\Omega} \cdot \vec{J} \quad \text{où} \quad \vec{\Omega} = -\gamma \vec{B} . \quad (33)$$

Il convient de noter qu'il s'agit d'un hamiltonien *non-standard*, c'est-à-dire pas de la forme $\mathcal{H} = K + V$, et même sans équivalent Lagrangien (classique) en termes de $\{q_i\}$ et de $\{p_i\}$.

Toutefois, on peut utiliser, au lieu des équation de Hamilton, la loi générale (32) et — en anticipant sur les développements de § 11.2 — les crochets de Poisson des composantes de \vec{J} (caractéristiques du groupe SO_3 des rotations de l'espace) :

$$\{J_i, J_j\} = \sum_k \varepsilon_{ijk} J_k$$

où ε_{ijk} est le tenseur totalement antisymétrique de Levi-Civita. On a alors :

$$\frac{dJ_i}{dt} = \left\{ J_i, \sum_j J_j \Omega_j \right\} + \frac{\partial \vec{J}}{\partial t} = \sum_j \varepsilon_{ijk} \Omega_j J_k \quad \text{soit} \quad \frac{d\vec{J}}{dt} = \vec{\Omega} \wedge \vec{J} . \quad (34)$$

Cette équation d'évolution bien connue décrit le mouvement de *précession* (dite ici « de Larmor ») du moment cinétique autour de l'axe du champ :

- (1) La projection de \vec{J} sur la direction de \vec{B} est constante
- (2) En conséquence l'énergie (33) est conservée.
- (3) La composante de \vec{J} projetée sur le plan perpendiculaire à l'axe, notée \vec{J}_\perp , décrit un cercle à la fréquence angulaire $\Omega = \|\vec{\Omega}\|$.
- (4) On note aussi que, puisque $d\vec{J}/dt \perp \vec{J}$, la norme $J = \|\vec{J}\|$ est elle-même constante.

Dans le cas où le champ magnétique est de direction variable, même s'il vérifie toujours les équations (34), \vec{J} a un mouvement plus compliqué. La proposition (4) ci-dessus est toujours valable. La (1) l'est aussi, au sens que si on repère la direction variable de \vec{B} par le vecteur unitaire \vec{u}_B , la projection $\vec{J} \cdot \vec{u}_B$ est conservée. La (3), si \vec{u}_B varie, n'a plus de sens, mais vu ce qui précède, la composante transverse \vec{J}_\perp garde une longueur constante. Pour (2), en se limitant au cas où le champ \vec{B} tourne avec un vecteur rotation $\vec{\omega}$, on a $d\mathcal{H}/dt = \vec{J} \cdot d\vec{\Omega}/dt = \vec{J} \cdot (\vec{\omega} \wedge \vec{\Omega})$, qui sera nul ou pas selon la direction initiale de \vec{J} .

³⁰ Pour un objet classique dont la masse m est distribuée comme la charge q (c'est-à-dire que les masse volumique et densité de charge sont proportionnelles), on a $\gamma_{cl} = q/2m$. Pour des particules classiques ne vérifiant pas cette dernière condition, ou pour des particules quantiques dont une partie au moins de \vec{J} provient du spin, on généralise cela en posant $\gamma = g \gamma_{cl}$ où le facteur sans dimension g , appelé « facteur de Landé », et caractéristique la particule. Notamment pour un électron, la théorie de Dirac donne $g_e = 2$ et on mesure pour le proton et le neutron respectivement $g_p \approx 5,586$ et $g_n \approx -3,826$ (en prenant pour les deux nucléons le $\gamma_{cl} = e/2m_p$ du proton).

Note : Le couplage aux fluctuations du vide, conduit toutefois pour l'électron à $g_e - 2 = \alpha/\pi + \mathcal{O}(\alpha/\pi)^2 \approx 0,0023119\dots$, où $\alpha \approx 1/137$ est la « constante de structure fine », et pour le muon $g_\mu - 2 \approx 0,002332\dots$.

10.8 Trajectoires et portraits de phase des systèmes autonomes

Le traitement sur un pied d'égalité des positions et impulsions permet de considérer avec un œil neuf les trajectoires dans l'espace des phases. Notons tout d'abord qu'une telle trajectoire doit être contenue dans l'espace de dimension $2D - 1$ défini par la valeur de l'énergie $\mathcal{H}(\{q_i\}, \{p_i\}) = \mathcal{E} = \text{cste}$. En outre, la donnée d'un point particulier de l'espace des phases détermine la totalité de la trajectoire à laquelle il appartient, par intégration formelle des équations de Hamilton. Il en résulte que deux trajectoires distinctes ne sauraient se couper, et qu'une trajectoire donnée ne revient au même point que si elle revient exactement sur ses pas (trajectoire « fermée » périodique dans l'espace des phases et donc aussi en position).

Ces trajectoires peuvent être vues comme les lignes de champ d'un écoulement *stationnaire* dont le champ de vitesse est donné par le vecteur à $2D$ composantes $\vec{V}_{\mathcal{H}} \equiv (-\frac{\partial \mathcal{H}}{\partial p_i}, \frac{\partial \mathcal{H}}{\partial q_i})$. Cet écoulement est appelé « flot de \mathcal{H} ». On note bien sûr que cet écoulement est par construction normal au gradient de \mathcal{H} (dont les $2D$ composantes sont $(\frac{\partial \mathcal{H}}{\partial q_i}, \frac{\partial \mathcal{H}}{\partial p_i})$) et qu'il est donc tangent aux surfaces iso-énergie. On peut encore noter que ce champ de vitesse vérifie :

$$\vec{V}_{\mathcal{H}} = -\frac{\partial^2 \mathcal{H}}{\partial q_i \partial p_i} + \frac{\partial^2 \mathcal{H}}{\partial p_i \partial q_i} = 0 ,$$

sans hypothèse particulière³¹ sur \mathcal{H} , ce qui est caractéristique d'un écoulement *incompressible*, ainsi que nous l'établirons ultérieurement d'une façon plus générale (cf. § 11.5).

Un autre intérêt de cette analogie hydrodynamique tient à ce qu'elle fournit une interprétation géométrique du crochet de Poisson. Ainsi le crochet de Poisson $\{f, \mathcal{H}\}$ s'écrit comme le produit scalaire du gradient de la fonction f avec le champ de vitesse $\vec{V}_{\mathcal{H}}$, et constitue donc une « dérivée convective » comme il est usuel en hydrodynamique ; si le terme $\frac{\partial f}{\partial t}$ est nul, l'évolution de f sera réduite à ce seul terme. Une constante du mouvement est donc une quantité « transportée » par le flot hamiltonien.

Un autre point de vue consiste à interpréter ce même crochet de Poisson $\{f, \mathcal{H}\}$ comme le débit du flot $\vec{V}_{\mathcal{H}}$ travers une surface $f = \text{cste}$, ou symétriquement celui du flot défini par $\vec{V}_f \equiv (-\frac{\partial f}{\partial p_i}, \frac{\partial f}{\partial q_i})$ à travers les surfaces iso-énergie. En particulier, lorsque deux quantités f et g ont un crochet de Poisson nul (on dit qu'elles sont « en involution »), le flot de f conserve la valeur de g et inversement, et plus généralement, une évolution selon le flot de f et une autre selon le flot de g commutent.

Lorsqu'il est possible, le tracé des lignes du flot hamiltonien dans l'espace des phases est appelé un « **portrait de phase** » du système considéré, et celui-ci permet généralement d'obtenir des informations générales importantes sur la dynamique du système. Dans la pratique, ce n'est réellement possible que pour des systèmes à un degré de liberté, ou marginalement, en projection sur un système à deux degrés de liberté. Il est par contre toujours possible de représenter une coupe à deux dimensions de l'espace des phases, que l'on appelle « **section de Poincaré** », et dans laquelle une trajectoire sera représentée par une suite (finie ou infinie) de points.

11 Structure symplectique et transformations canoniques

11.1 Définition

Dans un espace vectoriel, le choix d'un *produit scalaire* définit une structure *euclydienne*, permettant d'introduire les notions de distance, de projection orthogonale, etc.. De façon similaire,

³¹ Autre que la condition de classe \mathcal{C}^2 .

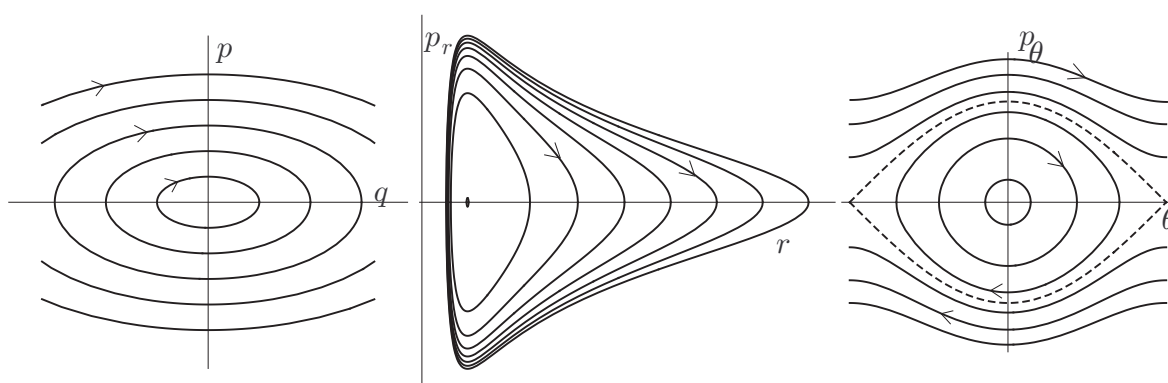


Fig. 10 – Portraits de phase (de gauche à droite) de l'oscillateur harmonique, du problème de Kepler (partie radiale), du pendule pesant

le crochet de Poisson dote l'espace des phases d'une structure géométrique dite *symplectique*. Il existe à la fois des analogies et de profondes différences entre ces deux types de géométrie, tenant essentiellement à ce que le produit scalaire est une forme bilinéaire symétrique, et que le crochet de Poisson est une « opération » bilinéaire antisymétrique et s'apparente donc davantage au produit vectoriel.

Cette structure symplectique sera notamment caractérisée, dans un jeu de coordonnées donné, par les **relations de conjugaison canoniques** :

$$\{q_i, p_j\} = \delta_{ij} \quad , \quad \{q_i, q_j\} = 0 \quad \text{et} \quad \{p_i, p_j\} = 0 \quad .$$

L'intérêt de la géométrie symplectique est d'achever et de systématiser la démarche que nous avons entreprise de traiter de la même façon les différentes coordonnées, qu'il s'agisse de positions ou d'impulsions. Pour ce faire, la première condition est de donner au crochet de Poisson lui-même une forme plus symétrique, ce qui peut être fait en écrivant :

$$\begin{aligned} \{f, g\} &= \sum_i \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} = \sum_{ij} \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_j} \delta_{ij} + \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} (-\delta_{i,j}) \\ &= \sum_{ij} \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_j} \{q_i, p_j\} + \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial p_i} \{p_i, q_j\} \end{aligned}$$

et si l'on introduit une notation commune x_n pour les $2D$ coordonnées de l'espace des phases :

$$\{f, g\} = \sum_{kl} \frac{\partial f}{\partial x_k} \frac{\partial g}{\partial x_l} \{x_k, x_l\} = \sum_{kl} \frac{\partial f}{\partial x_k} \frac{\partial g}{\partial x_l} S_{kl} = \left[\frac{\partial f}{\partial \mathbf{x}} \right]^\top \mathbf{S} \left[\frac{\partial g}{\partial \mathbf{x}} \right] , \quad (35)$$

puisque les termes où les deux dérivations seraient faites soit par rapport à q_i ou par rapport à p_j donnent des contributions nulles. Les coefficients $S_{kl} \equiv \{x_k, x_l\}$ sont les éléments de la matrice antisymétrique :

$$\mathbf{S} = \begin{matrix} & \begin{matrix} q_1 & \cdots & q_D & p_1 & \cdots & p_D \end{matrix} \\ \begin{matrix} q_1 \\ \vdots \\ q_D \\ p_1 \\ \vdots \\ p_D \end{matrix} & \left[\begin{array}{cc|cc} & & & \\ & \mathbf{0} & & \mathbf{1}_D \\ \hline & & & \\ -\mathbf{1}_D & & & \mathbf{0} \end{array} \right] \end{matrix} , \quad \text{où } \mathbf{1}_D \text{ est la matrice identité } (D \times D).$$

Cette matrice vérifie clairement les relations³² :

$$\mathbf{S}^\top = -\mathbf{S}, \quad \mathbf{S}^2 = -\mathbf{1}_{2D}, \quad \text{d'où } \mathbf{S}^{-1} = \mathbf{S}^\top = -\mathbf{S} \quad \text{et} \quad \det \mathbf{S} = 1 \quad (36)$$

NB : On pourra remarquer que le flot d'une fonction f quelconque, dont bien sûr \mathcal{H} , est défini simplement à l'aide de \mathbf{S} par $\vec{V}_f = \mathbf{S} [\partial f / \partial \mathbf{x}]$.

11.2 Transformations canoniques

11.2.1 Motivation et définition

De même que les propriétés géométrique d'un espace euclidien sont intrinsèques et ne dépendent pas de la base dans lequel le produit scalaire a été défini, la structure symplectique ne dépend pas du choix des variables canoniques utilisées. Dans le premier cas, on introduit donc des transformations qui respectent cette structure, dites *transformations orthogonales*, ou isométries.

Nous allons de même introduire des transformations respectant la structure symplectique, dites *transformations canoniques*, qui seront caractérisées par la propriété de conserver les crochets de Poisson³³. L'intérêt évident de ces transformations est qu'elle permettent de réaliser des changements de coordonnées, beaucoup plus généraux que les « transformations ponctuelles » utilisées dans le point de vue de Lagrange.

Soit donc une transformation de l'espace des phases $\mathbf{x} \mapsto \mathbf{X}(\mathbf{x}, t)$, supposée bijective, que nous pourrons encore écrire $(\{q_i\}, \{p_j\}) \mapsto (\{Q_i(q, p, t)\}, \{P_j(q, p, t)\})$ si nous voulons distinguer de façon typographique les impulsions des positions. Pour que cette transformation soit une transformation canonique, elle doit vérifier, à chaque instant t :

$$\{X_i, X_j\} = \{x_i, x_j\} = S_{ij} \quad \text{pour tous } i, j \in \{1, 2, \dots, 2D\}. \quad (37)$$

11.2.2 Caractérisation

Partant de la définition (35) du crochet de Poisson, nous déduisons de (37) :

$$\{X_i, X_j\} = S_{ij} = \sum_{kl} \frac{\partial X_i}{\partial x_k} \frac{\partial X_j}{\partial x_l} S_{kl},$$

ce qui peut être mis sous la forme matricielle, dite « condition symplectique » :

$$\mathbf{J} \mathbf{S} \mathbf{J}^\top = \mathbf{S} \quad \text{où } \mathbf{J} \text{ est la matrice jacobienne, d'éléments } J_{ij} = \frac{\partial X_i}{\partial x_j}. \quad (38)$$

Si cette condition sera réalisée, on aura pour toutes fonctions f et g de l'espace des phases :

$$\begin{aligned} \{f, g\}_{q,p} &= \sum_{kl} \frac{\partial f}{\partial x_k} \frac{\partial g}{\partial x_l} \{x_k, x_l\} = \sum_{kl} \left(\sum_i \frac{\partial f}{\partial X_i} J_{ik} \right) \left(\sum_j \frac{\partial g}{\partial X_j} J_{jl} \right) \{x_k, x_l\} \\ &= \sum_{ij} \frac{\partial f}{\partial X_i} \frac{\partial g}{\partial X_j} \underbrace{\left(\sum_{kl} J_{ik} \{x_k, x_l\} J_{jl} \right)}_{S_{ij}} = \sum_{ij} \frac{\partial f}{\partial X_i} \frac{\partial g}{\partial X_j} \{X_i, X_j\}, \end{aligned}$$

qui est bien l'expression du crochet de Poisson dans les nouvelles coordonnées.

³² Noter que $\mathbf{S}^2 = -\mathbf{1}$ et $\det \mathbf{S} = -1$ sont compatibles parce que la dimension de l'espace de phases est paire.

³³ De nombreux ouvrages les définissent en disant qu'elles « respectent la structure des équations de Hamilton », ce qui est assez vague et nécessite d'introduire un hamiltonien, puis disent que cette propriété doit être indépendante du hamiltonien, ce qui est peu cohérent...

11.2.3 Exemples simples

Voici quelques exemples élémentaires de transformations canoniques :

- (1) Des changements d'échelle $(q_i, p_j) \mapsto (Q_i, P_j) = (\lambda q_i, p_j/\lambda)$ où λ est un réel quelconque.
- (2) Toutes les transformations ponctuelles, pour lesquelles la matrice J prend la forme³⁴ :

$$J = \begin{array}{c} Q_1 \\ \vdots \\ Q_D \\ P_1 \\ \vdots \\ P_D \end{array} \left[\begin{array}{c|c} q_1 \cdots q_D & p_1 \cdots p_D \\ \hline \mathbf{F} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{F}^{\text{T}-1} \end{array} \right] \quad \text{avec} \quad F_{ij} = \frac{\partial Q_i}{\partial q_j} \quad (39)$$

- (3) Des échanges entre positions et impulsions $\{q_i, p_j\} \mapsto \{Q_i, P_j\} = \{p_i, -q_j\}$.

Les exemples (1) et (2) ne tirent pas réellement partie de la puissance du formalisme, puisqu'ils ne mélangent pas les impulsions et les positions, et auraient pu être faites dans le cadre lagrangien. Notons aussi que pour les cas exemples (1) ou (3), le changement d'échelle ou l'échange peut, bien sûr, ne porter que *certaines paires* de variables conjuguées.

11.2.4 Propriétés des transformations symplectiques

Les propriétés (36) de la matrice S et la condition (38) entraînent de nombreuses propriétés des transformations symplectiques que nous analysons brièvement ici. En premier lieu, elles indiquent de façon immédiate que^{35 36} :

$$\det J = \pm 1 \quad \text{et} \quad J^{-1} = -SJ^{\text{T}}S.$$

Si on sépare les $\{Q_i\}$ des $\{P_j\}$ comme à l'équation (39) et on décompose la matrice J en blocs, on peut expliciter J^{-1} :

$$J = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{C} \\ \hline \mathbf{B} & \mathbf{D} \end{array} \right], \quad \text{il vient :} \quad J^{-1} = \left[\begin{array}{c|c} \mathbf{D}^{\text{T}} & -\mathbf{C}^{\text{T}} \\ \hline -\mathbf{B}^{\text{T}} & \mathbf{A}^{\text{T}} \end{array} \right] \quad (40)$$

Si on forme alors les produits $J^{-1}J = JJ^{-1} = \mathbb{1}_{2D}$, on observe que :

- les produits $\mathbf{B}^{\text{T}}\mathbf{A}$, $\mathbf{A}\mathbf{B}^{\text{T}}$, $\mathbf{D}^{\text{T}}\mathbf{C}$ et $\mathbf{C}\mathbf{D}^{\text{T}}$ sont symétriques,
- $\mathbf{A}\mathbf{D}^{\text{T}} - \mathbf{C}\mathbf{B}^{\text{T}} = \mathbf{D}\mathbf{A}^{\text{T}} - \mathbf{B}\mathbf{C}^{\text{T}} = \mathbf{A}^{\text{T}}\mathbf{D} - \mathbf{B}^{\text{T}}\mathbf{C} = \mathbf{D}^{\text{T}}\mathbf{A} - \mathbf{C}^{\text{T}}\mathbf{B} = \mathbb{1}_D$.

³⁴ Une illustration importante est la transformation de Jacobi, qui permet de décomposer le problème à deux corps en deux mouvements indépendants. On pose en effet $\vec{R} = (m_1 \vec{r}_1 + m_2 \vec{r}_2)/(m_1 + m_2)$ et $\vec{p} = \vec{r}_2 - \vec{r}_1$, soit :

$$\mathbf{A} = \begin{bmatrix} m_1/M & m_2/M \\ -1 & 1 \end{bmatrix} \Rightarrow \mathbf{A}^{\text{T}-1} = \begin{bmatrix} 1 & 1/M \\ -m_2/M & m_1/M \end{bmatrix} \Rightarrow \vec{P} = \vec{p}_1 + \vec{p}_2 \quad \text{et} \quad \vec{\pi} = (m_1 \vec{p}_2 - m_2 \vec{p}_1)/M,$$

et on montre ensuite : $\vec{p}_1^2/m_1 + \vec{p}_2^2/m_2 = \vec{P}^2/M + \vec{\pi}^2(1/m_1 + 1/m_2)$.

³⁵ En fait, on peut même montrer que $\det J = \pm 1$.

³⁶ La quantité $J = |\det J|$, souvent notée $J = \frac{D(\{Q_i\}, \{P_j\})}{D(\{q_i\}, \{q_j\})}$ est le jacobien de la transformation, et sa valeur égale à 1 indique que la transformation canonique conserve l'élément de volume dans l'espace des phases.

11.2.5 Transformations canoniques infinitésimales

Des exemples moins triviaux seront fournis par des transformations *infinitésimales* que l'on écrit sous la forme :

$$x \mapsto X = (\mathbb{1} + \varepsilon \{*, G\})x = x + \varepsilon \{x, G\} ,$$

où G est une fonction de x (c'est à dire de $\{q_i\}$ et $\{p_j\}$), appelée « générateur » de la transformation considérée, et où l'étoile $*$ représente la variable ou fonction auquel l'opérateur est appliqué. Son caractère canonique découle de $J = \mathbb{1} + \varepsilon SG^{(2)}$ où $G^{(2)}$ est la matrice hessienne de G ou plus directement de :

$$\{X_i, X_j\} = \{x_i, x_j\} + \varepsilon \left(\{x_i, \{x_j, G\}\} + \{\{x_i, G\}, x_j\} \right) + \mathcal{O}(\varepsilon^2) = \{x_i, x_j\} + \mathcal{O}(\varepsilon^2) ,$$

car le terme entre crochets, d'après l'identité de Jacobi, vaut $\{G, \{x_i, x_j\}\}$ est identiquement nul puisque $\{x_i, x_j\}$ est une constante. Le point important est que G est absolument quelconque, et que l'on engendre ainsi une infinité de transformations indépendantes.

Ainsi, des choix possibles sont :

- La transformation $(q(t), p(t)) \mapsto (q(t + dt), p(t + dt))$ obtenue avec $G = \mathcal{H}$, et $\varepsilon = dt$, ce qui montre que l'évolution temporelle est une transformation canonique.
- la transformations $(\vec{r}, \vec{p}) \mapsto (\vec{r} + \vec{\varepsilon}, \vec{p})$ est une simple translations l'espace des positions, engendrée par $G = \vec{p} \cdot \vec{\varepsilon}$; de même, $(\vec{r}, \vec{p}) \mapsto (\vec{r}, \vec{p} + \vec{\varepsilon})$ est une translation dans l'espace des impulsions, engendrée par $G = -\vec{r} \cdot \vec{\varepsilon}$.
- La transformation $(\vec{r}, \vec{p}) \mapsto (\vec{r} + \vec{\varepsilon} \wedge \vec{r}, \vec{p} + \vec{\varepsilon} \wedge \vec{p})$ est une rotation dans l'espace des phases (sans mélange des impulsions et positions, engendrée par $G = \vec{L} \cdot \vec{\varepsilon}$;
- les changements d'échelle $(\vec{q}, \vec{p}) \mapsto ((1 + \varepsilon)\vec{q}, (1 - \varepsilon)\vec{p})$, obtenu avec $G = \sum_i p_i q_i$.

Généralisation Toutes ces transformations infinitésimales peuvent être formellement étendues à des transformations finies en utilisant l'exponentiation des opérateurs, définie par son développement en série entière. À titre d'exemple, on peut noter que, pour toute fonction $f(q, p)$ de l'espace des phases :

$$\{f, p\} = \frac{\partial f}{\partial q} \quad \text{d'où, par récurrence} \quad \{*, p\}^n = \underbrace{\{\dots\{*, p\}, p\}, \dots, p\}_n = \frac{\partial^n}{\partial q^n} * ,$$

ce qui permet d'écrire :

$$T(a) f \equiv \exp(a\{*, p\}) f = \sum_{n=0}^{\infty} \frac{a^n}{n!} \frac{\partial^n f}{\partial q^n} (q, p) ,$$

où l'on reconnaît immédiatement le développement de Taylor de $f(q+a, p)$. Ainsi l'opérateur exponentiel $T(a)$ utilisé généralise-t-il à une translation spatiale finie a la transformation infinitésimale précédemment évoquée.

Il en est bien sûr de même pour l'évolution temporelle entre l'instant $t_0 = 0$ et l'instant t avec le hamiltonien \mathcal{H} *indépendant du temps*, qui peut être décrite à l'aide de « l'opérateur d'évolution » :

$$U = \exp \left(t\{*, \mathcal{H}\} \right) = \mathbb{1} + t\{*, \mathcal{H}\} + \frac{t^2}{2!} \{\{*, \mathcal{H}\}, \mathcal{H}\} + \frac{t^3}{3!} \{\{\{*, \mathcal{H}\}, \mathcal{H}\}, \mathcal{H}\} + \dots , \quad (42)$$

tel que pour toute fonction ne dépendant pas explicitement du temps :

$$f(q(t), p(t)) = U(t)f(q(0), p(0)) .$$

Ce résultat peut être obtenu très simplement en notant que l'équation d'évolution (32), dans le cas où f ne dépend pas directement du temps, peut être écrite :

$$\forall f, \quad \frac{d}{dt} (U(t)f_0) = \{U(t)f_0, \mathcal{H}\} \quad \Rightarrow \quad \frac{d}{dt} U(t) = \{*, \mathcal{H}\} \circ U(t) ,$$

qui est une équation différentielle sur U dont la solution est donnée par l'équation (42).

11.3 Fonctions génératrices

Une méthode plus immédiatement opératoire, pour construire les transformations canoniques, repose sur l'indétermination du hamiltonien que nous avons soulignée plus haut (*cf.* § 10.6 à la page 48).

11.3.1 Présentation générale

De façon naturelle, la conservation des crochets de Poisson en général doit en particulier conserver les équations de Hamilton, qui n'en sont qu'un cas particulier, mais pas nécessairement le hamiltonien lui-même.

La seule fonction dont le choix est imposé par la physique est l'action et son caractère stationnaire, qui laisse la liberté d'ajouter la dérivée totale dG/dt d'une quelconque fonction de l'espace des phases. Si l'on considère deux jeux de composantes équivalents, (q, p) et (Q, P) , ceux-ci doivent vérifier³⁷ :

$$\delta \mathcal{S}(\gamma) = \delta \int_{\gamma} p dq - \mathcal{H}(q, p, t) dt = \delta \int_{\gamma} P dQ - \mathcal{K}(Q, P, t) dt = 0 , \quad (43)$$

l'intégrale étant prise sur un même chemin de l'espace des phases, mais paramétrisé de deux façons différentes. Une condition suffisante est donc que les deux intégrales diffèrent par l'introduction dans la seconde d'une dérivée $dG(q, p, t)/dt$ dont la contribution s'annule lorsqu'on évalue la variation $\delta \mathcal{S}$. La différentielle de la fonction G doit alors vérifier :

$$dG = (p dq - \mathcal{H}(q, p, t) dt) - (P dQ - \mathcal{K}(Q, P, t) dt) . \quad (44)$$

où $\mathcal{K}(Q, P, t)$ est bien sûr le nouveau hamiltonien.

Compte tenu de la présence de dq , dQ et dt dans cette équation, les variables les plus naturelles pour G sont q , Q et t , et en utilisant $G = G_1(q, Q, t)$ on aura donc :

$$\frac{\partial G_1}{\partial q_i} (q, Q, t) = p_i , \quad (45a)$$

$$\frac{\partial G_1}{\partial Q_i} (q, Q, t) = -P_i , \quad (45b)$$

$$\frac{\partial G_1}{\partial t} (q, Q, t) = -(\mathcal{H}(q, p, t) - \mathcal{K}(Q, P, t)) . \quad (45c)$$

Ces équations (45a) et (45b) ne valent bien sûr que si elles contiennent une caractérisation complète de la transformation, c'est à dire que la donnée de q et Q définit bien un point et un

³⁷ Les variables q , p , ou Q et P sont ici des raccourcis pour des jeux de variables multidimensionnels $\{q_i\}$, $\{p_i\}$, où les produits comme $p dq$ doivent être compris comme des sommations $\sum_i p_i dq_i$, ou encore $\sum_{\alpha} \vec{p}_{\alpha} \cdot d\vec{r}_{\alpha}$ dans le cas de coordonnées cartésiennes pour plusieurs corps.

seul dans l'espace des phases, ou, de façon équivalente, si (45a) peut être inversée pour déterminer Q en fonction de q et p .

Dans le cas contraire, l'une au moins des transformées de Legendre $G_2(q, P, t) = G_1(q, Q, t) + PQ$, ou $G_3(p, Q, t) = G_1(q, Q, t) - pq$ ou encore $G_4(p, P, t) = G(q, Q, t) + PQ - pq$ permettra d'y parvenir avec des équations extrêmement similaires.

La vérification du caractère canonique, au sens défini plus haut, de telles transformations, est un exercice mathématique laborieux mais sans difficulté, si l'on voit que la conservation du crochet de Poisson fondamental $\{q_i, p_j\} = \delta_{ij}$ est suffisante. Toutefois, cette vérification n'est pas vraiment nécessaire, puisque l'équation (43) assure que l'évolution des nouvelles variables satisfait aux équations de Hamilton, et donc conserve le crochet de Poisson (32), *indépendamment* du choix particulier qui est fait pour \mathcal{H} : il suffit donc de prendre formellement $\mathcal{H} = q_i$ ou p_i pour vérifier que les crochets de Poisson canoniques seront conservés.

On notera que chaque fois que la fonction génératrice G dépend du temps, le nouveau hamiltonien $\mathcal{H}(Q, P, t)$ ne s'obtient pas *seulement* par « invariance ponctuelle » en changeant de variables avec $\mathcal{H}'(Q, P, t) = \mathcal{H}(q(Q, P, t), p(Q, P, t), t)$: il faut en outre y ajouter un terme provenant de la variation temporelle de la fonction G . Ainsi on a :

$$\mathcal{H}(Q, P, t) = \mathcal{H}'(Q, P, t) + \frac{\partial G}{\partial t} = \mathcal{H}(q(Q, P, t), p(Q, P, t), t) + \frac{\partial G}{\partial t}(q(Q, P, t), Q, t).$$

Exemples élémentaires Les exemples simples et fondamentaux de fonctions génératrices sont les suivants :

- (1) La fonction $G_2(q, P, t) = \sum_j P_j F_j(q, t)$, pour laquelle on a $Q_i = \partial G_2 / \partial P_i = F_i(q, t)$ et $p_i = \partial G_2 / \partial q_i = \sum_j P_j \frac{\partial F_j}{\partial q_i}$, redonne une fois de plus les transformations *ponctuelles*, éventuellement dépendantes du temps ; un cas particulier est évidemment la transformation identité, obtenue avec $F_i(q, t) = q_i$,
- (2) La fonction $G_3(p, Q, t) = -\sum_j Q_j F_j(p, t)$ qui en est le pendant exact en échangeant les rôles joués par p et q ;
- (3) La fonction $G_1(q, Q, t) = qQ$ qui échange le rôle de p et q (attention au changement de signe !)
- (4) La fonction $G_2(q, P, t) = qP + \varepsilon G(q, P, t)$ qui redonne la transformation canonique infinitésimale engendrée par $G(q, P(q, p, t), t)$.

11.3.2 Un exemple en détail : le référentiel tournant

Un système mécanique est décrit par les variables canoniques q et p dans un référentiel galiléen \mathcal{R} . On se propose de décrire le mouvement dans le référentiel \mathcal{R}' qui est en rotation uniforme à la vitesse angulaire $\vec{\Omega}$ par rapport à \mathcal{R} . Ceci va nous fournir un exemple relativement simple et physiquement important de transformation canonique dépendant du temps ³⁸

Pour simplifier, nous nous concentrons sur une particule décrite par des coordonnées cartésiennes $q \equiv \vec{r}$ et $p \equiv \vec{p}$ dans \mathcal{R} , et nous notons par \vec{R} et \vec{P} les nouvelles variables canoniques dans \mathcal{R}' . Il est commode, pour les calculs qui suivent, de poser $\vec{\Omega}t = \vec{\Theta}$ et de noter $R_{\vec{\Theta}}$ la rotation correspondante. On aura bien sûr $\vec{R} = R_{-\vec{\Theta}}(\vec{r})$ et $\vec{P} = R_{-\vec{\Theta}}(\vec{p})$, mais nous allons le retrouver via le formalisme des fonctions génératrices.

³⁸ Cet exemple, mettant en jeu une transformation ponctuelle (sans mélange des positions et de impulsions), est en fait plus simple dans l'approche Lagrangienne. Il est traité ici à des fins pédagogiques.

Considérons donc la transformation canonique engendrée par la fonction $G_2(\vec{r}, \vec{P}) = \vec{P} \cdot \mathbf{R}_{-\vec{\Omega}t}(\vec{r})$. Les équivalents pour G_2 de (45a) et (45b) donnent :

$$\vec{R} = \frac{\partial G_2}{\partial \vec{P}} = \mathbf{R}_{-\vec{\Theta}}(\vec{r}) \quad \text{et} \quad \vec{p} = \frac{\partial G_2}{\partial \vec{r}} = \vec{P} \cdot \frac{d}{d\vec{r}} \mathbf{R}_{-\vec{\Theta}}(\vec{r}).$$

On a ainsi, en choisissant les axes orthonormés (x, y, z) tels que $\vec{\Theta} = \Theta \vec{u}_z$:

$$\vec{R} = \mathbf{R}_{-\vec{\Theta}}(\vec{r}) = \begin{bmatrix} x \cos \Theta - y \sin \Theta \\ x \sin \Theta + y \cos \Theta \\ z \end{bmatrix},$$

et de la première équation, on déduit la fonction génératrice :

$$G_2(\vec{r}, \vec{P}) = \vec{P} \cdot \mathbf{R}_{-\vec{\Theta}}(\vec{r}) = P_x(x \cos \Theta - y \sin \Theta) + P_y(x \sin \Theta + y \cos \Theta) + P_z z.$$

On peut en tirer l'expression de \vec{p} en fonction de \vec{P} , puis par inversion celle de \vec{P} en fonction de \vec{p} . Obtient ainsi comme attendu :

$$\vec{p} = \frac{\partial G_2}{\partial \vec{r}} = \begin{bmatrix} P_x \cos \Theta + P_y \sin \Theta \\ P_x \sin \Theta - P_y \cos \Theta \\ P_z \end{bmatrix} = \mathbf{R}_{\vec{\Theta}}(\vec{P}) \quad \xrightarrow{\text{Inversion}} \quad \vec{P} = \mathbf{R}_{-\vec{\Theta}}(\vec{p}),$$

Le point le plus intéressant est celui de la modification du hamiltonien, ainsi que des équations de Hamilton qui en dérivent. Nous supposeront à partir d'ici, car c'est très souvent le cas, que le hamiltonien $\mathcal{H}(\vec{r}, \vec{p}, t)$ est un hamiltonien standard (avec une énergie cinétique $K = \vec{p}^2/2m = \vec{P}^2/2m$) et que le potentiel V est lui-aussi *invariant par rotation*, ce qui se traduit par :

$$\mathcal{H}'(\vec{R}, \vec{P}, t) = \mathcal{H}(\vec{r}, \vec{p}, t) = \mathcal{H}(\mathbf{R}_{\vec{\Theta}}(\vec{R}), \mathbf{R}_{\vec{\Theta}}(\vec{P}), t) = \mathcal{H}(\vec{R}, \vec{P}, t)$$

En notant que :

$$\begin{aligned} \frac{\partial \vec{r}(\vec{R}, t)}{\partial t} &= \frac{\partial}{\partial t} (\mathbf{R}_{\vec{\Omega}t}(\vec{R})) = \vec{\Omega} \wedge \vec{R} \\ \text{donc} \quad \frac{\partial G_2(\vec{r}, \vec{P}, t)}{\partial t} &= \frac{\partial \vec{P} \cdot \vec{r}(\vec{R}, t)}{\partial t} = \vec{P} \cdot \frac{\partial \vec{r}(\vec{R}, t)}{\partial t} = \vec{P} \cdot (\vec{\Omega} \wedge \vec{R}) \end{aligned}$$

on obtient finalement :

$$\mathcal{K}(\vec{R}, \vec{P}, t) = \mathcal{H}'(\vec{R}, \vec{P}, t) + \frac{\partial G_2(\vec{r}, \vec{P}, t)}{\partial t} = \mathcal{H}(\vec{R}, \vec{P}, t) + \vec{\Omega} \cdot (\vec{R} \wedge \vec{P}) = \mathcal{H}(\vec{R}, \vec{P}, t) + \vec{\Omega} \cdot \vec{L}. \quad (46)$$

Les équations de Hamilton sont alors :

$$\frac{d\vec{R}}{dt} = \frac{\partial \mathcal{K}}{\partial \vec{P}} = \frac{\vec{P}}{m} + \vec{\Omega} \wedge \vec{R} \quad (47a)$$

$$\frac{d\vec{P}}{dt} = -\frac{\partial \mathcal{K}}{\partial \vec{R}} = -\frac{\partial V}{\partial \vec{R}} + \vec{\Omega} \wedge \vec{P} \quad (47b)$$

on observe sans surprise que $\vec{P} \neq m\dot{\vec{R}}$, et que le moment angulaire de l'impulsion $\vec{L} = \vec{R} \times \vec{P}$ qui figure dans (46) n'est donc pas celui du moment de la quantité de mouvement.

En reportant (47b) dans la dérivée de (47a), on obtient l'équation du mouvement « à la Newton » :

$$m\ddot{\vec{R}} = -\frac{\partial V}{\partial \vec{R}} + 2\vec{\Omega} \wedge \vec{P} + m\vec{\Omega} \wedge (\vec{\Omega} \wedge \vec{R}) .$$

On retrouve ainsi que l'équation du mouvement dans \mathcal{R}' est *formellement analogue* à celle que l'on aurait dans \mathcal{R} à condition d'ajouter la force (fictive) d'inertie d'entraînement $m\vec{\Omega} \wedge (\vec{\Omega} \wedge \vec{R})$ et la force (fictive) de Coriolis $2m\vec{\Omega} \wedge \vec{P}$. On notera que ces deux forces inertielles sont issues du seul terme $\vec{\Omega} \cdot \vec{L}$ dans (46).

Toutefois dans cette *analogie formelle*, il conviendra de faire comme si on avait $\dot{\vec{R}} = \vec{P}/m$ comme c'est le cas dans les équations de Newton dans \mathcal{R} .

11.3.3 Identité des deux points de vue ★

Comme exposé plus haut, il n'est pas nécessaire de vérifier *par le calcul* que les équations de transformation (45a)–(45c) conservent les crochets de Poisson.

Il est par contre très utile de faire le lien entre le point de vue géométrique développé au § 11.2 ci-dessus, et la méthode de construction que nous venons d'exposer. Nous allons dans ce but montrer que la condition (38) est *équivalente* à la relation de fermeture de la forme différentielle dG de (44), et donc à l'existence de la fonction génératrice G . Il nous faut pour cela comparer les dérivées croisées de G , à savoir $\partial p_i / \partial Q_j$ et $-\partial P_j / \partial q_i$, évaluées respectivement à $\{q_i\}$ fixé et à $\{Q_i\}$ fixé.

Pour un système à un degré de liberté, le calcul repose sur des différentiations assez élémentaires. On peut en effet écrire :

$$\left. \frac{\partial P}{\partial q} \right|_Q = \left. \frac{\partial P}{\partial q} \right|_p + \left. \frac{\partial P}{\partial p} \right|_q \left. \frac{\partial p}{\partial q} \right|_Q \quad \text{avec} \quad \left. \frac{\partial p}{\partial q} \right|_Q = - \left. \frac{\partial Q}{\partial q} \right|_p \left. \frac{\partial Q}{\partial p} \right|_q^{-1} = - \left. \frac{\partial Q}{\partial q} \right|_p \left. \frac{\partial p}{\partial Q} \right|_q .$$

En mettant le terme $\partial p / \partial Q|_q$ en facteur, il vient :

$$\left. \frac{\partial P}{\partial q} \right|_Q = - \left. \frac{\partial p}{\partial Q} \right|_q \left(\left. \frac{\partial P}{\partial p} \right|_q \left. \frac{\partial Q}{\partial q} \right|_p - \left. \frac{\partial P}{\partial q} \right|_p \left. \frac{\partial Q}{\partial p} \right|_q \right) ,$$

On reconnaît dans la parenthèse le crochet de Poisson $\{Q, P\}$, ce qui confirme que dG est fermée si et seulement si ce crochet de Poisson canonique prend la valeur attendue 1.

En d'autres termes, la jacobienne J de (40) est ici une matrice 2×2 , les blocs A, B, C et D sont des simples nombres, et le terme entre parenthèses est $\det J = AD - BC$. On vérifie donc dans ce cas simple l'équivalence entre la fermeture de dG et les conditions du § 11.2.4 sur la jacobienne J .

On peut généraliser ce résultat au cas *multidimensionnel*. En prenant en compte le caractère matriciel des équations, nous pouvons partir des différentielles des nouvelles variables par rapport aux anciennes, qui s'écrivent :

$$dQ_i = \sum_{i,j} A_{ij} dq_j + C_{ij} dp_j , \quad (48a)$$

$$dP_i = \sum_{i,j} B_{ij} dq_j + D_{ij} dp_j , \quad (48b)$$

où nous avons introduit les quatre matrices blocs composant la matrice J , comme dans (40). Nous supposons, comme précédemment, que la donnée des $\{Q_i\}$ et des $\{q_i\}$ détermine entièrement les $\{p_i\}$, ce qui équivaut à supposer que

la matrice C est inversible. En imposant $dq_i = 0$ dans (48a), et en reportant l'expression de dp_i ainsi obtenue dans (48b), on obtient successivement :

$$\left. \frac{\partial Q_i}{\partial p_j} \right|_{q_k} = [C]_{ij}, \quad \left. \frac{\partial p_i}{\partial Q_j} \right|_{q_k} = [C^{-1}]_{ij} \quad \text{et} \quad \left. \frac{\partial P_i}{\partial q_j} \right|_{Q_k} = B_{ij} - \sum_{lm} D_{il} C_{lm}^{-1} A_{mj}.$$

Or le second membre de la seconde équation peut s'écrire comme le produit de matrices :

$$B - DC^{-1}A = (B - DC^{-1}A) C^T C^T^{-1} = (B C^T - DC^{-1} A C^T) C^T^{-1},$$

que les relations (41) entre les quatre blocs permettent de réécrire :

$$\left(B C^T - DC^{-1} (A C^T)^T \right) C^T^{-1} = \left(B C^T - DC^{-1} (C A^T) \right) C^T^{-1} = \left(\underbrace{B C^T - D A^T}_{-1_D} \right) C^T^{-1} = -C^T^{-1}$$

ce qui se traduit par la relation :

$$\left. \frac{\partial P_i}{\partial q_j} \right|_{Q_k} = -[C^T^{-1}]_{ij} \quad \text{et donc} \quad \left. \frac{\partial P_i}{\partial q_j} \right|_{Q_k} = -[C^{-1}]_{ji} = -\left. \frac{\partial p_j}{\partial Q_i} \right|_{q_k},$$

qui est bien la relation de Schwarz pour la forme différentielle dG . □

11.4 Transformation canonique engendrée par l'action

Au paragraphe § 9 nous avons établi l'expression de la différentielle de la *fonction* action $S(q_0, t_0; q_1, t_1)$:

$$dS = p_1 dq_1 - p_0 dq_0 - \mathcal{H}(q_1, p_1, t_1) dt_1 + \mathcal{H}(q_0, p_0, t_0) dt_0 \quad (49)$$

sans montrer qu'il s'agissait d'une différentielle totale. Or nous pouvons maintenant constater que les relations d'identité des dérivées croisées (théorème de Schwarz) :

$$\frac{\partial(p_1)_i}{\partial t_1} = -\frac{\partial \mathcal{H}_1}{\partial (q_1)_i} \quad \text{et} \quad \frac{\partial(p_0)_i}{\partial t_0} = -\frac{\partial \mathcal{H}_0}{\partial (q_0)_i}$$

ne font que ré-exprimer la seconde équation de Hamilton aux instants initiaux et finaux, tandis que les autres relations entre dérivées croisées résultent simplement de l'indépendance des variables canoniques.

Cette fonction action, aussi désignée comme « fonction principale de Hamilton », nous fournit un exemple subtil et très important de transformation canonique. Considérons en effet la fonction génératrice :

$$G_1(q, Q, t) = S(Q, t_0; q, t)$$

construite à l'aide de la fonction action S ci-dessus, les coordonnées $q_1 \equiv q$ à l'instant t sont les « anciennes » variables q , et les coordonnées $q_0 \equiv Q$ à l'instant initial t_0 sont les « nouvelles » variables. On a alors :

$$p_i = +\frac{\partial S}{\partial (q_1)_i} = (p_1)_i \quad \text{et} \quad P_i = -\frac{\partial S}{\partial (q_0)_i} = (p_0)_i \quad (50)$$

qui réalisent effectivement la transformation attendue sur les impulsions. Le résultat important est l'expression du nouveau hamiltonien en utilisant (49) :

$$\mathcal{H}(Q, P, t) = \mathcal{H}(q_1(Q, P, t), p_1(Q, P, t), t) + \frac{\partial S}{\partial t_1}(q_0, t_0; q_1, t_1) = 0. \quad (51)$$

Ceci signifie que le remplacement (évidemment dépendant du temps) des variables « usuelles » q et p par leurs valeurs initiales est une transformation canonique, que celle-ci conduit à des variables qui n'évoluent plus, et que la fonction action, qui était jusque là un peu dépourvue de sens physique, est la fonction génératrice de cette transformation très particulière.

On pourrait aussi bien considérer la transformation canonique engendrée par la fonction $G_1(q, Q, t) = -S(q, t; Q, t_1)$ où les rôles des anciennes et nouvelles variables est échangé par rapport au cas précédent. Cela signifie que les variables canoniques Q et P sont à chaque instant identifiées aux valeurs prises par $q(t)$ et $p(t)$, ce qui revient à suivre le mouvement : l'évolution temporelle disparaît et on doit donc de nouveau trouver un nouveau hamiltonien identiquement nul. C'est ce que l'on vérifie en écrivant :

$$\mathcal{K}(Q, P, t) = \mathcal{H}'(Q, P, t) + \frac{\partial G_1}{\partial t}(q_0, q_1, t) = \mathcal{H}(q_0, p_0, t) - \frac{\partial S}{\partial t_0}(q_0, t_0; q_1, t_1) = 0, \quad (52)$$

Nous pouvons conclure de cette étude le résultat très important que l'évolution temporelle est une transformation canonique, ainsi que nous l'avons pressenti en constatant que le flot de \mathcal{H} « emportait » les constantes du mouvement, et que nous l'avons aussi montré en exhibant l'opérateur d'évolution U au § 11.2.

11.5 Théorèmes de Liouville et de Poincaré

Le théorème de Liouville affirme que, pour un système autonome, le flot hamiltonien conserve le volume dans l'espace des phases ³⁹. En réalité, cela résulte simplement de ce que – ainsi que nous venons de l'établir – l'évolution temporelle est un cas particulier de transformation canonique. En effet, l'égalité (38) implique que le jacobien $J = |\det(J)|$ d'une transformation canonique, qui exprime le rapport entre les éléments de volume, est nécessairement égal à 1.

Ce théorème de conservation est d'une grande importance en physique statistique. Il permet en particulier de montrer que si l'on définit une densité de probabilité (ou plus simplement une densité de particules) dans l'espace des phases $\rho(q, p, t)$, celle-ci est « emportée » par le flot hamiltonien, et qu'elle évolue donc selon la loi $\frac{\partial \rho}{\partial t} + \{\rho, \mathcal{H}\} = 0$. Ceci conduit à l'équation dite « de Boltzmann » :

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial \vec{r}} \cdot \dot{\vec{r}} + \frac{\partial \rho}{\partial \vec{p}} \cdot \dot{\vec{p}} = C(\rho),$$

où les deuxième et troisième termes décrivent l'évolution « à une particule » sous l'effet de \mathcal{H} , tandis que le second membre $C(\rho)$ décrit les interactions entre les particules, non prises en compte dans \mathcal{H} , et est appelé « terme de collisions ».

Le théorème du retour (ou « de récurrence ») de Poincaré assure que, pour un système autonome borné (en position et en impulsion), tout voisinage, aussi petit soit-il, d'un point d'une trajectoire sera nécessairement visité à nouveau par la trajectoire. Il résulte à peu près directement de la conservation du volume. Soit en effet l'application g décrivant l'évolution hamiltonienne durant un temps fini δt , et U_0 un voisinage du point x_0 ; on introduit alors la suite infinie de domaines $U_n \equiv g^n(U_0)$, qui ne peuvent pas être tous disjoints puisque ils ont tous le même volume, et que le volume accessible dans l'espace des phases est fini. Il y a donc deux entiers $n > m \geq 0$ tels que $U_n \cap U_m = g^n(U_0) \cap g^m(U_0) \neq \emptyset$, et on en déduit que $g^{n-m}(U_0) \cap U_0 \neq \emptyset$: au bout d'un certain temps, appelé « temps de récurrence de Poincaré », certains points du voisinage

³⁹ Ce résultat est souvent démontré à grand renfort de calculs plus ou moins rigoureux sur les jacobiens de transformations partielles qui ne sont pas toujours définies (on trouvera cependant une justification mathématique de cette approche dans l'ouvrage d'Arnold).

U_0 y reviennent. Le côté paradoxal de cet énoncé, contradictoire avec notre expérience pratique et avec l'irréversibilité de physique statistique, tient à ce que ce temps croît extrêmement vite avec le nombre de degrés de liberté, et que dans bien des systèmes macroscopiques il dépasse, de nombreux ordres de grandeur, l'âge de l'univers.

12 Systèmes intégrables ★

Compte tenu du format de ce cours, et des outils mathématiques dont nous disposons, les notions qui vont être exposées dans ce qui suit ne le seront que de façon qualitative ou informelle et le plus souvent sans démonstration. Pour le lecteur légitimement désireux d'un exposé plus rigoureux et prêt à en consentir à l'effort requis, nous renvoyons principalement à l'ouvrage d'Arnold cité dans la bibliographie (originellement écrit en russe, il a été traduit en anglais et même en français).

12.1 Notion de système intégrable

Du point de vue qualitatif, un système intégrable est d'un système dont l'évolution est suffisamment régulière, ou « ordonnée », pour qu'on puisse prédire son évolution à long terme, et qu'il soit possible de prédire l'évolution à des temps arbitrairement longs d'un point quelconque de l'espace des phases choisi comme condition initiale. Malgré le déterminisme de la mécanique classique qui n'est pas mis en cause ici, de tels systèmes sont relativement rares, et il n'est pas inutile de s'interroger à leur sujet. Nous montrerons plus bas que les systèmes autonomes (c'est-à-dire de hamiltonien indépendant du temps) à un degré de liberté sont intégrables, mais il suffit de deux degrés de liberté pour que la question soit beaucoup plus ardue.

Reformulée en termes géométriques, la question est de savoir si une trajectoire donnée du flot hamiltonien est susceptible d'explorer toute la variété de dimension $2D - 1$ définie par la conservation de l'énergie (on parle alors de comportement ergodique), ou s'il existe des contraintes plus fortes sur ces trajectoires.

Nous admettrons le théorème dû à Liouville selon lequel un système autonome à D degrés de liberté est intégrable s'il existe D constantes du mouvement C_1, \dots, C_D (dont l'énergie) indépendantes et en involution, c'est-à-dire dont le crochets de Poisson soient nuls lorsqu'on les prend deux à deux soit $\{C_i, C_j\} = 0$ pour tout couple (i, j) . Dans ces conditions, on constate que le mouvement est contraint à se déplacer sur une variété de dimension D dans l'espace de phases.

Nous admettrons encore le théorème dû à Arnold selon lequel, pour un système intégrable dont l'espace des phases accessible est *borné* (par exemple par la donnée de l'énergie), la variété explorée par le système a la topologie d'un *tore* à D dimensions \mathbb{T}^D dans l'espace des phases. Plus précisément, l'évolution du système peut être paramétrée continûment par D angles indépendants ϕ_1, \dots, ϕ_D , tels que le point de l'espace des phases soit revienne au même point lorsqu'on change un (ou plusieurs) de ces angles d'un multiple entier de 2π . Ce tore est appelé « tore invariant », parce qu'il est conservé par le flot de $\mathcal{H} = \mathcal{E} = C_1$, mais aussi par le flot associé aux $D - 1$ autres constantes du mouvement $\{C_2, \dots, C_D\}$. De même que deux trajectoires dans l'espace des phases ne peuvent se couper, deux tores invariants associé à des valeurs distinctes des constantes ne peuvent se couper, et la variété définie par la valeur \mathcal{E} de l'énergie a donc une structure « feuilletée » par l'ensemble des $D - 1$ tores de dimension $2D - 1$ associés aux autres constantes du mouvements.

12.2 Variables d'action

Sur le tore invariant $\mathbb{T}^D(C_1, \dots, C_D)$, on peut définir la notion d'homotopie : deux lacets (courbes fermées) sont homotopes si l'on peut transformer l'un en l'autre par déformation continue sans sortir du tore. Ainsi, sur un tore à deux dimensions que nous pouvons visualiser dans l'espace à trois dimensions, tous les « parallèles » sont homotopes, de même que tous les « méridiens », mais il n'est pas possible de passer de l'un à l'autre, ni à un lacet qui s'enroulerait trois fois autour du petit diamètre tout en faisant une révolution autour du grand diamètre ⁴⁰.

Pour chaque classe d'homotopie, on peut définir un ensemble de D indices n_1, \dots, n_D entiers relatifs, qui caractérisent le nombre de tours fait par chacun des angles avant que le lacet ne se referme. On est ainsi amené à définir les lacets élémentaires γ_i qui correspondent à des indices tous nuls sauf pour l'indice n_i qui vaut 1 : un tel lacet peut être obtenu en faisant varier l'un des angles ϕ_i de 2π tout en gardant les autres constants.

Pour chacune des D classes de lacets élémentaires (qui ne sont bien-sûr pas des trajectoires) on définit alors la variable d'action correspondante par l'intégrale, dont la dimension est visiblement celle de l'action :

$$J_i = \frac{1}{2\pi} \oint_{\gamma_i} \sum_j p_j dq_j . \quad (53)$$

On peut montrer que cette intégrale est *indépendante du lacet* choisi dans sa classe d'homotopie en raison de la propriété *d'involutions* des constantes définissant le tore invariant ; il suffit pour cela de vérifier que la circulation de la forme $\sum_j p_j dq_j$ est nulle pour tout lacet homotope à un point sur \mathbb{T}^D .

Il résulte bien sûr de ces définitions et propriétés que l'intégrale sur un lacet quelconque γ d'indices n_1, \dots, n_D s'écrit ⁴¹ :

$$\oint_{\gamma} \sum_i p_i dq_i = \sum_j n_j(\gamma) J_j .$$

En résumé, les D variables d'action sont des constantes du mouvement, sont indépendantes et ont des crochets de Poisson deux à deux qui sont nuls : ces deux propriétés résultent clairement de ce qu'elles sont des fonctions des D constantes du mouvement C_i qui possèdent elles-mêmes ces propriétés. De la même façon, le choix des valeurs prises par les constantes du mouvement $\{C_i\}$ ou le choix des valeurs prises par les variables d'action $\{J_j\}$ sont des choix équivalents. Ils déterminent entièrement le tore invariant considéré, et fixent en particulier la valeur de l'énergie que l'on peut écrire :

$$\mathcal{H}(\{q_i\}, \{p_i\}) = \mathcal{E}(\{C_i\}) = \mathcal{E}(\{J_i\}) . \quad (54)$$

Mouvement à un degré de liberté Un exemple simple et important qui illustre notre propos est celui d'un mouvement à un degré de liberté. Dans ce cas, les $D = 1$ constantes du mouvement en involution sont simplement l'énergie $\mathcal{E} = \mathcal{H}(q, p)$, ce qui assure l'intégrabilité du problème. Le tore invariant se confond avec l'unique trajectoire compatible avec la valeur de l'énergie (*cf.* figure ??). Le cas prototype est celui du pendule pesant :

– Pour une énergie inférieure à la valeur critique $\mathcal{E}_c = 2mgl$, on a un mouvement de libration, la trajectoire est une courbe fermée clairement assimilable au cercle unité \mathbb{T}^1 . La variable d'action

⁴⁰ Notons que sur une sphère, à la différence d'un tore, il n'y a qu'une seule classe d'homotopie, puisque que tous les lacets sont homotopes à un point.

⁴¹ Il s'agit d'une identité globale, sans égalité terme à terme.

est visiblement l'aire encerclée par cette trajectoire. Elle s'écrit analytiquement :

$$J(E) = \frac{1}{2\pi} \int p_\theta d\theta = \frac{1}{2\pi} \int \frac{p_\theta^2}{ml^2} dt = \frac{T}{2\pi} \times 2\langle K \rangle ,$$

où T est la période d'oscillation et $\langle K \rangle$ est la valeur moyenne de l'énergie cinétique sur une période. Cette valeur moyenne, de même que la période, ne peuvent être exprimées qu'en termes de fonctions elliptiques, sauf dans la limite harmonique $|\theta| \ll 1$ où on a bien sûr $T/2\pi = \omega^{-1} = \sqrt{l/g}$, $\langle K \rangle = E/2$ et donc $J = E/\omega$.

- Pour $\mathcal{E} > \mathcal{E}_c =$, on a un mouvement de rotation, et la trajectoire se referme seulement parce que θ est un angle qui reprend les mêmes valeurs à chaque tour. La variable d'action est alors donnée par l'aire comprise entre la trajectoire et l'axe horizontal $p_\theta = 0$ divisée par 2π , ainsi que par l'expression formelle obtenue ci-dessus.

Cas des variables cycliques Un cas autre important est celui d'une variable cyclique, c'est-à-dire qui n'apparaît pas dans l'expression du Hamiltonien (c'est notamment le cas de l'angle polaire ϕ pour un mouvement plan à force centrale dans un plan). Soit donc q_1 cette variable, et p_1 l'impulsion conjuguée, qui fournit donc immédiatement l'une des constante du mouvement. Cela conduit à résoudre le problème à $D - 1$ degrés de liberté de hamiltonien $\mathcal{H}(q_2, \dots, q_D, p_2, \dots, p_D; \alpha_1)$, où la valeur α_1 prise par p_1 est un simple paramètre. Si ce nouveau problème est intégrable, le tore invariant total \mathbb{T}^d sera clairement le produit du tore invariant \mathbb{T}^{d-1} par le cercle \mathbb{T}^1 associé à la variable q_1 , et on pourra donc calculer la première variable d'action :

$$J_1 = \frac{1}{2\pi} \oint p_1 dq_1 = p_1 \oint \frac{dq_1}{2\pi} ,$$

où l'intégrale est une constante (qui pourra ultérieurement être réduite à l'unité).

Ce cas particulier des variables cycliques est de loin le plus simple, et il serait très satisfaisant de pouvoir considérer dans le cas général que l'expression (54) puisse être celle du nouveau hamiltonien dans une représentation où les variables d'action seraient les impulsions, et les variables canoniquement conjuguées seraient toutes des variables cycliques. C'est cette représentation que nous allons construire dans le chapitre § 14 ci-dessous.

D – Hamilton–Jacobi et variables actions-angles ★

13 Formalisme de Hamilton–Jacobi

Nous donnons ici un bref aperçu du formalisme de Hamilton–Jacobi, qui est une nouvelle approche de la mécanique analytique, et peut être parfois vue comme une méthode de résolution des équations de Hamilton, fondée sur l’exploitation systématique des transformations canoniques. Elle permettra en particulier de construire la représentation « actions-angles » des systèmes intégrables évoquée qualitativement dans la section précédente. Il s’agit d’une technique relativement avancée et dont l’intérêt ici réside essentiellement dans les rapprochements qu’elle permet avec la mécanique quantique.

13.1 Équation de Hamilton–Jacobi

Nous revenons à la *fonction* action $S(q, t)$, vue comme une fonction de la position (finale) et du temps, en principe obtenue par intégration sur les *trajectoires réelles*, que nous avons introduite dans la section § 9. Rappelons que la formule (27) s’écrit dans le point de vue de Hamilton, le long d’une trajectoire (cf. équation (49)) :

$$dS = \sum_i p_i dq_i - \mathcal{H}(q, p, t) dt .$$

Son caractère « fermé » résulte des équations de Hamilton, vérifiées par hypothèse puisqu’on se restreint ici aux trajectoires réelles. Il serait toutefois malaisé et inefficace de devoir préalablement résoudre les équations du mouvement pour connaître les trajectoires et calculer ensuite la fonction S . Si en revanche s’il était possible de calculer *a priori* la fonction S , ne serait-il pas plus aisé d’en déduire les trajectoires ?

On peut apporter une réponse à cette question en utilisant les équations (51) et (50), pour transformer l’expression différentielle (49) en une équation aux dérivées partielles, dite « de Hamilton–Jacobi » :

$$\frac{\partial S}{\partial t} + \mathcal{H}\left(q, \frac{\partial S}{\partial q}, t\right) = 0 , \quad (55)$$

ou en coordonnées cartésiennes pour une particule⁴² :

$$\frac{\partial S}{\partial t} + \mathcal{H}\left(\vec{r}, \frac{\partial S}{\partial \vec{r}}, t\right) = 0 .$$

Mais est-il bien raisonnable, en vue de résoudre un problème, de remplacer des équations différentielles (de Hamilton) par une équation aux dérivées partielles (même du premier ordre) ? Si la réponse est en général négative, il se trouve que c’est intéressant ici, comme nous allons l’expliquer dans ce qui suit. Le principal intérêt de cette approche est qu’elle permet d’obtenir d’un coup *toutes les trajectoires*. En outre, dans certaines circonstances, le surcoût est modique, car cette équation aux dérivées partielles peut être *résolue par quadratures*.

⁴² Rappelons que $\partial S / \partial \vec{r}$ dénote le gradient de S par rapport à \vec{r} , noté aussi $\vec{\nabla} S$.

13.2 Notion de solution complète

Ce dont nous avons besoin pour résoudre le problème de mécanique n'est pas la solution la plus générale générale de l'EDP, mais ce que l'on appelle une *solution complète*. Il s'agit d'une solution particulière qui prend la forme :

$$S = F(q_1, \dots, q_D, \alpha_1, \dots, \alpha_D, t)$$

où les $\{\alpha_j\}_{j=1\dots D}$ sont D constantes d'intégration indépendantes.

Une telle solution est alors formellement une fonction génératrice de la forme $G_2(\{q_j\}, \{P_j\}, t)$, si l'on peut identifier les constantes $\{\alpha_j\}_{j=1\dots D}$ aux nouvelles impulsions, et on peut alors écrire :

$$\alpha_j \equiv p_j = \frac{\partial F}{\partial q_j}(\{q_i\}, \{\alpha_i\}, t) \quad \text{et} \quad \beta_j \equiv Q_j = \frac{\partial F}{\partial \alpha_j}(\{q_i\}, \{\alpha_i\}, t).$$

La première de ces équations permet d'établir l'expression des constantes α_n en fonction des constantes du mouvement que sont les p_i et q_i à l'instant initial. La seconde fait apparaître les nouvelles coordonnées Q_i **qui sont aussi des constantes**, puisque la transformation canonique considérée annule le hamiltonien (*cf.* (52)). Ce sont donc D constantes d'intégration supplémentaires qui manquaient pour finir de spécifier une trajectoire donnée et que l'on peut alors – en principe – exprimer aussi en fonction des q_i et p_i .

13.3 Exemple de l'oscillateur harmonique

A titre d'illustration, nous appliquons la méthode de Hamilton–Jacobi au hamiltonien : $\mathcal{H} = p^2/2m + m\omega^2 q^2/2$ de l'oscillateur harmonique. L'équation de Hamilton Jacobi s'écrit :

$$\frac{\partial S}{\partial t} + \frac{1}{2m} \left(\frac{\partial S}{\partial q} \right)^2 + \frac{m\omega^2}{2} q^2 = 0,$$

et on cherche une solution de la forme $S(q, \mathcal{E}, t) = \tilde{S}(q, \mathcal{E}) - \mathcal{E}t$, où \tilde{S} vérifie l'EDO plus simple :

$$\frac{1}{2m} \left(\frac{\partial \tilde{S}}{\partial q} \right)^2 + \frac{m\omega^2}{2} q^2 = \mathcal{E},$$

dont la solution est, en introduisant l'amplitude $q_0 = \sqrt{2\mathcal{E}/m\omega^2}$:

$$\tilde{S}(q, \mathcal{E}) = \pm \int^q \sqrt{2m\mathcal{E} - m^2\omega^2 q^2} dq = \pm m\omega \int^q \sqrt{q_0^2 - q^2} dq \quad (56)$$

même si cette intégrale peut être obtenue analytiquement, ce n'est pas utile puisque c'est sa dérivée $Q = \beta$ qui nous intéresse (avec $\alpha = \mathcal{E}$), soit en dérivant sous le signe somme :

$$\beta = \frac{\partial(\tilde{S} - \alpha t)}{\partial \alpha} = \pm \frac{1}{\omega} \int^q \frac{dq}{\sqrt{q_0^2 - q^2}} - 1 = \pm \frac{1}{\omega} \arcsin(q/q_0) - t = cste \quad (57)$$

Cela donne la solution attendue : $q = \pm \sqrt{2\mathcal{E}/m\omega^2} \sin(\omega(t + \beta))$. Si l'on utilise alors la seconde équation définissant la transformation canonique $p = \partial S/\partial q = \partial \tilde{S}/\partial q$, on obtient bien sûr l'intégrand de l'expression de \tilde{S} ci-dessus, et moyennant l'expression de q , il vient $p = \pm \sqrt{2m\mathcal{E}} \cos(\omega(t + \beta))$.

13.4 Séparation des variables

Un moyen très efficace, bien que pas totalement général, pour obtenir une solution complète consiste à utiliser la méthode de séparation des variables, lorsqu'elle est possible.

13.4.1 Équation de H-J stationnaire

Considérons dans un premier temps que le hamiltonien \mathcal{H} est indépendant du temps. On peut alors chercher l'action S sous la forme :

$$S(q, t) = S_0(t) + \tilde{S}(q_1, \dots, q_D) ,$$

ce qui conduit à mettre l'équation de Hamilton–Jacobi sous la forme :

$$\left\{ \begin{array}{l} \frac{\partial S_0}{\partial t} = -\mathcal{E} = cste , \\ \mathcal{H}(q_1, \dots, q_D, \frac{\partial \tilde{S}}{\partial q_1}, \dots, \frac{\partial \tilde{S}}{\partial q_D}) = \mathcal{E} = cste , \end{array} \right. \quad (58a)$$

ce qui ramène le problème à une EDP indépendante du temps pour \tilde{S} , et à une intégration triviale donnant : $S_0(t) = -\mathcal{E}(t - t_0)$. La fonction \tilde{S} est dite « action réduite » ou « fonction caractéristique de Hamilton ».

13.4.2 Cas « générique »

Le cas est « générique » dans la mesure où l'on s'intéresse à une variable q_1 quelconque, mais il n'est en rien « général », au sens où il s'appliquerait systématiquement à tous les problèmes.

Il arrive que l'équation (55) ou (58a) puisse être mise sous la forme :

$$f\left(q_1, \frac{\partial S}{\partial q_1}\right) = g\left(q_2, \dots, q_D, \frac{\partial S}{\partial q_2}, \dots, \frac{\partial S}{\partial q_D}\right) ,$$

où q_1 apparaît dans le membre de gauche et pas dans celui de droite. Cela introduit naturellement une constante α_1 dite « constante de séparation » telle que :

$$f\left(q_1, \frac{\partial S}{\partial q_1}\right) = \alpha_1 , \quad (59a)$$

$$g\left(q_2, \dots, q_D, \frac{\partial S}{\partial q_2}, \dots, \frac{\partial S}{\partial q_D}\right) = -\alpha_1 , \quad (59b)$$

La solution de l'équation de Hamilton–Jacobi (stationnaire) peut alors être cherchée sous la forme $S = S_1(q_1) + S_2(q_2, \dots, q_D)$, où les deux fonctions S_1 et S_2 sont respectivement solution des équations plus simples (59a) et (59b).

Cette procédure s'applique de façon triviale aux variables cycliques éventuelles, disons q_1 , puisque qu'on peut alors écrire :

$$\frac{\partial S_1}{\partial q_1} = \alpha_1 = cste \quad \text{et} \quad \tilde{S} = \alpha_1 (q_1 - q_1^{(0)}) + S_2(q_2, \dots; q_D) .$$

Notons que l'équation (59a) contient le plus souvent une simple dépendance quadratique en $\partial S/\partial q_1$, qui permettra donc de déterminer $S_1(q_1)$ par une simple intégration : pour cette raison, un système totalement séparable sera dit « intégrable par quadratures ».

13.5 Application au potentiel central

À titre d'exemple fondamental, nous allons montrer que le problème du mouvement d'une particule dans un champ central est, en coordonnées sphériques, totalement séparable.

Le hamiltonien s'écrit :

$$\mathcal{H} = \frac{p_r^2}{2m} + \frac{1}{2mr^2} \left(p_\theta^2 + \frac{p_\varphi^2}{\sin^2 \theta} \right) + V(r) .$$

La variable φ étant visiblement une variable cyclique, on a $p_\varphi \equiv M = cste$, et l'équation (58a) se réduit à

$$\frac{1}{2m} \left(\frac{\partial S}{\partial r} \right)^2 + \frac{1}{2mr^2} \left(\left(\frac{\partial S}{\partial \theta} \right)^2 + \frac{M^2}{\sin^2 \theta} \right) + V(r) = \mathcal{E} . \quad (60)$$

Dans cette dernière équation, le terme de la seconde parenthèse est le seul où figure la variable θ . Celle-ci peut donc être séparée à son tour en posant :

$$\left(\frac{\partial S}{\partial \theta} \right)^2 + \frac{M^2}{\sin^2 \theta} = L^2 ,$$

ce qui ramène l'EDP (60) aux deux équations différentielles ordinaires :

$$\begin{cases} S'_\theta = \pm \sqrt{L^2 - M^2 / \sin^2 \theta} \\ S'_r = \pm \sqrt{2m(\mathcal{E} - V(r)) - L^2 / r^2} . \end{cases}$$

On vérifie aisément que les constantes de séparation sont ici la projection $M = \vec{\sigma} \cdot \vec{u}_z$ du moment cinétique $\vec{\sigma}$ sur l'axe Oz des coordonnées sphériques, et le module carré $L^2 = \vec{\sigma}^2$ du moment cinétique.

La solution complète de l'équation de Hamilton–Jacobi est alors :

$$S(r, \theta, \varphi; \mathcal{E}, L, M) = S_r(r; \mathcal{E}, L) + S_\theta(\theta; L, M) + S_\varphi(\varphi; M)$$

$$\text{où } S_\varphi(\varphi; M) = M \varphi$$

$$S_\theta(\theta; L, M) = \pm L \left(\sin \Theta \arcsin \left(\frac{\tan \Theta}{\tan \theta} \right) - \arcsin \left(\frac{\cos \theta}{\cos \Theta} \right) \right)$$

avec $\Theta = \arcsin(|M|/L)$, tandis que l'action S_r dépend bien sûr du potentiel $V(r)$ considéré.

14 Variables canoniques « action-angle »

Nous avons précédemment considéré la transformation canonique engendrée par l'action, conduisant à des variables canoniques indépendantes du temps. Pour les problèmes autonomes intégrables, il est plus avantageux de considérer la transformation canonique engendrée par l'action réduite \tilde{S} . Nous supposons dans un premier temps que le problème est *entièrement séparable*.

14.1 Transformation canonique engendrée par l'action réduite

On choisit de prendre la solution complète $\tilde{S}(q_1, \dots, q_D; \alpha_1, \dots, \alpha_D)$ comme fonction génératrice $G_3(q, P)$, où les constantes de séparation $\{\alpha_i\}$ sont les nouvelles impulsions $\{P_i\}$. Ces dernières sont nécessairement en involution en raison de l'équation (59a), qui montre que chaque constante

α_i peut s'écrire en fonction de p_i et q_i uniquement. On exhibe ainsi les D constantes du mouvement requises pour l'application du théorème de Liouville-Arnold. Comme la transformation canonique considérée est indépendante du temps, le nouveau hamiltonien est inchangé, à ceci près qu'il s'exprime maintenant en fonction des seules impulsions α_i . Les nouvelles positions Q_i vérifient alors :

$$Q_i = \frac{\partial G_3}{\partial P_i} = \frac{\partial \tilde{S}}{\partial \alpha_i} \quad \text{et} \quad \dot{Q}_i = \frac{\partial \mathcal{E}}{\partial \alpha_i} = \text{cste}$$

ce qui signifie que (i) elles sont cycliques et (ii) leur vitesse d'évolution est constante, soit $Q_i(t) = Q_i(t_0) + (t-t_0)\partial\mathcal{E}/\partial\alpha_i$. Cela n'est évidemment compatible avec l'hypothèse du mouvement borné dans l'espace des phases, que si les q_i et p_i sont des fonctions périodiques des Q_i .

Nous constatons ainsi que nous avons presque réalisé le programme que nous nous proposons à la fin du chapitre précédent. En particulier, le tore invariant \mathbb{T}^d est – dans le cas séparable – le produit des D cercles associés à chacune des coordonnées, et pour chacune la variable d'action est donnée par l'intégrale sur ce « cercle » :

$$J_i(\alpha_1, \dots, \alpha_D) = \frac{1}{2\pi} \oint p_i dq_i, \quad (62)$$

où le point important est l'absence de sommation sur i , contrairement au cas de l'équation (53). On constate que, au facteur 2π près, J_i est l'accroissement de l'action réduite sur une période de la variable correspondante q_i .

14.2 Interprétation géométrique

Cette équation (62), pour un cas à une dimension, ou pour un cas séparable, possède une interprétation géométrique simple : l'intégrale sur un circuit fermé de $p dq$ n'est autre que l'aire de ce circuit dans l'espace de phase de dimension 2 correspondant, c'est-à-dire l'intérieur des « trajectoires » fermées dessinées sur les portraits de phase de la figure 10.

14.3 Variables d'angle

Pour achever ce programme, pour un système intégrable même éventuellement non séparable, il suffit de reprendre la transformation canonique $G_3 = \tilde{S}$ précédente, mais en choisissant cette fois comme impulsions constantes les variables d'action. Les nouvelles positions obtenues sont alors les « variables d'angle », notées w_i , telles que :

$$w_i = \frac{\partial \tilde{S}}{\partial J_i}, \quad \dot{w}_i = \frac{\partial \mathcal{E}}{\partial J_i} = \text{Cste}$$

On peut alors revenir à la définition des variables d'action en considérant le lacet défini par la variation de w_i jusqu'à revenir au point de départ, les autres variables $w_{j \neq i}$ étant gardées constantes. On peut alors écrire :

$$J_i = \frac{1}{2\pi} \oint \sum_j p_j dq_j = \frac{1}{2\pi} \oint \sum_j J_j dw_j = J_i \oint \frac{dw_i}{2\pi}.$$

Cela montre que le lacet considéré est un lacet élémentaire défini précédemment et que la périodicité de w_i est de 2π , d'où son nom de « variable d'angle ». Ainsi les variables canoniquement conjuguées ($\{w_i\}, \{J_i\}$) constituent-elles la paramétrisation naturelle du tore invariant. En particulier le flot de \mathcal{H} s'écrit alors comme la combinaison (commutative) du flot de chacun

des J_i dont les paramètres w_i défilent linéairement en temps. Il en résulte que les q_i et les p_i sont des fonctions des fonctions 2π -périodiques des angles w_i et sont donc des fonctions « quasi-périodiques » ou « multi-périodiques » du temps, dont les fréquences angulaires sont celles des w_i soit $\omega_i = \partial\mathcal{E}/\partial J_i$.

Retour sur le cas harmonique 1D L'unique variable d'action J , donnée par l'incrément sur un tour de (56) est :

$$J = \frac{m\omega}{\pi} \int_{-q_0}^{q_0} \sqrt{q_0^2 - q^2} dq = \frac{m\omega q_0^2}{\pi} \int_{-1}^1 \sqrt{1 - x^2} dx = \frac{m\omega q_0^2}{\pi} \underbrace{\int_{-\pi/2}^{\pi/2} \cos^2(u) du}_{\pi/2} = \frac{\mathcal{E}}{\omega}$$

$$\text{et donc } \Omega = \frac{\partial\mathcal{E}}{\partial J} = \frac{\partial(J\omega)}{\partial J} = \omega. \quad (63)$$

Dans le même esprit, l'équation (57) anticipait déjà sur la variable d'angle $w = \omega(\beta + t) = \pm \arcsin(q/q_0)$ qui varie bien de π entre les deux points d'élongation maximale $q = \pm q_0$. Et a ainsi, à nouveau, compte tenu de (63) :

$$w = w_0 + \Omega t \quad \text{et donc} \quad q = \pm \sqrt{\frac{2J}{m\omega}} \sin(\omega t + w_0).$$

14.4 Évolution des variables d'action

La décomposition d'un problème de mécanique, grâce à l'approche de Hamilton–Jacobi, en ses variables action-angle, fournit — lorsqu'elle est possible — un nouvel outil puissant pour étudier l'évolution du système si le hamiltonien s'écarte un peu de sa valeur initiale. Deux cas particuliers peuvent être envisagés, qui ont en commun le fait que les variables d'action qui étaient constantes sont alors susceptibles de subir une évolution lente, dite « évolution séculaire »⁴³, à savoir le problème des perturbations et celui de l'évolution adiabatique. Malgré l'importance considérable historique, pratique et conceptuelle du problème des perturbations, nous n'étudierons ici que le problème de l'invariance adiabatique, en raison de son rôle dans la connexion entre la mécanique classique et la mécanique quantique.

Invariance adiabatique des variables d'action Le problème étudié ici est celui où le hamiltonien \mathcal{H} est susceptible de varier temporellement sous l'effet de la modification extérieure d'un paramètre λ . Un exemple simple est l'évolution d'un système physique dans un champ magnétique que l'on fait varier. Si la variation du paramètre est rapide, l'énergie va vraisemblablement dépendre du temps, et la description en termes de variables action-angle perd donc son intérêt.

Si en revanche la variation de λ est « adiabatique », c'est à dire lente à l'échelle des temps caractéristiques du mouvement (soit $\dot{\lambda}/\lambda \ll \Omega_i$), on peut considérer que le mouvement est à chaque instant très proche de celui qu'on aurait avec λ constant, mais que ses caractéristiques évoluent lentement du fait de la variation de λ . Étudions dans quelle mesure on peut continuer à utiliser les variables action-angle construites pour λ constant. Nous supposons pour ce faire que, pour toute valeur fixée de λ , il est possible de déterminer l'action réduite $\tilde{S}(q, J, \lambda)$, et de construire les variables action-angle $w(\lambda)$ et $J(\lambda)$.

⁴³ En raison de l'application de ces techniques aux évolutions à long terme des paramètres des orbites planétaires.

Considérons alors la transformation canonique dépendant du temps engendrée par la fonction :

$$S^\dagger(q, J, t) = \tilde{S}(q, J, \lambda(t)) ,$$

où les nouvelles impulsions sont à chaque instant les variables d'action que l'on aurait à λ fixé. Nous noterons $v_i = \partial S^\dagger / \partial J_i$ les variables correspondantes, qui *ne sont pas* exactement des variables d'angle, puisque S^\dagger ne vérifie pas l'équation de Hamilton–Jacobi si $\dot{\lambda} \neq 0$. Comme la transformation utilisée dépend explicitement du temps, le nouveau hamiltonien K diffère de l'ancien et s'écrit :

$$K(v, J, t) = E(J, \lambda(t)) + \Lambda(v, J, \lambda(t)) \cdot \dot{\lambda} \quad (64)$$

où l'on a posé :

$$\Lambda(v, J, \lambda) = \left. \frac{\partial \tilde{S}}{\partial \lambda} \right|_{q, J} (q(v, J, \lambda), J, \lambda) .$$

Il est essentiel de noter que, si \tilde{S} est une fonction multivaluée de la position, en ce sens qu'elle se voit augmentée de $2\pi J_k$, chaque fois que l'angle w_k a varié de 2π , la fonction Λ , obtenue par dérivation, est elle une « vraie fonction » 2π -périodique des v_k .

La première équation de Hamilton tirée de (64) montre que les v_k évoluent à peu près comme les w_k à λ fixé, c'est à dire que ce sont toujours des variables rapides à l'échelle de temps considérée. La seconde équation de Hamilton s'écrit quant à elle :

$$\dot{J} = - \frac{\partial K}{\partial v} = - \frac{\partial \Lambda}{\partial v} \cdot \dot{\lambda} ,$$

et décrit *a priori* une variation lente des « variables d'action », en accord l'image qualitative que que l'on se fait du mouvement.

En raison de la variation lente de λ , J va rester constant au moins sur quelques périodes, et on peut donc moyenner l'équation ci-dessus sur un temps long devant la (ou les) période(s) du mouvement, mais suffisamment court pour pouvoir négliger la variation de λ et de $\dot{\lambda}$; on obtient ainsi :

$$\dot{J} = - \left\langle \frac{\partial \Lambda}{\partial v} \right\rangle \cdot \dot{\lambda} \simeq 0 .$$

En effet la valeur moyenne dans le membre de droite peut être évaluée à λ (et donc J) constant, et la valeur moyenne de la dérivée $\partial \Lambda / \partial v$ sur un grand nombre de périodes de v est nulle. On obtient en définitive le résultat $J = cste$.

Ce résultat constitue une propriété très importante des variables d'action, qui fut discutée à l'initiative de Ehrenfest au tout début du XX^e siècle, mais ne fut réellement établie qu'en 1916, dans le contexte de la première théorie des quanta.

15 Liens avec la mécanique quantique

Il n'aura pas échappé au lecteur de nombreuses « ressemblances » entre certains résultats obtenus ici et leurs analogues en mécanique quantique, que nous allons préciser ici. Nous nous intéresserons successivement aux points suivants :

1. la quantification canonique ;
2. les crochets de Poisson et les commutateurs ;
3. les transformations canoniques et les transformations unitaires ;
4. l'équation de Hamilton–Jacobi, et le sens physique du principe de Hamilton ;
5. les variables action-angle et la quantification semi-classique.

15.1 Quantification canonique

C'est un postulat souvent non explicité dans les cours de MQ, car souvent considéré comme purement historique. Il stipule que les quantités physiques de la mécanique quantiques, qui sont des opérateurs, se déduisent de leurs homologues classiques grâce à la substitution de \vec{p} par \vec{P} et de \vec{r} par \vec{R} . Cette substitution doit en outre être suivie d'une symétrisation afin que l'opérateur obtenu soit hermitique. Ainsi en est-il en particulier du hamiltonien, avec les cas où \vec{p} n'est pas la quantité de mouvement, et donc \vec{P}/m n'est pas la vitesse. Ainsi pour le moment cinétique $\vec{L} = \vec{R} \times \vec{P} = -\vec{P} \times \vec{R}$ mais $\vec{A} = (\vec{L} \times \vec{P} - \vec{L} \times \vec{P})/2mK - \vec{R}/\|\vec{R}\|$ pour le vecteur de Laplace (ou de Runge-Lenz).

15.2 Crochets de Poisson

Le crochet de Poisson et le commutateur quantique sont deux opérations bilinéaires anti-symétriques pour les grandeurs physiques qui possèdent nombre de propriétés mathématiques communes. Les opérateurs fondamentaux de la MQ vérifient des relation de commutation canoniques $[R_i, P_j] = i\hbar \delta_{ij}$ qui sont les analogues des crochets de Poisson canoniques pour $\{r_i, p_j\} = \delta$, en utilisant la correspondance $\{*, *\} \leftrightarrow [*, *]/i\hbar$. Cette règle s'applique à toutes les quantités physiques construites sur \vec{r} et \vec{p} et transposées dans le domaine quantique par quantification canonique. Cela peut être obtenu par le calcul en utilisant les similarités des deux opérations, mais reflète plus fondamentalement la structure des algèbres de Lie sous-jacentes. Il résulte bien sûr les règles similaires sur l'évolution temporelle et l'invariance au cours du mouvement. Plus généralement, les règles d'involution des grandeurs physiques compatibles (c'est-à-dire dont les flots commutent) sont en rapport avec les règles de commutation des opérateurs représentant des grandeurs physique compatibles, et il existe une forte analogie entre le théorème de Liouville-Arnold et la construction d'un « ECOC » en MQ.

15.3 Transformations canoniques

La grande flexibilité dans la construction des TC pour représenter un changement de point de vue ou pour simplifier les équations du mouvement trouve son analogue quantique dans les transformations unitaires. En particulier, les opérateurs exponentiels construits pour représenter les translations ou l'évolution temporelle possèdent des analogues exacts en MQ; les règles de modification du hamiltonien quantique lors d'une transformation unitaire dépendant du temps sont d'ailleurs très analogues à celles étudiées en mécanique classique. La transformation canonique engendrée par l'action qui rend les variables indépendantes du temps est en étroite connexion avec la transformation unitaire qui permet de passer du point de vue de Heisenberg au point de vue de Schrödinger.

Les inégalités de Heisenberg n'ont bien sûr pas d'équivalent classique, mais dans la mesure où elles expriment qu'un état quantique « occupe » un volume $(\hbar/2)^D$ dans l'espace des phase, elles ne seraient pas consistantes sans le théorème de Liouville.

15.4 Équation de Hamilton-Jacobi

Elle joue un rôle dans la « limite des courtes longueurs d'onde », c'est-à-dire lorsque la longueur d'onde de de Broglie est courte devant les échelle de variation de la fonction d'onde (ou du potentiel).

Soit une particule quantique de masse m , et évoluant dans un potentiel $V(\vec{r}, t)$. Son hamiltonien quantique s'écrit :

$$H = \frac{\vec{p}^2}{2m} + V(\vec{R}, t)$$

et son évolution dans la représentation de position (dans le point de vue de Schrödinger) :

$$i\hbar \frac{\partial \psi}{\partial t}(\vec{r}, t) = -\frac{\hbar^2}{2m} \Delta \psi(\vec{r}, t) + V(\vec{r}, t).$$

Nous cherchons une solution sous la forme générale :

$$\psi(\vec{r}, t) = A(\vec{r}, t) \exp(i\Phi(\vec{r}, t)),$$

où A est l'amplitude, et Φ la phase de la fonction d'onde, dont nous supposons qu'elles ont des échelles de variation très différentes, A variant lentement par rapport à Φ . En reportant cette expression dans l'équation de Schrödinger et en isolant la partie réelle, on obtient l'équation :

$$\hbar \frac{\partial \Phi}{\partial t} + \frac{\hbar^2}{2m} \left(\frac{\partial \Phi}{\partial \vec{r}} \right)^2 + V(\vec{r}, t) = \frac{\hbar^2}{2m} \frac{\Delta A}{A}.$$

Par hypothèse, le second membre est, en ordre de grandeur, négligeable devant les termes du premier⁴⁴. La variation rapide de la phase peut être prise en compte plus clairement en posant $\Phi = S/\hbar$, où S est maintenant à variation lente puisque $1/\hbar$ est très grand, on obtient :

$$\frac{\partial S}{\partial t} + \frac{1}{2m} \left(\frac{\partial S}{\partial \vec{r}} \right)^2 + V(\vec{r}, t) = \frac{\hbar^2}{2m} \frac{\Delta A}{A}$$

qui redonne exactement l'équation de Hamilton–Jacobi si le second terme est effectivement négligeable⁴⁵. L'équation de Hamilton–Jacobi est donc la forme asymptotique de l'équation de Schrödinger à la limite des courtes longueurs d'onde, de même que, en optique, l'équation eikonale $(\nabla S(\vec{r}))^2 = n^2(\vec{r})$ est la limite de l'équation de d'Alembert pour les courtes longueurs d'onde.

Dans ce contexte, on comprend le rôle joué en mécanique par l'action et en optique par le chemin optique, et le principe de Hamilton et de Fermat apparaissent simplement comme des principes de phase stationnaire, c'est à dire d'interférence constructive, qui tend à localiser l'onde au voisinage de la trajectoire « classique ». C'est sur ce principe qu'est fondée la formulation de la MQ en « intégrales de chemin » développée par Feynman, où les propagateurs quantiques sont calculés en sommant sur tous les chemins γ possibles les amplitudes élémentaires données par $\exp(iS(\gamma)/\hbar)$.

15.5 Variables action-angle

La « première théorie des quanta » développée par Bohr, Wilson, Sommerfeld (et quelques autres, dont Schrödinger) dans les années 1913-1925 postulait la quantification des variables d'action en multiples entiers de \hbar (quantification BWS). Elle fut étendue et reformulée par Einstein, Brillouin et Keller à partir de 1917 (quantification EBK). Ces hypothèses sont aussi

⁴⁴ Si \mathcal{L} est l'échelle de variation de A et λ_{dB} celle de Φ , ce second membre est de l'ordre de $\hbar^2/2m\mathcal{L}^2$ alors que le premier est de l'ordre de $\hbar^2/2m\lambda_{dB}^2$.

⁴⁵ Cette approximation cesse d'être valide au voisinage des caustiques (c'est-à-dire pour un problème séparable au niveau des points de rebroussement) où λ_{dB} devient localement grand, ou lorsque V varie trop vite pour qu'un λ_{dB} « local » garde un sens.

retrouvées par la résolution de l'équation de Schrödinger à la limite semi-classique (quantification WKBJ), au moins dans le cas séparable.

De façon plus qualitative, mais bien plus simple, cette règle de quantification des variables d'action prend du sens si on se réfère à l'interprétation géométrique de celles-ci introduite au § 14.2. Si on considère que, en raison des inégalités de Heisenberg, un état quantique occupe une cellule de taille h dans l'espace des phase, le nombre quantique entier est le nombre d'état contenus dans la trajectoire considérée.

Plusieurs observations importantes doivent être faites à cet égard :

- Tout d'abord, en imposant aux variables d'action des valeurs bien définies, on conçoit que les inégalités de Heisenberg conduisent à une indétermination complète des variables conjuguées, à savoir les angles. Ceci est donc intimement lié aux propriétés de délocalisation de la fonction d'onde, comme la position de l'électron sur son « orbite » circulaire.
- En second lieu, l'expression classique des variables q ou p en fonctions des variables d'angles conduisent à un comportement multipériodique à des fréquences $\omega_i = \partial\mathcal{E}/\partial J_i$. En mécanique quantique, même si les valeurs instantanées des angles ne sont pas définies, leur évolution temporelle subsiste et se manifeste par les « fréquences de Bohr » qui peuvent s'écrire sous la forme très analogue $\Omega_i = 1/\hbar \Delta\mathcal{E}/\Delta n = \Delta\mathcal{E}/\Delta J$.
- Enfin, la quantification par un entier de certaines grandeurs physique pose *a priori* un problème conceptuel si on envisage une évolution lente du hamiltonien sous l'effet d'un paramètre extérieur. Comment rendre compte alors d'une évolution *continue* des grandeurs quantifiées ? C'est ici qu'intervient le théorème d'invariance adiabatique des variables d'action : comme celles-ci n'évoluent pas pour une évolution assez lente, on conçoit qu'elles puissent prendre des valeur constantes discrètes. Par contre, pour des variations plus rapide, on n'a plus conservation des variables d'action, et la MQ nous prédit l'observation de « sauts quantiques ». Mais ceci est une autre histoire...

Références

- [1] Lev Landau and Evguéni Lifchitz. *Physique Théorique – Mécanique*. Éditions MIR / Ellipses, 5^e édition (1994).
- [2] Vladimir I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer-Verlag (1989).
- [3] Richard P. Feynman and Robert Leighton avd Matthew Sands. *Le cours de physique de Feynman : Electromagnétisme 1*. InterEditions (1997). chap. 19.
- [4] Herbert Goldstein. *Classical Mechanics*. Addison Wesley, 2nd édition (1980).
- [5] Claude Gignoux and Bernard Silvestre-Brac. *Mécanique : de la formulation lagrangienne au chaos hamiltonien*. EDP Sciences (2002).
- [6] Jean-Michel Raimond. “Electromagnétisme et relativité”. HAL ID cel-00092954 (2006).
- [7] Jean-Louis Basdevant. *Les principes variationnels en physique*. Vuibert (2014).